Identification of an antiviral component from the venom of the scorpion Liocheles australasiae using transcriptomic and mass spectrometric analyses

AUTHOR(S):

Miyashita, Masahiro; Mitani, Naoya; Kitanaka, Atsushi; Yakio, Mao; Chen, Ming; Nishimoto, Sachiko; Uchiyama, Hironobu; ... Hotta, Hak; Nakagawa, Yoshiaki; Miyagawa, Hisashi

CITATION:

Miyashita, Masahiro ...[et al]. Identification of an antiviral component from the venom of the scorpion Liocheles australasiae using transcriptomic and mass spectrometric analyses. Toxicon 2021, 191: 25-37

ISSUE DATE:

2021-02

URL:

http://hdl.handle.net/2433/266820

Identification of an antiviral component from the venom of the scorpion *Liocheles australasiae* using transcriptomic and mass spectrometric analyses

Masahiro Miyashita[1]*, Naoya Mitani[1], Atsushi Kitanaka[1], Mao Yakio[1], Ming Chen[2], Sachiko Nishimoto[3], Hironobu Uchiyama[4], Masayuki Sue[5], Hak Hotta[2,3], Yoshiaki Nakagawa[1], and Hisashi Miyagawa[1]

[1]Graduate School of Agriculture, Kyoto University, Kyoto 606-8502, Japan

[2]Graduate School of Health Sciences, Kobe University, Kobe 650-0047, Japan

[3]Faculty of Clinical Nutrition and Dietetics, Konan Women's University, Kobe 658-0001, Japan

[4]NODAI Genome Research Center, Tokyo University of Agriculture, Tokyo 156-8502, Japan

[5]Department of Agricultural Chemistry, Tokyo University of Agriculture, Tokyo 156-8502, Japan

*Corresponding author: e-mail; miyashita.masahiro.6e@kyoto-u.ac.jp

Postal address; Kitashirakawa Oiwakecho Sakyo-ku, Kyoto 606-8502, Japan

Abstract

Scorpion venom contains a variety of biologically active peptides. Among them, neurotoxins are major components in the venom, but it also contains peptides that show antimicrobial activity. Previously, we identified three insecticidal peptides from the venom of the *Liocheles australasiae* scorpion, but activities and structures of other venom components remained unknown. In this study, we performed a transcriptome analysis of the venom gland of the scorpion *L. australasiae* to gain a comprehensive understanding of its venom components. The result shows that potassium channel toxin-like peptides were the most diverse, whereas only a limited number of sodium channel toxin-like peptides were observed. In addition to these neurotoxin-like peptides, many non-

disulfide-bridged peptides were identified, suggesting that these components have some critical roles in the *L. australasiae* venom. In this study, we also isolated a component with antiviral activity against hepatitis C virus using a bioassay-guided fractionation approach. By integrating mass spectrometric and transcriptomic data, we successfully identified LaPLA$_2$-1 as an anti-HCV component. LaPLA$_2$-1 is a phospholipase A$_2$ having a heterodimeric structure that is N-glycosylated at the N-terminal region. Since the antiviral activity of LaPLA$_2$-1 was inhibited by a PLA$_2$ inhibitor, the enzymatic activity of LaPLA$_2$-1 is likely to be involved in its antiviral activity.

Keywords

bioactive peptide; glycosylation; hepatitis C; phospholipase; venom gland

1. Introduction

Scorpions use their venom to capture their prey and defend themselves against predators. For this purpose, scorpion venom is composed of a variety of biologically active peptides (Ahmadi et al., 2020). Among them, neurotoxins are the most effective components in capturing their prey, such as insects. These toxins can immediately stop the movement of prey by acting on their ion channels (Schwartz et al., 2012, Smith, J. J. et al., 2013). In addition, these neurotoxins often show selective toxicity between insects and mammals, which is conferred at the ion channel level (Gordon et al., 2007, Housley et al., 2017).

Scorpion toxins acting on ion channels are peptides cross-linked with multiple disulfide bonds (Quintero-Hernandez et al., 2013). These peptides can be classified into two groups based on their molecular size. The long-chain group consists of peptides with 60-80 amino acid residues, which include toxins that act on Na^+ channels. The short-chain group consists of peptides with 20-50 amino acid residues, which include toxins that act on K^+, Ca^{2+}, and Cl^- channels. The scorpion venom also contains many peptides without disulfide bonds. Most of these peptides show antimicrobial activity (Almaaytah and Albalas, 2014), some of which also exhibit insecticidal activity and/or synergistically enhance the activity of other toxins (Miyashita et al., 2010, Wullschleger et al., 2005). In addition to peptides, enzyme proteins are present in the scorpion venom, although their contribution to the biological activity of the venom remains unclear (Smith, J.J. and Alewood, 2015). Because of these diverse components, scorpion venom has been studied as a rich source of bioactive molecules, which provide useful information for the development of novel pesticides and drugs (Ghosh et al., 2019, Smith, J. J., et al., 2013). Scorpion toxins can also serve as an important probe for the study of ion channels (Herzig et al., 2020, Zhao, Y. et al., 2019).

Traditionally, the search for bioactive peptides/proteins has been conducted mainly relying on the bioassay-guided isolation procedure, in which high performance liquid chromatography (HPLC) fractionation and bioassay is repeated until a single active

component is obtained (Vetter et al., 2011). This approach has led to the discovery of a number of novel bioactive molecules with unique structures. However, it is often difficult to identify minor components in the venom using this approach. In such cases, cDNAs obtained by reverse transcription of the mRNAs in the venom gland have been amplified by polymerase chain reaction (PCR) using a primer constructed based on a partial sequence to determine the entire structure (Quintero-Hernandez et al., 2011). Recently, the advent of next-generation high-throughput sequencing technologies allows us to obtain sequences of all mRNAs expressed in the venom gland (transcriptome), which provides a comprehensive understanding of structures of venom components (Oldrati et al., 2016). Furthermore, the combination of transcriptome analysis with proteome analysis has accelerated the determination of mature structures of each component in the venom (Fu et al., 2018, Walker et al., 2020). Although bioactive components can be estimated based on their structurally similarity with reported molecules using this approach, biological functions of many components remain unclear due to the absence of similar molecules reported. In this regard, the classical bioassay-guided approach is still effective for the discovery of bioactive components with unique structural characteristics, and its combination with transcriptome analysis should accelerate the structural determination.

It is known that infection with hepatitis C virus (HCV) often causes liver disease, including cirrhosis and hepatocellular carcinoma. Although highly effective direct-acting antivirals can cure the vast majority of HCV infections, in the absence of a vaccine, there is a continued demand for antiviral drugs against HCV. Of the various enemies and pathogens scorpions cope with, virus represents one of major threats, and actually the scorpion venom contains antiviral components to prevent infection via venom glands (da Mata et al., 2017, El-Bitar et al., 2015, Yacoub et al., 2020). In this context, the scorpion venom has been studied as one of the promising sources of antiviral molecules. Previously, we identified three insecticidal peptides from the venom of the *Liocheles australasiae* scorpion based on toxicity against insects (Juichi et al., 2019, Matsushita et al., 2009,

Matsushita et al., 2007). However, the venom has not been evaluated for other biological activities, including antiviral effects. Thus, in this study, we first performed a transcriptome analysis of the venom gland of *L. australasiae* to gain a comprehensive understanding of its venom components. Using this information coupled with the bioassay-guided approach, we identified a new component with antiviral activity against HCV from the venom of *L. australasiae*.

2. Materials and methods

2.1. Biological materials

The scorpions *L. australasiae* were collected in Ishigaki Island, located at the southern end of the Ryukyu Islands in Japan. They were reared in the laboratory under humid conditions at 25°C and fed crickets. The venom was obtained by mechanical stimulation as previously reported (Miyashita et al., 2007). The venom secreted on Parafilm was dissolved in aqueous 2% acetic acid and filtered, which was lyophilized and stored at –80°C.

2.2. RNA extraction and sequencing

The telsons were dissected from six specimens anesthetized on ice and placed in a glass tube of a micro tissue grinder. The total RNA was extracted using RNAiso Plus (Takara Bio, Kusatsu, Japan) and further purified using the RNeasy Mini Kit (Qiagen, Venlo, The Netherlands) according to the manufacturer's instructions. The integrity of the RNA was verified using a Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA). The mRNA was isolated using the NEBNext Poly(A) mRNA Magnetic Isolation Module (New England Biolabs, Ipswich, MA, USA). A cDNA library was prepared from the purified mRNA using the NEBNext Ultra RNA Library Prep Kit for Illumina (New England Biolabs). The 100 bp paired-end sequencing was performed on an Illumina HiSeq2500 platform (San Diego, CA, USA). The short-read data were deposited to the Read Archive of DDBJ (accession number DRA010798).

128

129 *2.3. De novo assembly and functional annotations*

130 After adapter- and quality-trimming using the software TagDust (Lassmann et al.,
131 2009) and Fastx-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit), the cleaned reads were
132 assembled into contigs with Trinity (2.06) (Haas et al., 2013) and Bridger (r2014-12-01)
133 (Chang et al., 2015) software using the standard protocol. Coding regions were predicted
134 by TransDecoder (2.0). The predicted sequences obtained using two different assembly
135 software were separately submitted to similarity searches using the BLASTP program
136 against a database containing only sequences identified from scorpion venom (the
137 UniProt Animal Toxin Annotation Project) to annotate the functions of identified peptides
138 and proteins through the local BLAST tool (Altschul et al., 1990). Coding regions of each
139 component were further inspected and corrected manually by comparing them with the
140 reported sequences. Multiple sequence alignments were performed using MAFFT online
141 (Katoh et al., 2019) and Clustal Omega (Madeira et al., 2019).

142

143 *2.4. HPLC purification*

144 The crude venom dissolved in distilled water was applied to a C4 semi-preparative
145 column (10 × 250 mm, Grace Vydac, Deerfield, IL, USA). The column was eluted with
146 0.1% trifluoro acetic acid (TFA) in water (solvent A) and 0.08% TFA in acetonitrile
147 (solvent B) at a flow rate of 2 mL/min, using a linear gradient of 5–60% of solvent B over
148 55 min. Elution was monitored by the UV absorbance at 215 nm. Fractions were collected
149 every 5 min during gradient elution. Each fraction was submitted to the antiviral activity
150 test after lyophilization, and the fraction showing activity was then applied to a C18
151 microbore column (1.0 × 250 mm, Grace Vydac). The column was eluted with solvent A
152 and B at a flow rate of 0.05 mL/min, using a linear gradient of 20–50% of solvent B over
153 60 min. Each HPLC peak was individually collected and lyophilized. The most active
154 fraction was further purified on the same C18 microbore column using a different solvent
155 system. The column was eluted with 0.1% formic acid in water (solvent C) and 0.1%

formic acid in acetonitrile (solvent D) at a flow rate of 0.05 mL/min using a linear gradient of 5–55% solvent D over 50 min. The purity was checked by liquid chromatography/mass spectrometry (LC/MS) analysis as described below.

2.5. Mass spectrometric analysis

LC/MS and LC/MS/MS measurements were carried out in a positive ion mode on an LCMS-IT-TOF mass spectrometer (Shimadzu, Kyoto, Japan) equipped with an electrospray ion source. Precursor ions were manually selected, and a collision-induced dissociation (CID) spectrum was obtained by using argon as a collision gas. Reversed-phase (RP)-HPLC separation was performed on a C18 microbore column (TSKgel ODS-100V 3 μm, 1.0 × 150 mm, Tosoh, Tokyo, Japan). The column was eluted with solvent C and solvent D at a flow rate of 0.05 mL/min, using a linear gradient of 5–70% of solvent D over 65 min. The mass scale was calibrated externally using sodium trifluoroacetate cluster ions.

2.6. Enzymatic digestion

The protein was dissolved in a buffer containing 0.2 M Tris-HCl (pH 8.5), 6 M guanidine hydrochloride and 10 mM dithiothreitol (DTT), and the mixture was incubated at 50ºC for 1 h. The reaction mixture was then mixed with iodoacetic acid (20 mM final concentration) and incubated at 28ºC for 1 h. The solution was diluted three-fold with water, and the Cys-alkylated protein was digested with endoproteinase Lys-C (FUJIFILM Wako Pure Chemical Corporation, Osaka, Japan) or Glu-C (FUJIFILM Wako Pure Chemical Corporation) at 37 ºC for 18 h at a peptide/enzyme ratio of 20:1 (w/w).

2.7. Deglycosylation

The protein was dissolved in a solution containing 0.5% SDS and 40 mM DTT, and the mixture was incubated at 100ºC for 10 min. The solution was diluted two-fold with a buffer containing 50 mM sodium acetate and 1% NP-20, then mixed with 5 units of

PNGase A (New England Biolabs, Ipswich, MA, USA) at 37°C for 1 h.

2.8. Cell culture and viruses

Huh7it-1 cells were cultivated in Dulbecco's modified Eagle's medium (FUJIFILM Wako Pure Chemical Corporation, Osaka, Japan) supplemented with fetal bovine serum (Biowest, Nuaille, France), non-essential amino acids (Thermo Fisher Scientific, Waltham, MA, USA), penicillin (100 IU/ml) and streptomycin (100 μg/ml) (Thermo Fisher Scientific). Cells were grown at 37°C in a 5% CO_2 incubator. The J6/JFH1-P47 strain of HCV (El-Bitar, et al., 2015), the Trinidad 1751 strain of dengue virus (DENV) (El-Bitar, et al., 2015), the Nakayama strain of Japanese encephalitis virus (JEV) (Song et al., 1999), and the CHR3 strain of herpes simplex virus type 1 (HSV-1) (Aoki-Utsubo et al., 2018) prepared in Huh7it-1 cells were used in this study. Infectivity of the stock virus was 1.5×10^6, 9.0×10^5, 5.0×10^5, and 2.0×10^4 cell-infecting units (CIU)/ml for HCV, DENV, JEV and HSV-1, respectively.

2.9. Antiviral activity test

Antiviral activity was tested according to two different experimental procedures as described below.

(i) Pretreatment of virus with venom components before and during virus inoculation (pretreatment). This procedure is used to determine virucidal (neutralizing) activity of a test sample (Aoki-Utsubo, et al., 2018, El-Bitar, et al., 2015). In brief, Huh7it-1 cells were seeded in 24-well plates (1.6×10^5 cells/well). A fixed amount of the virus (either 10-fold diluted or undiluted stock virus) was mixed with serial dilutions of the whole venom or separated HPLC fractions of the venom and inoculated to the cells for 2 h. The cells were then washed with medium to remove the residual virus and venom components, and further cultured in medium. Culture supernatants were obtained at 24–48 h post-infection and titrated for virus infectivity as described previously (El-Bitar, et al., 2015). After 24 h, the virus-infected cells were washed with phosphate-buffered saline (PBS), fixed with

4% paraformaldehyde for 20 min, and permeabilized with 0.1% Triton X-100 in PBS for 15 min at room temperature. After being washed three times with PBS, the cells were incubated with UV-inactivated HCV-infected patient's serum for 1 h, followed by incubation with FITC-conjugated goat anti-human IgG (Medical & Biological Laboratories Co., Ltd., Nagoya, Japan). The cells were counterstained with Hoechst 33342 (Molecular Probes, Eugene, OR, USA) for 5 min, and HCV-infected cells were counted under a BX53LED-43FLD fluorescence microscope (Olympus Corporation, Tokyo, Japan). Virus and cells treated with medium served as controls. Percent inhibition of virus infectivity by the samples was calculated by comparing with the controls, and 50% inhibitory concentrations (IC_{50}) were determined. Experiments were performed in duplicate and repeated three times.

(ii) Treatment of virus-infected cells after virus has been entered the cells (post-entry treatment). This procedure is used to determine viral replication-inhibiting activity of a test sample in the infected cells (Aoki-Utsubo, et al., 2018, El-Bitar, et al., 2015) In brief, cells were inoculated with virus in the absence of the venom components for 2 h. The virus-infected cells were cultured in medium containing serial dilutions of the venom components. The virus-infected cells cultured in medium without the venom components served as a control. After 48 h, RNA was extracted from the cells and subjected to reverse transcription-quantitative PCR (RT-qPCR) as described below. Experiments were performed in duplicate and repeated three times.

2.10. Reverse transcription-quantitative PCR (RT-qPCR)

Total cellular RNA was extracted from the cells using NucleoSpin RNA extraction kit (TaKaRa Bio, Inc.) according to the manufacturer's instructions. RNA (1 µg) was reverse transcribed using a GoScript Reverse Transcription system (Promega) with random primers. The cDNA products were subjected to quantitative real-time PCR analysis using Power SYBR Green PCR Master kit (Thermo Fisher Scientific Corp.) and a StepOne qPCR system (Thermo Fisher). The primers used to amplify an NS5A region of the HCV

genome were 5'-AGACGTATTGAGGTCCATGC-3' (sense) and 5'-CCGCAGCGACGGTGCTGATAG-3' (antisense) (Deng, L. et al., 2011). As an internal control, human glyceraldehyde-3-phosphate dehydrogenase (GAPDH) gene expression levels were measured using primers 5'-GCCATCAATGACCCCTTCATT-3' (sense) and 5' TCTCGCTCCTGGAAGATGG-3'. Relative quantity of HCV NS5A cDNA was calculated and expressed as an arbitrary unit for each sample.

2.11. Cytotoxicity test

Cytotoxicity of $LaPLA_2$-1 was evaluated using the WST-1 reagent (Roche, Mannheim, Germany) as reported previously (El-Bitar, et al., 2015). Briefly, Huh7it-1 cells plated in each well of a 96-well plate (0.1 ml/well) were treated with serial dilutions (0.1 to 1,000 ng/ml) of $LaPLA_2$-1 at 37 °C for 24 h. Untreated cells served as a control. After this treatment, 10 µl of the WST-1 reagent was added to each well, and the cells were cultured for 4 h. The WST-1 reagent is converted to formazan by living cells. The amount of formazan, which correlates with the number of living cells, was determined by measuring the absorbance at 450 and 630 nm using a microplate reader. The percent cell viability compared to the untreated control was calculated, and the 50% cytotoxic concentration (CC_{50}) was determined.

2.12. Phospholipase A_2 activity test

L-α-Phosphatidylcholine (from egg yolk, Nacalai Tesque, Kyoto, Japan) was suspended in a buffer containing 0.1 M Tris-HCl (pH 8.0), 10 mM $CaCl_2$, and 0.2% TritonX-100 at a concentration of 2 mg/ml. The reaction was started with the addition of the PLA_2 solution to the buffer (total 100 µl) at 37°C. After incubation for 10 min at 37°C, the mixture was analyzed by LC/MS to quantitate the amount of 1-palmitoyl-*sn*-glycero-3-phosphocholine (P-lysoPC, *m/z* 540.3, $[M+HCOOH-H]^-$) generated by the reaction. Conditions used for LC/MS analysis (LCMS-8030, Shimadzu) were as follows: ionization, ESI-negative; column, COSMOCORE 2.6C_{18} (2.1×75 mm, Nacalai Tesque);

flow rate, 0.3 mL/min; mobile phase, solvent C and solvent D; gradient, 35–95% of solvent D over 30 min. The amount of P-lysoPC generated in the absence of phospholipase A_2 (PLA_2) was subtracted from those generated in each experiment. Experiments were repeated three times. PLA_2 activity was expressed relative to that of PLA_2 from honeybee (*Apis mellifera*) venom (Sigma-Aldrich, St. Louis, MO, USA).

3. Results

3.1. Transcriptome analysis

A total of 253,691,496 raw reads were obtained by sequencing. After adapter- and quality-trimming, the clean reads were assembled in a de novo fashion using Trinity, which resulted in 133,715 contigs with N50 of 849 bp (224–14,739 bp), and Bridger software, which resulted in 98,577 contigs with N50 of 1447 bp (201–21,016 bp). BLASTP search was performed for predicted coding regions against a database consisting of reported sequences identified from scorpion venom. This resulted in identification of 77 transcripts encoding peptides and proteins similar to those of scorpion venom as shown in Fig. 1-5 and Table S1 and S2.

3.2. Components identified by venom gland transcriptome analysis

3.2.1. Non-disulfide bridged peptides

Scorpion venoms are rich in non-disulfide bridged peptides (NDBPs) (Almaaytah and Albalas, 2014). NDBPs are known to display diverse biological functions to defend from pathogen infection via the venom gland by showing antimicrobial activity and/or to capture prey by showing insecticidal activities (Almaaytah and Albalas, 2014, Dias et al., 2018, Miyashita, et al., 2010). NDBPs are classified into five subfamilies based on the sequence similarity and the molecular size. In this study, we identified nine transcripts coding for NDBPs (Fig. 2a); two transcripts (LaNDBP2-1 and LaNDBP2-2) sharing sequence similarity to those of group 2 consisting of long-chain multifunctional peptides, two transcripts (LaNDBP3-1 and LaNDBP3-2) to those of group 3 consisting of medium-

296 length antimicrobial peptides, and five transcripts (LaNDBP4-1, LaNDBP4-2,
297 LaNDBP4-3, LaNDBP4-4, and LaNDBP4-5) to those of group 4 consisting of short
298 antimicrobial peptides.
299
300 3.2.2. Invertebrate defensins
301 Defensins are cationic peptides stabilized with three disulfide bonds, which are
302 observed in a wide variety of organisms, including plants, insects, and mammals (Holly
303 et al., 2017, Yi et al., 2014). These peptides play an important role in innate immunity by
304 showing antimicrobial activity. Peptides similar to defensins are also found in scorpion
305 venom (Cheng et al., 2020, Harrison et al., 2014, Meng et al., 2016). Although a limited
306 number of scorpion venom peptides were reported as a defensin, a similar structural motif
307 exists in potassium channel toxins as described below (Zhu, S. Y. et al., 2014). In this
308 study, we found three transcripts (LaDefensin1, LaDefensin2, and LaDefensin3) coding
309 for invertebrate defensins, which share sequence similarity to AbDef-1 that was identified
310 from the venom of *Androctonus bicolor* (Fig. 2b) (Zhang et al., 2015).
311
312 3.2.3. Potassium channel toxin-like peptides
313 Potassium channel toxins (KTx) are cysteine-rich peptides that act on potassium
314 channels (Jimenez-Vargas et al., 2017). To date, a large number of KTx peptides have
315 been identified from scorpion venom. KTx peptides are currently classified into seven
316 groups (α-, β-, γ-, δ-, ε-, κ-, and λ-KTx) based on the sequence similarity and disulfide-
317 bonding patterns (Chen, Z. Y. et al., 2012, Jimenez-Vargas, et al., 2017). In this study, we
318 identified 18 transcripts coding for KTx peptides, and the majority of them were α-KTx
319 peptides (La-alphaKTx1, La-alphaKTx2, La-alphaKTx3, La-alphaKTx4, La-alphaKTx5,
320 La-alphaKTx6, La-alphaKTx7, La-alphaKTx8, and La-alphaKTx9, Fig. 3a). α-KTx
321 adopts a typical fold consisting of an α-helix and three β-strands stabilized by three or
322 four disulfide bonds (CS-αβ fold). We also found three transcripts (La-betaKTx1, La-
323 betaKTx2, and La-betaKTx3) coding for β-KTx peptides (Fig. 3b). These peptides adopt

a CS-αβ fold as observed in α-KTx peptides, but they have an additional α-helical structure without disulfide bonds at the N-terminal region. La-betaKTx1 and La-betaKTx2 correspond to LaIT2 and LaIT3, which were previously identified from the *L. australasiae* venom at a peptide level, respectively (Juichi, et al., 2019, Matsushita, et al., 2009). Five transcripts (La-kappaKTx1, La-kappaKTx2, La-kappaKTx3, La-kappaKTx4, and La-kappaKTx5) coding for κ-KTx peptides were also found (Fig. 3c). Unlike α- and β-KTx peptides, κ-KTx peptides do not adopt a CS-αβ fold, but instead they have two parallel α-helices stabilized with two disulfide bonds. The inhibitory activity of κ-KTx on potassium channels is known to be relatively weak compared with that of α-KTx peptides. Furthermore, one transcript (La-deltaKTx1) coding for δ-KTx was identified, which has a similar structure with Kunitz-type protease inhibitors consisting of two antiparallel β-strands and an α-helix stabilized by three or four disulfide bonds (Fig. 3d). Some δ-KTx peptides, such as BmKTT-2 from the *Mesobuthus martensii* venom, are known to show inhibitory activity on both potassium channels and proteases (Chen, Z. Y., et al., 2012).

3.2.4. Disulfide-directed hairpin peptides

Peptides with disulfide-directed hairpin (DDH) motif were identified from the venom of the limited number of scorpion species (Horita et al., 2011, Smith, J. J. et al., 2011). Although DDH peptides contain only two disulfide bonds, their structures are similar to the inhibitory cystine knot (ICK) motif, which is stabilized by three disulfide bonds. LaIT1, an insecticidal peptide previously identified from the venom of *L. australasiae*, is the first example of DDH peptides (Matsushita, et al., 2007). In this study, two transcripts (LaDDH1 and LaDDH2) coding for DDH peptides were identified, in which LaDDH1 corresponds to LaIT1 (Fig. 4a).

3.2.5. Sodium channel toxin-like peptides

Sodium channel toxins (NaTxs) adopt a CS-αβ fold as observed in α-KTx peptides,

but their sequences are relatively long (58-76 residues) (Housley, et al., 2017, Quintero-Hernandez, et al., 2013). NaTx peptides were identified mainly from the venom of Buthidae scorpions, and their sodium channel modulating activity is thought to be responsible for the relatively high toxicity of Buthidae scorpion venom (Cid-Uribe et al., 2019, de Oliveira et al., 2015, Luna-Ramirez, K. et al., 2015, Ward et al., 2018, Zhao, R. M. et al., 2010, Zhong et al., 2017). In addition to ion channel modulating activity, several members of this family are known to induce adipocyte lipolysis by forming a homodimer (Soudani et al., 2005, Zhu, S. and Gao, 2006). In the present study, two transcripts (LaLAP1 and LaLAP2) coding for the peptides similar to lipolysis-activating peptides were identified (Fig. 4b). However, these peptides could have other biological functions because the total number or the position of Cys residues are different from those of the known lipolysis-activating peptides (Zhu, S. and Gao, 2006).

3.2.6. Serine protease inhibitor-like peptides

Animal venom generally contains protease inhibitors probably for protecting the venom components from enzymatic degradation. Two types of inhibitors having different structural motifs (Kunitz- and Ascaris-type) are known to exist in scorpion venom (Chen, Z. Y. et al., 2013, Ranasinghe and McManus, 2013). In this study, we found one transcript (La-deltaKTx1) coding for Kunitz-type inhibitor-like peptides, which was classified as δ-KTx peptides as described above (Fig. 3d). In addition, we found eight transcripts (LaAPI1, LaAPI2, LaAPI3, LaAPI4, LaAPI5, LaAPI6, LaAPI7, and LaAPI8) coding for Ascaris-type inhibitor-like peptides (Fig. 4c). Interestingly, LaAPI1 contains only eight Cys residues that can form four disulfide bonds, whereas a typical Ascaris-type motif is stabilized by five disulfide bonds.

3.2.7. La1-like peptides

La1 was previously identified as the most abundant peptide in the *L. australasiae* venom (Miyashita, et al., 2007). To date, peptides similar to La1 have been identified

from the venom of diverse scorpion species, although their biological significance remains unknown. La1-like peptides adopt a single domain von Willebrand factor type C (SVWC) motif (Sheldon et al., 2007). Some peptides adopting this motif based on the position of Cys residues are classified as SVWC peptides. In the present study, we identified eight transcripts (La1, La1-1, La1-2, La1-3, La1-4, La1-5, La1-6, and La1-7) coding for La1-like or SVWC peptides (Fig. 5).

3.2.8. Enzymes

In scorpion venom, enzymes have not been recognized as major bioactive components. However, recent progress in venom gland transcriptome analysis revealed that there are various enzymes, such as phospholipase, metalloprotease, serine protease, and hyaluronidase, in the venom (Bordon et al., 2015, Carmo et al., 2014, Gao, R. et al., 2008, Krayem and Gargouri, 2020). In this study, we identified three transcripts (LaPLA$_2$-1, LaPLA$_2$-2, and LaPLA$_2$-3) coding for phospholipase A$_2$, and eight transcripts (LaSP1, LaSP2, LaSP3, LaSP4, LaSP5, LaSP6, LaSP7, and LaSP8) coding for serine protease (Table S2). In addition to these enzymes, we identified one transcript (La-alpha-amylase) coding for α-amylase. Since only a limited number of α-amylase sequences have been identified from scorpion venom at a transcript level, there is little information on their biological functions.

3.2.9. Other proteins

We identified 10 transcripts (LaCRVP1, LaCRVP2, LaCRVP3, LaCRVP4, LaCRVP5, LaCRVP6, LaCRVP7, LaCRVP8, LaCRVP9, and LaCRVP10) coding for cysteine-rich secretory proteins, and three transcripts (LaVP1, LaVP2, and LaVP3) coding for proteins with an insulin-like growth factor binding motif (Table S2). Although these proteins are widely observed in scorpion venom, their biological functions remain unknown (Amorim et al., 2019, de Oliveira, et al., 2015, Yang et al., 2014).

3.3. Isolation of an anti-HCV component

We first evaluated anti-HCV activity of the venom of *L. australasiae*. The result showed that the venom has a significant inhibitory activity against HCV infection at extremely low concentrations (IC_{50} = 0.01 μg/ml). To isolate a component responsible for the anti-HCV activity, the venom was separated using RP-HPLC on a C4 column (Fig. 6a). Each fraction was tested for anti-HCV activity, and the fraction eluting at 37-42 min was found to be active. This fraction was further separated using a C18 column (Fig. 6b). The HPLC peak eluting at 25.5-26.5 min showed the most significant activity, and the main component in this peak was found to be responsible for anti-HCV activity (Fig. 6c). Mass spectrometric analysis revealed the molecular mass of this component as 13,079.8 Da.

3.4. Structure determination of the anti-HCV component

The anti-HCV component was subjected to reduction and alkylation reactions to examine the number of disulfide bonds in this molecule. LC/MS analysis showed that the molecule was separated into two parts after the reactions, indicating that the anti-HCV component consists of two subunits cross-linked with disulfide bonds (Fig. 7a). Molecular masses of each subunit after alkylation of Cys residues were determined as 1145.2 Da (small subunit) and 12,524.7 Da (large subunit). A total increase of the molecular mass after Cys-alkylation was approximately 590 Da [(1145.2 + 12,524.7) – 13,079.8 = 590.1]. Considering a mass shift due to carboxymethylation of Cys residues (118 Da per disulfide bond), we determined that the anti-HCV component contains five disulfide bonds (118 x 5 = 590).

To obtain partial sequence information from the anti-HCV component, MS/MS de novo sequencing analysis was performed for the small subunit. As shown in Fig. 7b, the sequence of the small subunit was determined as KCVAHWKES. This sequence was searched from all transcripts obtained in this study, and the anti-HCV component was identified as $LaPLA_2$-1, one of the phospholipase A_2 proteins. As observed in other

436 scorpion PLA_2s, $LaPLA_2$-1 forms a heterodimeric structure through processing. However,
437 a sequence region that corresponds to the large subunit could not be determined based on
438 its observed molecular mass (12,524.7 Da). It was assumed that the large subunit
439 undergoes some types of post-translational modifications (PTMs). To determine a mature
440 structure of the large subunit of $LaPLA_2$-1 including PTMs, it was digested with Lys-C
441 or Glu-C, which was analyzed by LC/MS (Table 1). The N- and C-terminal sites of the
442 large subunit were identified based on the sequences of two Lys-C digests (LIFPGTK and
443 CFVLDCD), because these peptides cannot be generated by Lys-C digestion (–
444 IHQR/LIFPG– and –VLDCD/KRRF–). The calculated molecular mass of the deduced
445 sequence of the large subunit is 11,826.2 Da, which is still 698.5 Da lower than the
446 measured value (12,524.7 Da). When the LC/MS data of Lys-C and Glu-C digests were
447 carefully examined, the molecular masses of two digested peptides
448 (AANYSDLGSAAETDK and LIFPGTKWCGAGDKAANYSDLGSAAE) were found
449 to be 698 Da higher than those calculated without modifications. These peptides share the
450 same sequence region containing an N-linked glycosylation consensus motif (NXS)
451 (Marshall, 1972). This suggests that the Asn residue in this region is glycosylated. To
452 confirm the existence of an N-linked glycan at this position, the protein was treated with
453 PNGase A, which can cleave the N-glycan moiety from the Asn residue (Plummer and
454 Tarentino, 1981). LC/MS analysis showed that the measured molecular mass of the large
455 subunit after removal of N-glycan was identical to the calculated one (11,304.9 Da, Fig.
456 S1). The large subunit without N-glycan was then treated with Lys-C after reduction and
457 alkylation of Cys residues. The existence of the peptide containing an Asp residue instead
458 of the N-glycosylated Asn residue was confirmed by LC/MS analysis (Table 1). To obtain
459 structural information of the N-glycan moiety, MS/MS analysis was performed for the
460 Lys-C digested peptide containing the N-glycan moiety. As shown in Fig. 8, several
461 fragment ions formed through regular and successive losses of 203 and 145 Da, which
462 correspond to an N-acetyl hexose and a deoxyhexose, respectively, were observed in the
463 product ion spectrum. Based on the fragmentation pattern and the molecular masses, the

N-glycan moiety is supposed to consist of two N-acetyl hexoses and two deoxyhexoses. Considering the N-glycan structures and the biosynthetic pathways reported for arthropods including scorpions and insects (Hassani et al., 1999, Staudacher et al., 1992, Walski et al., 2017), the N-glycan structure in LaPLA$_2$-1 was estimated as shown in Fig. 8. Homology searches revealed that LaPLA$_2$-1 shares high sequence similarity to hemilipin from *Hemiscorpius lepturus*, HgPLA2 from *Hadrurus gertschi*, and phaiodactylipin from *Anuroctonus phaiodactylus* (Fig. 9) (Jridi et al., 2015, Schwartz et al., 2007, Valdez-Cruz et al., 2004).

3.5. Antiviral activity of LaPLA$_2$-1

LaPLA$_2$-1 inhibited infectivity of HCV in a concentration-dependent manner from 0.1 to 100 ng/ml (Fig. 10). Antiviral activity of LaPLA$_2$-1 against several viruses was further examined. Accordingly, IC_{50} values of LaPLA$_2$-1 against HCV, DENV and JEV, which belong to the family *Flaviviridae*, were 2.0, 3.4 and 5.7 ng/ml, respectively (Table 2). On the other hand, LaPLA$_2$-1 did not exhibit virucidal activity against HSV-1, which belongs to the family *Herpesviridae* even at a much higher concentration (1,000 ng/ml) (Fig. 10 and Table 2). To determine whether the post-entry step of the viral life cycle is affected by LaPLA$_2$-1, possible effect of post-entry treatment with LaPLA2-1 on virus replication in infected cells was examined. The result demonstrated that post-entry treatment with LaPLA$_2$-1 barely exhibited anti-HCV activity even at a high concentration (100 ng/ml), whereas pretreatment exhibited highly potent anti-HCV activity (Table S3).

To examine whether the virucidal activity of LaPLA$_2$-1 against HCV is associated with its enzymatic activity, the effect of PLA$_2$ inhibitor on the anti-HCV activity was evaluated. The result obtained revealed a dramatic decrease of the virucidal activity in the presence of manoalide, the PLA$_2$ inhibitor (Lombardo and Dennis, 1985). In addition, LaPLA$_2$-1 exhibited the significant phospholipase activity at a comparable level that was observed for bee venom PLA$_2$ (Fig. 11). These results suggest that the virucidal activity of LaPLA$_2$-1 against HCV is closely associated with its enzymatic activity. In the present study, the

concentration ranges of LaPLA$_2$-1 that showed virucidal and PLA$_2$ activities appear to be different. However, this may be due to differences in states of the substrate (membrane versus solution) and incubation time used for measurements (2 h versus 10 min) between virucidal and PLA$_2$ activity tests.

We also evaluated the toxicity of LaPLA$_2$-1 to the host cells. LaPLA$_2$-1 showed no cytotoxic effect at concentrations of up to 100 ng/ml and only a marginal cytotoxic effect at 1,000 ng/ml (Table S4), suggesting that observed LaPLA$_2$-1 action is specific to the viruses.

4. Discussion

4.1. Peptides identified by venom gland transcriptome analysis

Previously, we isolated and characterized four peptides (LaIT1, LaIT2, LaIT3, and La1) from the venom of the scorpion *L. australasiae*, mainly based on toxicity against insects (Juichi, et al., 2019, Matsushita, et al., 2009, Matsushita, et al., 2007, Miyashita, et al., 2007). In the present study, we performed the transcriptome analysis of the venom gland of *L. australasiae* to comprehensively understand the components in the venom. This resulted in the identification of 77 transcripts coding for peptides and proteins similar to those from the venom of other scorpion species. Among them, KTx-like peptides, including insecticidal toxins LaIT2 and LaIT3, were more diverse than other families of venom components (Fig. 1). On the other hand, only three NaTx-like peptides were identified in this study. This is consistent with the previous observations obtained by mass spectrometric analysis (Miyashita, et al., 2007). The number of the components with molecular masses of 6,000-9,000 Da, which may correspond to NaTx-like peptides, detected at the peptide level was relatively small in the *L. australasiae* venom when compared with that observed in the Buthidae scorpion venom. On the other hand, the number of the components with molecular masses of 3,000-5,000 Da, which may correspond to KTx-like peptides, was relatively large at the peptide level. A similar trend was noted in the transcriptome analysis of the venom glands of *Hadogenes troglodytes* of

the family Hormuridae (Zhong, et al., 2017). The study showed the existence of many KTx-like peptides in the venom, whereas no NaTx-like peptides were found. It is known that NaTx peptides have been found mainly from the venom of Buthidae scorpions (Cid-Uribe, et al., 2019, de Oliveira, et al., 2015, Luna-Ramirez, K., et al., 2015, Quintero-Hernandez, et al., 2013, Ward, et al., 2018, Zhao, R. M., et al., 2010, Zhong, et al., 2017). This suggests that sodium channels are not a main target of the venom components in non-Buthidae scorpions.

LaIT2 and LaIT3 belong to the β-KTx peptide group, which is structurally characterized by the α-helical region attached to the N-terminal of a CS-αβ fold. In addition to these toxins, we found another peptide having a structure similar to β-KTx peptides (Fig. 3b). This peptide (La-betaKTx3) has the sequence similar to scorpine-like peptides, such as SCl1 from *Urodacus yaschenkoi*, which are classified as one of the subfamilies of β-KTx peptides (Luna-Ramirez, Karen et al., 2016, Luna-Ramirez, K. et al., 2013). Since scorpine-like peptides are known to show antibacterial activity, La-betaKTx3 is likely to have a similar biological function.

LaIT1 is the first insecticidal toxin identified from the *L. australasiae* venom (Matsushita, et al., 2007). In the present study, another peptide (LaDDH2) that has the sequence similar to LaIT1 was found (Fig. 4a). Two basic residues (R13 and R15) that are important for the expression of the insecticidal activity of LaIT1 are conserved in LaDDH2, suggesting that this peptide may also have insecticidal activity (Horita, et al., 2011). To date, peptides similar to LaIT1 have been identified from the family Hormuridae (*L. waigiensis* and *Opisthacanthus cayaporum*) (Silva et al., 2009, Smith, J. J., et al., 2011) and the family Hemiscorpiidae (*Hemiscorpius lepturus*) (Kazemi-Lomedasht et al., 2017). This suggests that LaIT1 and its related peptides have been evolved independently in limited scorpion species.

La1 was previously identified as the most abundant component in the *L. australasiae* venom, although its biological function remains unknown (Miyashita, et al., 2007). Unlike LaIT1, peptides similar to La1 have been found in a wide variety of scorpion

548 species from seven families; Buthidae (Zeng et al., 2013, Zhao, R. M., et al., 2010),
549 Hemiscorpiidae (Kazemi-Lomedasht, et al., 2017), Hormuridae (Silva, et al., 2009,
550 Zhong, et al., 2017), Scorpionidae (Abdel-Rahman et al., 2013, Deng, Y. C. et al., 2018,
551 Diego-Garcia et al., 2012, Luna-Ramirez, K., et al., 2015), Superstitioniidae (Santibanez-
552 Lopez et al., 2016), and Vaejovidae (Quintero-Hernandez et al., 2015, Romero-Gutierrez
553 et al., 2018). Particularly, many La1-like peptides were observed in the venom gland
554 transcriptome of *H. troglodytes* (Zhong, et al., 2017). Therefore, it is possible that La1-
555 like peptides are abundant in the venom of Hormuridae scorpions, as observed in this
556 study (Fig. 5). La1 adopts an SVWC motif, which has been observed in peptides from a
557 wide variety of arthropods, including insects (Sheldon, et al., 2007). Recently, it has been
558 reported that SVWC peptides may play a role in protection against entomopathogenic
559 fungi in the epidermis of the silkworm and in basal AMPs expression of bumblebee (Han
560 et al., 2017, Wang et al., 2017). Although the mechanisms of action of these peptides are
561 still unknown, La1 might play a role in preventing pathogenic infections in the venom
562 glands as well.

563 In addition to the peptides having disulfide bonds, many NDBPs were identified in
564 this study (Fig. 2a). These peptides generally show membrane-disrupting activity by
565 forming an amphipathic α-helical structure (Harrison, et al., 2014). Among five
566 subfamilies of scorpion NDBPs, peptides of four subfamilies were found in this study.
567 The reasons for the presence of the various types of NDBPs in the venom are unknown,
568 but they are likely to have different biological functions. For example, four NDBPs
569 isolated from *Isometrus maculatus* show different antimicrobial spectra, and some of
570 them also exhibit insecticidal and hemolytic activities (Miyashita et al., 2017).
571 Furthermore, some NDBPs may have a synergistic effect on other neurotoxins in venom
572 to enhance their activity (Gao, B. et al., 2018). Defensins are another type of antimicrobial
573 peptides observed in a wide variety of organisms (Holly, et al., 2017, Yi, et al., 2014). In
574 this study, three defensin-like peptides were found (Fig. 2b). Since invertebrate defensins
575 have the same disulfide-bonding pattern as peptides that act on potassium channels, it is

576 also possible that defensins identified in this study could show some ion channel-
577 modulating activity (Meng, et al., 2016).

578 A number of peptides similar to serine protease inhibitors were also found in the
579 present study (Fig. 4c). Kunitz-type protease inhibitors have been found in many
580 organisms (Mishra, 2020). The peptides having a Kunitz-type motif have been reported
581 to be present in the venom of various scorpion species (Ranasinghe and McManus, 2013).
582 In addition to Kunitz-type peptides, Ascaris-type protease inhibitors have also been
583 identified from the scorpion venom (Chen, Z. Y., et al., 2013). These peptides are likely
584 responsible for preventing degradation of venom components by proteases of its own or
585 other organisms. In addition, some of the protease inhibitor peptides from scorpion venom
586 are known to act on potassium channels. This is thought to be a result of divergent
587 evolution. For example, BmKTT-2, a Kunitz-type peptide identified from *Mesobuthus*
588 *martensii*, shows inhibitory activity on Kv1.3 channels as well as inhibition of trypsin
589 activity (Chen, Z. Y., et al., 2012). Since La-deltaKTx1 identified in this study (Fig. 3d)
590 shares sequence similarity to BmKTT-2, it could show potassium channel inhibition
591 activity.

592

593 *4.2. Anti-HCV component in the* L. australasiae *venom*

594 In the present study, we found that the *L. australasiae* venom has a potent virucidal
595 activity against HCV and successfully identified LaPLA$_2$-1 as an anti-HCV component
596 in the venom using a bioassay-guided approach. PLA$_2$ enzymes can be found in a wide
597 variety of organisms, including snake and bee venom (Dennis et al., 2011). They catalyze
598 the hydrolysis of glycerophospholipids at the *sn*-2 position to release free fatty acids and
599 lysophospholipids. PLA$_2$ molecules can be classified into six types, and scorpion venom
600 PLA$_2$ (scvPLA$_2$) belongs to group III of secreted PLA$_2$ (sPLA$_2$) (Krayem and Gargouri,
601 2020). This group also includes PLA$_2$ from bee venom (bvPLA$_2$). Some of sPLA$_2$
602 molecules are known to exhibit antiviral activity, and their enzymatic activity is suggested
603 to be associated with antiviral activity (Chen, M. et al., 2017). For example, PLA$_2$

molecules from snake venom (svPLA$_2$) specifically act on *Flaviviridae* viruses. This is likely because they can degrade the viral envelope membranes, particularly those derived from endoplasmic reticulum (ER). In fact, LaPLA$_2$-1 efficiently inhibited the infection of *Flaviviridae* viruses such as HCV, DENV and JEV (IC_{50} = 2.0, 3.4 and 5.7 ng/ml, respectively). A comparable degree of virucidal activity was reported for PLA$_2$s obtained from snake venom (Teixeira et al., 2020). However, no significant antiviral activity was observed for a *Herpesviridae* virus such as HSV-1, which has the plasma membrane (PM)-derived envelope. This suggests that LaPLA$_2$-1 has the same mechanism of action as svPLA$_2$ (Chen, M., et al., 2017). It has been shown that phaiodactylipin can hydrolyze phosphatidylcholine more preferably than phosphatidylethanolamine and phosphatidylserine (Valdez-Cruz, et al., 2004). Since the ER-budded viruses contain a relatively high proportion of phosphatidylcholine than PM-budded viruses (Callens et al., 2016), LaPLA$_2$ may have a similar substrate preference as observed for phaiodactylipin. The possible action specificity of LaPLA$_2$-1 on the ER membrane implies that it may have no adverse effect on the host cells, which is further supported by its low cytotoxicity with CC_{50} being >1,000 ng/ml. Although PLA$_2$s, in general, are known to be cytotoxic and inflammatory or neurotoxic, LaPLA$_2$-1 or its derivative(s) might be a good candidate as a lead compound for the development of an antiviral drug against HCV, DENV and other viruses budding from the ER membranes.

It is known that the Ca^{2+}-binding motif (XCGXG) and the catalytic center with a His-Asp dyad are important for the expression of enzyme activity of PLA$_2$ (Dennis, et al., 2011). In the structures of scvPLA$_2$s, these motifs are conserved except for phaiodactylipin, in which one of the Gly residues in the Ca^{2+}-binding motif is missing (Fig. 9). The structure of scvPLA$_2$s is further characterized by the formation of a heterodimer consisting of a large and a small subunit cross-linked with a disulfide bond, although its relevance to biological function remains unknown. It is also known that many bvPLA$_2$s contain N-glycans, most of which are core-fucosylated (Kubelka et al., 1993). Glycosylation of scvPLA$_2$ molecules has been investigated particularly for

phaiodactylipin (Valdez-Cruz, et al., 2004). Three N-glycosylation sites were recognized in the structure of phaiodactylipin, one of which is located at the N-terminal region of the large subunit, as observed in LaPLA$_2$-1 (Fig. 9). Hemilipin, HgPLA$_2$, and bvPLA$_2$ also contain one or two possible N-glycosylation sites in their sequences, although the presence of glycans is not experimentally examined in hemilipin and HgPLA$_2$. Among the N-glycosylation sites in these PLA$_2$ molecules, that at the N-terminal region of the large subunit is commonly observed. Since this site is close to the catalytic center, it may play an important role in enzymatic and/or antiviral activity.

5. Conclusion

Scorpion venom is known to be a rich source of bioactive peptides and proteins. Previously, we identified three insecticidal peptides from the venom of the *L. australasiae* scorpion, but the activities and structures of the other venom components remained unknown. In the present study, we performed a transcriptomic analysis of the venom gland of *L. australasiae* to elucidate a comprehensive picture of the venom components. As a result, 77 transcripts coding for venom peptides and proteins, including four previously reported peptides, were successfully identified. Among them, KTx-like peptides were the most diverse, suggesting that these peptides play a crucial role in this venom. On the other hand, a relatively small number of NaTx-like peptides were found. This is consistent with previous findings that NaTx-like peptides are predominantly found in the venom of Buthidae scorpions. In addition, a relatively large number of peptides without disulfide bonds were identified, suggesting that these peptides have some critical functions in the *L. australasiae* venom.

These transcriptome data facilitated the structural determination of the anti-HCV component in the venom, which was obtained by a bioassay-guided approach. Mass spectrometric analysis revealed that this component is one of the PLA$_2$s (LaPLA$_2$-1). This is the first report of a PLA$_2$ with antiviral activity from scorpion venom, although PLA$_2$s from snake and bee venom are known to show antiviral activity. Since the anti-HCV

activity of LaPLA$_2$-1 was inhibited by a PLA$_2$ inhibitor, the enzymatic activity of LaPLA$_2$-1 is likely to be involved in its expression of the anti-HCV activity. It is known that scvPLA$_2$s, including LaPLA$_2$-1, adopt a heterodimeric structure and are likely to be post-translationally modified by N-glycosylation at the N-terminal region. Elucidation of the relationship between the characteristic structures and enzymatic/antiviral activity of svcPLA$_2$s will be a subject of future research.

Acknowledgements

The authors are grateful to Dr. C. M. Rice (The Rockefeller University, New York, NY, U.S.A.) for providing pFL-J6/JFH1. We also thank Dr. Hajime Ono (Kyoto University) for valuable advice on the RNA extraction. This study was supported by the Cooperative Research Program of the Genome Research for BioResource, NODAI Genome Research Center, Tokyo University of Agriculture. This study was also supported in part by the Program on the Innovative Development and the Application of New Drugs for Hepatitis B from the Japan Agency for Medical Research and Development (AMED) under the grant number JP19fk0310103h2103 and by JSPS KAKENHI grant number JP19K05842.

References

Abdel-Rahman, M.A., Quintero-Hernandez, V., Possani, L.D., 2013. Venom proteomic and venomous glands transcriptomic analysis of the Egyptian scorpion *Scorpio maurus palmatus* (Arachnida: Scorpionidae). Toxicon 74, 193-207, 10.1016/j.toxicon.2013.08.064

Ahmadi, S., Knerr, J.M., Argemi, L., Bordon, K.C.F., Pucca, M.B., Cerni, F.A., Arantes, E.C., Caliskan, F., Laustsen, A.H., 2020. Scorpion venom: Detriments and benefits. Biomedicines 8, 118, 10.3390/biomedicines8050118

Almaaytah, A., Albalas, Q., 2014. Scorpion venom peptides with no disulfide bridges: A review. Peptides 51, 35-45, 10.1016/J.Peptides.2013.10.021

688 Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local
689 alignment search tool. J. Mol. Biol. 215, 403-410, 10.1006/jmbi.1990.9999
690 Amorim, F.G., Longhim, H.T., Cologna, C.T., Degueldre, M., De Pauw, E., Quinton, L.,
691 Arantes, E.C., 2019. Proteome of fraction from *Tityus serrulatus* venom reveals new
692 enzymes and toxins. J. Venom Anim. Toxins 25, e148218, 10.1590/1678-9199-
693 JVATITD-1482-18
694 Aoki-Utsubo, C., Chen, M., Hotta, H., 2018. Time-of-addition and temperature-shift
695 assays to determine particular step(s) in the viral life cycle that is blocked by antiviral
696 substance(s). Bio-Protocol 8, e2830, 10.21769/BioProtoc.2830
697 Bordon, K.C., Wiezel, G.A., Amorim, F.G., Arantes, E.C., 2015. Arthropod venom
698 hyaluronidases: biochemical properties and potential applications in medicine and
699 biotechnology. J. Venom Anim. Toxins Incl. Trop. Dis. 21, 43, 10.1186/s40409-015-
700 0042-7
701 Callens, N., Brugger, B., Bonnafous, P., Drobecq, H., Gerl, M.J., Krey, T., Roman-Sosa,
702 G., Rumenapf, T., Lambert, O., Dubuisson, J., Rouille, Y., 2016. Morphology and
703 molecular composition of purified bovine viral diarrhea virus envelope. PLOS Pathog.
704 12, e1005476, 10.1371/journal.ppat.1005476
705 Carmo, A.O., Oliveira-Mendes, B.B., Horta, C.C., Magalhaes, B.F., Dantas, A.E.,
706 Chaves, L.M., Chavez-Olortegui, C., Kalapothakis, E., 2014. Molecular and functional
707 characterization of metalloserrulases, new metalloproteases from the *Tityus serrulatus*
708 venom gland. Toxicon 90, 45-55, 10.1016/j.toxicon.2014.07.014
709 Chang, Z., Li, G.J., Liu, J.T., Zhang, Y., Ashby, C., Liu, D.L., Cramer, C.L., Huang,
710 X.Z., 2015. Bridger: a new framework for de novo transcriptome assembly using
711 RNA-seq data. Genome Biol. 16, 30, 10.1186/s13059-015-0596-2
712 Chen, M., Aoki-Utsubo, C., Kameoka, M., Deng, L., Terada, Y., Kamitani, W., Sato, K.,
713 Koyanagi, Y., Hijikata, M., Shindo, K., Noda, T., Kohara, M., Hotta, H., 2017. Broad-
714 spectrum antiviral agents: secreted phospholipase A2 targets viral envelope lipid
715 bilayers derived from the endoplasmic reticulum membrane. Sci. Rep. 7, 15931,

10.1038/s41598-017-16130-w

Chen, Z.Y., Hu, Y.T., Yang, W.S., He, Y.W., Feng, J., Wang, B., Zhao, R.M., Ding, J.P., Cao, Z.J., Li, W.X., Wu, Y.L., 2012. Hg1, novel peptide inhibitor specific for Kv1.3 channels from first scorpion Kunitz-type potassium channel toxin family. J. Biol. Chem. 287, 13813-13821, 10.1074/jbc.M112.343996

Chen, Z.Y., Wang, B., Hu, J., Yang, W.S., Cao, Z.J., Zhuo, R.X., Li, W.X., Wu, Y.L., 2013. SjAPI, the first functionally characterized Ascaris-type protease inhibitor from animal venoms. PLOS ONE 8, e57529, 10.1371/journal.pone.0057529

Cheng, Y.T., Sun, F., Li, S., Gao, M.J., Wang, L.Y., Sarhan, M., Abdel-Rahman, M.A., Li, W.X., Kwok, H.F., Wu, Y.L., Cao, Z.J., 2020. Inhibitory activity of a scorpion defensin BmKDfsin3 against hepatitis C virus. Antibiotics 9, ARTN 33, 10.3390/antibiotics9010033

Cid-Uribe, J.I., Meneses, E.P., Batista, C.V.F., Ortiz, E., Possani, L.D., 2019. Dissecting toxicity: The venom gland transcriptome and the venom proteome of the highly venomous scorpion *Centruroides limpidus* (Karsch, 1879). Toxins 11, ARTN 247, 10.3390/toxins11050247

da Mata, E.C., Mourao, C.B., Rangel, M., Schwartz, E.F., 2017. Antiviral activity of animal venom peptides and related compounds. J. Venom Anim. Toxins Incl. Trop. Dis. 23, 3, 10.1186/s40409-016-0089-0

de Oliveira, U.C., Candido, D.M., Dorce, V.A.C., Junqueira-de-Azevedo, I.D.M., 2015. The transcriptome recipe for the venom cocktail of *Tityus bahiensis* scorpion. Toxicon 95, 52-61, 10.1016/j.toxicon.2014.12.013

Deng, L., Shoji, I., Ogawa, W., Kaneda, S., Soga, T., Jiang, D., Ide, Y.H., Hotta, H., 2011. Hepatitis C virus infection promotes hepatic gluconeogenesis through an NS5A-Mediated, FoxO1-dependent pathway. J Virol 85, 8556-8568, 10.1128/Jvi.00146-11

Deng, Y.C., Gu, J.W., Yan, Z.P., Wang, M.D., Ma, C.Q., Zhang, J.F., Jiang, G.X., Ge, M.X., Xu, S.G., Xu, Z., Xiao, L., 2018. De novo transcriptomic analysis of the venomous glands from the scorpion *Heterometrus spinifer* revealed unique and

744 extremely high diversity of the venom peptides. Toxicon 143, 1-19,
745 10.1016/j.toxicon.2017.12.051
746 Dennis, E.A., Cao, J., Hsu, Y.H., Magrioti, V., Kokotos, G., 2011. Phospholipase A2
747 enzymes: physical structure, biological function, disease implication, chemical
748 inhibition, and therapeutic intervention. Chem. Rev. 111, 6130-6185,
749 10.1021/cr200085w
750 Dias, N.B., de Souza, B.M., Cocchi, F.K., Chalkidis, H.M., Dorcec, V.A.C., Palma,
751 M.S., 2018. Profiling the short, linear, non-disulfide bond-containing peptidome from
752 the venom of the scorpion *Tityus obscurus*. J. Proteomics 170, 70-79,
753 10.1016/j.jprot.2017.09.006
754 Diego-Garcia, E., Peigneur, S., Clynen, E., Marien, T., Czech, L., Schoofs, L., Tytgat, J.,
755 2012. Molecular diversity of the telson and venom components from *Pandinus*
756 *cavimanus* (Scorpionidae Latreille 1802): Transcriptome, venomics and function.
757 Proteomics 12, 313-328, 10.1002/pmic.201100409
758 El-Bitar, A.M., Sarhan, M.M., Aoki, C., Takahara, Y., Komoto, M., Deng, L., Moustafa,
759 M.A., Hotta, H., 2015. Virocidal activity of Egyptian scorpion venoms against
760 hepatitis C virus. Virol. J. 12, 47, 10.1186/s12985-015-0276-6
761 Fu, Y., Li, C., Dong, S., Wu, Y., Zhangsun, D., Luo, S., 2018. Discovery methodology
762 of novel conotoxins from Conus species. Mar. Drugs 16, 417, 10.3390/md16110417
763 Gao, B., Dalziel, J., Tanzi, S., Zhu, S.Y., 2018. Meucin-49, a multifunctional scorpion
764 venom peptide with bactericidal synergy with neurotoxins. Amino Acids 50, 1025-
765 1043, 10.1007/s00726-018-2580-0
766 Gao, R., Zhang, Y., Gopalakrishnakone, P., 2008. Purification and N-terminal sequence
767 of a serine proteinase-like protein (BMK-CBP) from the venom of the Chinese
768 scorpion (*Buthus martensii* Karsch). Toxicon 52, 348-353,
769 10.1016/j.toxicon.2008.06.003
770 Ghosh, A., Roy, R., Nandi, M., Mukhopadhyay, A., 2019. Scorpion venom-toxins that
771 aid in drug development: A review. Int. J. Pept. Res. Ther. 25, 27-37, 10.1007/s10989-

018-9721-x

Gordon, D., Karbat, I., Ilan, N., Cohen, L., Kahn, R., Gilles, N., Dong, K., Stuhmer, W., Tytgat, J., Gurevitz, M., 2007. The differential preference of scorpion alpha-toxins for insect or mammalian sodium channels: implications for improved insect control. Toxicon 49, 452-472, 10.1016/j.toxicon.2006.11.016

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., MacManes, M.D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C.N., Henschel, R., Leduc, R.D., Friedman, N., Regev, A., 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protoc. 8, 1494-1512, 10.1038/nprot.2013.084

Han, F., Lu, A., Yuan, Y., Huang, W., Beerntsen, B.T., Huang, J., Ling, E., 2017. Characterization of an entomopathogenic fungi target integument protein, *Bombyx mori* single domain von Willebrand factor type C, in the silkworm, *Bombyx mori*. Insect Mol. Biol. 26, 308-316, 10.1111/imb.12293

Harrison, P.L., Abdel-Rahman, M.A., Miller, K., Strong, P.N., 2014. Antimicrobial peptides from scorpion venoms. Toxicon 88, 115-137, 10.1016/j.toxicon.2014.06.006

Hassani, O., Loew, D., Van Dorsselaer, A., Papandreou, M.J., Sorokine, O., Rochat, H., Sampieri, F., Mansuelle, P., 1999. Aah VI, a novel, N-glycosylated anti-insect toxin from *Androctonus australis* hector scorpion venom: isolation, characterisation, and glycan structure determination. FEBS Lett. 443, 175-180, 10.1016/S0014-5793(98)01710-4

Herzig, V., Cristofori-Armstrong, B., Israel, M.R., Nixon, S.A., Vetter, I., King, G.F., 2020. Animal toxins - Nature's evolutionary-refined toolkit for basic research and drug discovery. Biochem. Pharmacol. 181, ARTN 114096, 10.1016/j.bcp.2020.114096

Holly, M.K., Diaz, K., Smith, J.G., 2017. Defensins in viral infection and pathogenesis. Annu. Rev. Virol. 4, 369-391, 10.1146/annurev-virology-101416-041734

Horita, S., Matsushita, N., Kawachi, T., Ayabe, R., Miyashita, M., Miyakawa, T.,

Nakagawa, Y., Nagata, K., Miyagawa, H., Tanokura, M., 2011. Solution structure of a short-chain insecticidal toxin LaIT1 from the venom of scorpion *Liocheles australasiae*. Biochem. Biophys. Res. Commun. 411, 738-744, 10.1016/j.bbrc.2011.07.016

Housley, D.M., Housley, G.D., Liddell, M.J., Jennings, E.A., 2017. Scorpion toxin peptide action at the ion channel subunit level. Neuropharmacology 127, 46-78, 10.1016/j.neuropharm.2016.10.004

Jimenez-Vargas, J.M., Possani, L.D., Luna-Ramirez, K., 2017. Arthropod toxins acting on neuronal potassium channels. Neuropharmacology 127, 139-160, 10.1016/j.neuropharm.2017.09.025

Jridi, I., Catacchio, I., Majdoub, H., Shahbazeddah, D., El Ayeb, M., Frassanito, M.A., Ribatti, D., Vacca, A., Borchani, L., 2015. Hemilipin, a novel *Hemiscorpius lepturus* venom heterodimeric phospholipase A2, which inhibits angiogenesis in vitro and in vivo. Toxicon 105, 34-44, 10.1016/j.toxicon.2015.08.022

Juichi, H., Miyashita, M., Nakagawa, Y., Miyagawa, H., 2019. Isolation and characterization of the insecticidal, two-domain toxin LaIT3 from the *Liocheles australasiae* scorpion venom. Biosci. Biotechnol. Biochem. 83, 2183-2189, 10.1080/09168451.2019.1654849

Katoh, K., Rozewicki, J., Yamada, K.D., 2019. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief. Bioinform. 20, 1160-1166, 10.1093/bib/bbx108

Kazemi-Lomedasht, F., Khalaj, V., Bagheri, K.P., Behdani, M., Shahbazzadeh, D., 2017. The first report on transcriptome analysis of the venom gland of Iranian scorpion, *Hemiscorpius lepturus*. Toxicon 125, 123-130, 10.1016/j.toxicon.2016.11.261

Krayem, N., Gargouri, Y., 2020. Scorpion venom phospholipases A2: A minireview. Toxicon 184, 48-54, 10.1016/j.toxicon.2020.05.020

Kubelka, V., Altmann, F., Staudacher, E., Tretter, V., Marz, L., Hard, K., Kamerling, J.P., Vliegenthart, J.F.G., 1993. Primary structures of the N-linked carbohydrate chains

828 from honeybee venom phospholipase A2. Eur. J. Biochem. 213, 1193-1204,
829 10.1111/j.1432-1033.1993.tb17870.x
830 Lassmann, T., Hayashizaki, Y., Daub, C.O., 2009. TagDust-a program to eliminate
831 artifacts from next generation sequencing data. Bioinformatics 25, 2839-2840,
832 10.1093/bioinformatics/btp527
833 Lombardo, D., Dennis, E.A., 1985. Cobra venom phospholipase A2 inhibition by
834 manoalide. A novel type of phospholipase inhibitor. J. Biol. Chem. 260, 7234-7240,
835 Luna-Ramirez, K., Jimenez-Vargas, J.M., Lourival, D.P., 2016. Scorpine-like peptides.
836 Single Cell Biol. 5, 138, 10.4172/2168-9431.1000138
837 Luna-Ramirez, K., Quintero-Hernandez, V., Juarez-Gonzalez, V.R., Possani, L.D., 2015.
838 Whole transcriptome of the venom gland from *Urodacus yaschenkoi* scorpion. PLOS
839 ONE 10, e0127883, 10.1371/journal.pone.0127883
840 Luna-Ramirez, K., Quintero-Hernandez, V., Vargas-Jaimes, L., Batista, C.V.F., Winkel,
841 K.D., Possani, L.D., 2013. Characterization of the venom from the Australian scorpion
842 *Urodacus yaschenkoi*: Molecular mass analysis of components, cDNA sequences and
843 peptides with antimicrobial activity. Toxicon 63, 44-54,
844 10.1016/J.Toxicon.2012.11.017
845 Madeira, F., Park, Y.M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P.,
846 Tivey, A.R.N., Potter, S.C., Finn, R.D., Lopez, R., 2019. The EMBL-EBI search and
847 sequence analysis tools APIs in 2019. Nucleic Acids Res. 47, W636-W641,
848 10.1093/nar/gkz268
849 Marshall, R.D., 1972. Glycoproteins. Annu. Rev. Biochem. 41, 673-702,
850 10.1146/annurev.bi.41.070172.003325
851 Matsushita, N., Miyashita, M., Ichiki, Y., Ogura, T., Sakuradani, E., Nakagawa, Y.,
852 Shimizu, S., Miyagawa, H., 2009. Purification and cDNA cloning of LaIT2, a novel
853 insecticidal toxin from venom of the scorpion *Liocheles australasiae*. Biosci.
854 Biotechnol. Biochem. 73, 2769-2772, 10.1271/bbb.90509
855 Matsushita, N., Miyashita, M., Sakai, A., Nakagawa, Y., Miyagawa, H., 2007.

856 Purification and characterization of a novel short-chain insecticidal toxin with two
857 disulfide bridges from the venom of the scorpion *Liocheles australasiae*. Toxicon 50,
858 861-867, 10.1016/j.toxicon.2007.06.014
859 Meng, L.X., Xie, Z.L., Zhang, Q., Li, Y., Yang, F., Chen, Z.Y., Li, W.X., Cao, Z.J., Wu,
860 Y.L., 2016. Scorpion potassium channel-blocking defensin highlights a functional link
861 with neurotoxin. J. Biol. Chem. 291, 7097-7106, 10.1074/jbc.M115.680611
862 Mishra, M., 2020. Evolutionary aspects of the structural convergence and functional
863 diversification of Kunitz-domain inhibitors. J. Mol. Evol., 537-548, 10.1007/s00239-
864 020-09959-9
865 Miyashita, M., Kitanaka, A., Yakio, M., Yamazaki, Y., Nakagawa, Y., Miyagawa, H.,
866 2017. Complete de novo sequencing of antimicrobial peptides in the venom of the
867 scorpion *Isometrus maculatus*. Toxicon 139, 1-12, 10.1016/j.toxicon.2017.09.010
868 Miyashita, M., Otsuki, J., Hanai, Y., Nakagawa, Y., Miyagawa, H., 2007.
869 Characterization of peptide components in the venom of the scorpion *Liocheles*
870 *australasiae* (Hemiscorpiidae). Toxicon 50, 428-437, 10.1016/j.toxicon.2007.04.012
871 Miyashita, M., Sakai, A., Matsushita, N., Hanai, Y., Nakagawa, Y., Miyagawa, H., 2010.
872 A novel amphipathic linear peptide with both insect toxicity and antimicrobial activity
873 from the venom of the scorpion *Isometrus maculatus*. Biosci. Biotechnol. Biochem.
874 74, 364-369, 10.1271/bbb.90723
875 Oldrati, V., Arrell, M., Violette, A., Perret, F., Sprungli, X., Wolfender, J.L., Stocklin, R.,
876 2016. Advances in venomics. Mol. Biosyst. 12, 3530-3543, 10.1039/c6mb00516k
877 Plummer, T.H., Tarentino, A.L., 1981. Facile cleavage of complex oligosaccharides
878 from glycopeptides by almond emulsin peptide - N-glycosidase. J. Biol. Chem. 256,
879 243-246,
880 Quintero-Hernandez, V., Jimenez-Vargas, J.M., Gurrola, G.B., Valdivia, H.H., Possani,
881 L.D., 2013. Scorpion venom components that affect ion-channels function. Toxicon
882 76, 328-342, 10.1016/J.Toxicon.2013.07.012
883 Quintero-Hernandez, V., Ortiz, E., Rendon-Anaya, M., Schwartz, E.F., Becerril, B.,

884 Corzo, G., Possani, L.D., 2011. Scorpion and spider venom peptides: gene cloning and
885 peptide expression. Toxicon 58, 644-663, 10.1016/j.toxicon.2011.09.015
886 Quintero-Hernandez, V., Ramirez-Carreto, S., Romero-Gutierrez, M.T., Valdez-
887 Velazquez, L.L., Becerril, B., Possani, L.D., Ortiz, E., 2015. Transcriptome analysis of
888 scorpion species belonging to the *Vaejovis* genus. PLOS One 10, e0117188,
889 10.1371/journal.pone.0117188
890 Ranasinghe, S., McManus, D.P., 2013. Structure and function of invertebrate Kunitz
891 serine protease inhibitors. Dev. Comp. Immunol. 39, 219-227,
892 10.1016/j.dci.2012.10.005
893 Romero-Gutierrez, M.T., Santibanez-Lopez, C.E., Jimenez-Vargas, J.M., Batista, C.V.F.,
894 Ortiz, E., Possani, L.D., 2018. Transcriptomic and proteomic analyses reveal the
895 diversity of venom components from the vaejovid scorpion *Serradigitus gertschi*.
896 Toxins 10, 359, 10.3390/toxins10090359
897 Santibanez-Lopez, C.E., Cid-Uribe, J.I., Batista, C.V., Ortiz, E., Possani, L.D., 2016.
898 Venom gland transcriptomic and proteomic analyses of the enigmatic scorpion
899 *Superstitionia donensis* (Scorpiones: Superstitioniidae), with insights on the evolution
900 of its venom components. Toxins 8, 367, 10.3390/toxins8120367
901 Schwartz, E.F., Diego-Garcia, E., Rodriguez de la Vega, R.C., Possani, L.D., 2007.
902 Transcriptome analysis of the venom gland of the Mexican scorpion *Hadrurus*
903 *gertschi* (Arachnida: Scorpiones). BMC Genomics 8, 119, 10.1186/1471-2164-8-119
904 Schwartz, E.F., Mourao, C.B., Moreira, K.G., Camargos, T.S., Mortari, M.R., 2012.
905 Arthropod venoms: a vast arsenal of insecticidal neuropeptides. Biopolymers 98, 385-
906 405, 10.1002/bip.22100
907 Sheldon, T.J., Miguel-Aliaga, I., Gould, A.P., Taylor, W.R., Conklin, D., 2007. A novel
908 family of single VWC-domain proteins in invertebrates. FEBS Lett 581, 5268-5274,
909 10.1016/J.Febslet.2007.10.016
910 Silva, E.C.N., Camargos, T.S., Maranhao, A.Q., Silva-Pereira, I., Silva, L.P., Possani,
911 L.D., Schwartz, E.F., 2009. Cloning and characterization of cDNA sequences encoding

912 for new venom peptides of the Brazilian scorpion *Opisthacanthus cayaporum*.
913 Toxicon 54, 252-261, 10.1016/J.Toxicon.2009.04.010
914 Smith, J.J., Alewood, P.F., 2015. Modern venom profiling: Mining into scorpion venom
915 biodiversity, in: Gopalakrishnakone, P., Possani, L.D., Schwartz, E.F., Rodríguez de la
916 Vega, R.C., (Eds), Scorpion Venoms. Springer, Dordrecht.
917 Smith, J.J., Herzig, V., King, G.F., Alewood, P.F., 2013. The insecticidal potential of
918 venom peptides. Cell. Mol. Life Sci. 70, 3665-3693, 10.1007/s00018-013-1315-3
919 Smith, J.J., Hill, J.M., Little, M.J., Nicholson, G.M., King, G.F., Alewood, P.F., 2011.
920 Unique scorpion toxin with a putative ancestral fold provides insight into evolution of
921 the inhibitor cystine knot motif. P. Natl. Acad. Sci. USA 108, 10478-10483,
922 10.1073/pnas.1103501108
923 Song, J., Fujii, M., Wang, F., Itoh, M., Hotta, H., 1999. The NS5A protein of hepatitis C
924 virus partially inhibits the antiviral activity of interferon. J Gen Virol 80, 879-886,
925 10.1099/0022-1317-80-4-879
926 Soudani, N., Gharbi-Chihi, J., Srairi-Abid, N., Martin-El Yazidi, C., Planells, R.,
927 Margotat, A., Torresani, J., El Ayeb, M., 2005. Isolation and molecular
928 characterization of LVP1 lipolysis activating peptide from scorpion *Buthus occitanus*
929 *tunetanus*. BBA-Proteins Proteom. 1747, 47-56, 10.1016/j.bbapap.2004.09.020
930 Staudacher, E., Altmann, F., Marz, L., Hard, K., Kamerling, J.P., Vliegenthart, J.F.G.,
931 1992. Alpha-1-6(alpha-1-3)-difucosylation of the asparagine-bound N-
932 acetylglucosamine in honeybee venom phospholipase-A2. Glycoconjugate J. 9, 82-85,
933 10.1007/Bf00731703
934 Teixeira, S.C., Borges, B.C., Oliveira, V.Q., Carregosa, L.S., Bastos, L.A., Santos, I.A.,
935 Jardim, A.C.G., Melo, F.F., Freitas, L.M., Rodrigues, V.M., Lopes, D.S., 2020. Insights
936 into the antiviral activity of phospholipases A2 (PLA2s) from snake venoms. Int. J.
937 Biol. Macromol. 164, 616-625, 10.1016/j.ijbiomac.2020.07.178
938 Valdez-Cruz, N.A., Batista, C.V.F., Possani, L.D., 2004. Phaiodactylipin, a glycosylated
939 heterodimeric phospholipase A2 from the venom of the scorpion *Anuroctonus*

940 *phaiodactylus*. Eur. J. Biochem. 271, 1453-1464, 10.1111/j.1432-1033.2004.04047.x
941 Vetter, I., Davis, J.L., Rash, L.D., Anangi, R., Mobli, M., Alewood, P.F., Lewis, R.J.,
942 King, G.F., 2011. Venomics: a new paradigm for natural products-based drug
943 discovery. Amino Acids 40, 15-28, 10.1007/s00726-010-0516-4
944 Walker, A.A., Robinson, S.D., Hamilton, B.F., Undheim, E.A.B., King, G.F., 2020.
945 Deadly proteomes: A practical guide to proteotranscriptomics of animal venoms.
946 Proteomics 20, ARTN 1900324, 10.1002/pmic.201900324
947 Walski, T., De Schutter, K., Van Damme, E.J.M., Smagghe, G., 2017. Diversity and
948 functions of protein glycosylation in insects. Insect Biochem. Mol. Biol. 83, 21-34,
949 10.1016/j.ibmb.2017.02.005
950 Wang, H.D., Smagghe, G., Meeus, I., 2017. The role of a single gene encoding the
951 single von Willebrand factor C-domain protein (SVC) in bumblebee immunity extends
952 beyond antiviral defense. Insect. Biochem. Mol. Biol. 91, 10-20,
953 10.1016/j.ibmb.2017.10.002
954 Ward, M.J., Ellsworth, S.A., Rokyta, D.R., 2018. Venom-gland transcriptomics and
955 venom proteomics of the Hentz striped scorpion (*Centruroides hentzi*; Buthidae)
956 reveal high toxin diversity in a harmless member of a lethal family. Toxicon 142, 14-
957 29, 10.1016/j.toxicon.2017.12.042
958 Wullschleger, B., Nentwig, W., Kuhn-Nentwig, L., 2005. Spider venom: enhancement
959 of venom efficacy mediated by different synergistic strategies in *Cupiennius salei*. J
960 Exp Biol 208, 2115-2121, 10.1242/jeb.01594
961 Yacoub, T., Rima, M., Karam, M., Fajloun, J., 2020. Antimicrobials from venomous
962 animals: An overview. Molecules 25, 2402, 10.3390/molecules25102402
963 Yang, Y., Zeng, X.C., Zhang, L., Nie, Y., Shi, W.X., Liu, Y.C., 2014. Androcin, a novel
964 type of cysteine-rich venom peptide from *Androctonus bicolor*, Induces akinesia and
965 anxiety-like symptoms in nice. IUBMB Life 66, 277-285, 10.1002/iub.1261
966 Yi, H.Y., Chowdhury, M., Huang, Y.D., Yu, X.Q., 2014. Insect antimicrobial peptides
967 and their applications. Appl. Microbiol. Biot. 98, 5807-5822, 10.1007/s00253-014-

968 5792-6
969 Zeng, X.C., Nie, Y., Luo, X.S., Wu, S.F., Shi, W.X., Zhang, L., Liu, Y.C., Cao, H.J.,
970 Yang, Y., Zhou, J.P., 2013. Molecular and bioinformatical characterization of a novel
971 superfamily of cysteine-rich peptides from arthropods. Peptides 41, 45-58,
972 10.1016/j.peptides.2012.10.004
973 Zhang, L., Shi, W.X., Zeng, X.C., Ge, F., Yang, M.K., Nie, Y., Bao, A., Wu, S.F., E,
974 G.J., 2015. Unique diversity of the venom peptides from the scorpion *Androctonus*
975 *bicolor* revealed by transcriptomic and proteomic analysis. J. Proteomics 128, 231-
976 250, 10.1016/j.jprot.2015.07.030
977 Zhao, R.M., Ma, Y.B., He, Y.W., Di, Z.Y., Wu, Y.L., Cao, Z.J., Li, W.X., 2010.
978 Comparative venom gland transcriptome analysis of the scorpion *Lychas mucronatus*
979 reveals intraspecific toxic gene diversity and new venomous components. BMC
980 Genomics 11, 452, 10.1186/1471-2164-11-452
981 Zhao, Y., Chen, Z., Cao, Z., Li, W., Wu, Y., 2019. Diverse structural features of
982 potassium channels characterized by scorpion toxins as molecular probes. Molecules
983 24, 2045, 10.3390/molecules24112045
984 Zhong, J., Zeng, X.C., Zeng, X., Nie, Y., Zhang, L., Wu, S., Bao, A., 2017.
985 Transcriptomic analysis of the venom glands from the scorpion *Hadogenes troglodytes*
986 revealed unique and extremely high diversity of the venom peptides. J. Proteomics
987 150, 40-62, 10.1016/j.jprot.2016.08.004
988 Zhu, S., Gao, B., 2006. Molecular characterization of a new scorpion venom lipolysis
989 activating peptide: Evidence for disulfide bridge-mediated functional switch of
990 peptides. FEBS Lett. 580, 6825-6836, 10.1016/j.febslet.2006.11.040
991 Zhu, S.Y., Peigneur, S., Gao, B., Umetsu, Y., Ohki, S., Tytgat, J., 2014. Experimental
992 conversion of a defensin into a neurotoxin: Implications for origin of toxic function.
993 Mol. Biol. Evol. 31, 546-559, 10.1093/molbev/msu038
994

Table 1 List of the peptide fragments generated by enzymatic digestions

Peptide sequence	Molecular mass (Da)		Mass difference
	Obs.	Calc.	
Lys-C digestion			
LIFPGTK	774.46	774.47	−0.01
WCGAGDK	793.30	793.31	−0.01
AANYSDLGSAAETDK	2209.90	1511.68	698.22
CCRAHDHCDNIAAGETK	2016.74	2016.78	−0.04
YGLENSYWFTK	1406.63	1406.66	−0.03
CEESFRNCLAEAK	1614.64	1614.67	−0.03
FSYFNVYGPK	1220.57	1220.59	-0.02
CFVLDCD	929.30	929.32	-0.02
Lys-C digestion after PNGase A treatment			
AADYSDLGSAAETDK	1512.69	1512.67	0.02
Glu-C digestion			
LIFPGTKWCGAGDKAANYSDLGSAAE	3397.39	2699.25	698.14
TDKCCRAHDHCDNIAAGETKYGLE	2823.04	2823.16	-0.12
NSYWFTKLNCKCEE	1879.70	1879.78	-0.08
SFRNCLAE	996.40	996.44	-0.04
TTAKFVKFSYFNVYGPKCFVLDCD	2907.22	2907.36	-0.14

Table 2 Virucidal activity of LaPLA$_2$-1

Virus	IC$_{50}$ (ng/ml)[a]
HCV	2.0 ± 0.3
HCV (+ manoalide)	>500
DENV	3.4 ± 0.6
JEV	5.7 ± 0.8
HSV-1	>1,000

[a] The data represent means ± SEM from three independent experiments.

Table S1 Sequences of peptides identified by the transcriptome analysis of the venom gland.

Name	Sequence	Similar peptide (Accession number, species)
Non-disulfide-bridged peptide		
LaNDBP2-1	MNIKVVLVVCLITLLVTEQVEGGWFSRIKSLAKKAWKSNLAKDLRKMAGKAARNYAAKVLNVSPEEQVQLDNLLRYLD	Vejovine (F1AWB0, *Vaejovis mexicanus*)
LaNDBP2-2	MNGKVLVVCLIVAMLVMEPAEAGIWGWIKKTAKKVWNSDIANQLKNKALNAAKNYVAEKVGATPAEAGQMPFDEFMDIYYS	Con22 (L0GBQ6, *Urodacus yaschenkoi*)
LaNDBP3-1	MQYKTFLVIFLAYLLVTEEAQAFWWALAKGAAKLLPSVVGAFTKRKREIEKIFDPYQSSLDLEMERFLRQLQ	Heterin 2 (AGK88595, *Heterometrus spinifer*)
LaNDBP3-2	MKLRTLMVIFLAYLIVTDEAEAFWGFLAKAAAKLLPSLFSSKKEKEKREIEHFFDPYEKELDSELDQLLYELQ	Heterin 2 (AGK88595, *Heterometrus spinifer*)
LaNDBP4-1	MKIQLAILVITVVLMQMLVQTEAGFWGKLWEGVKNVIGKRGLRNKDQLDDLFDSDLSDADAKLLREMFK	OcyC2 (C5J887, *Opisthacanthus cayaporum*)
LaNDBP4-2	MKIQLAILVITVVLMQMLVQTEASFWGKLWEGVKSAIGKRGLRDVDQMDDLFDSDLSDADAKLLREIFK	OcyC2 (C5J887, *Opisthacanthus cayaporum*)
LaNDBP4-3	MKNQFVILLLAIVFLQLFSQSEAFLSALWGVAKSLFGKRGLKNLDQLDDLFDGEVSEADLDFLKELMR	UyCT3 (L0GCI6, *Urodacus yaschenkoi*)
LaNDBP4-4	MKAQIVLLVITVVLMQMFAQSEAGFWGKLWEGVKSAIGKRGLRNLDQFDDELFDSDLSDADAKLLKEIFK	UyCT1 (L0GCV8, *Urodacus yaschenkoi*)
LaNDBP4-5	MRLVSLTPLFLILLIAVDYCQSFPFLLSLIPSAISALKKLGKRSTDFQRQLDFQRRYLNSDLDFDLDELEEFLDQLPDY	VpAmp1.0 (ALG64974, *Vaejovis punctatus*)
Invertebrate defensin		
LaDefensin1	MKAIATLLLILVAFSVLEFGIVDAGFGCPFNRYQCHSHCQSIKRRGGFCAGTFRTTCTCYKSK	AbDef-1 (AIX87626, *Androctonus bicolor*)
LaDefensin2	MKAIATLLLLLVAFSILEDGIVDAGFGCPFNRYQCHSHCRSIGRRGGYCSGPFRFTCTCYK	AbDef-1 (AIX87626, *Androctonus bicolor*)
LaDefensin3	MKAIATILLLLAAFSILEFGIVGAGFGCPHNRYQCHSHCQSIGRNGGYCGGTFRTTCTCYNSGGPTSSP	AbDef-1 (AIX87626, *Androctonus bicolor*)
Potassium channel toxin-like peptide (alpha)		
La-alphaKTx1	MNAKFIYILLLTAVMFALYEASVPPNIPCQVTNQCPKPCREATGRPNSKCINGRCKCYG	Urotoxin (P0DL37, *Urodacus yaschenkoi*)
La-alphaKTx2	MNTKFVFLLLVISTLMPTFDASAEDISCSSSKECYDPCEEETGCSSAKCVGGWCKCYGCRG	Urotoxin (P0DL37, *Urodacus yaschenkoi*)
La-alphaKTx3	MNRNFVFLLLLIVTLMPMLDAATEDINCDNWRDCLKPCKDETGCPNSKCEEGNCLCYGCNRLTV	Urotoxin (P0DL37, *Urodacus yaschenkoi*)
La-alphaKTx4	MNKKFIFLLLVVTTLMPMFDAATEAISCSNPNDCREPCKKQTGCSGGKCMNRKCKCHRCNG	OcKTx2 (Q6XLL8*Opistophthalmus carinatus*)
La-alphaKTx5	MNAKLVCIVLLTAVMFAPDEASLPPIRIPCYVSKDCRKPCLYLTGTPRSKCINRRCKCYG	Hemitoxin (P85528, *Hemiscorpius lepturus*)
La-alphaKTx6	MNKPFCAIFLVVLIMFAVSVLPAESTGGCPVDSLCKSYCKSNKFGTEGKCDGTSCKCAIG	LmKTx8 (A9QLM3, *Lychas mucronatus*)
La-alphaKTx7	MNKLACYILICVMVSCLFKVPVAEGISAGCPLTAKLCTIYCKKHRFGREGKCIGPTRFRCKCYV	LmKTx8 (A9QLM3, *Lychas mucronatus*)

La-alphaKTx8	MRLVIILLLMTTLVLAVGAPLGGAKCSSSTQCTRPCRYAGGTHGKCMNGRCRCYG	St20 (P0DP36, *Scorpiops tibetanus*)
La-alphaKTx9	MELKYLLVLLAVTCLVSCQDNSLLPSGSCSRTGICMESCAPFLYQPKYHRRCPAGYVCCTLIY	Kbot55 (P0DL62, *Buthus occitanus tunetanus*)
Potassium channel toxin-like peptide (beta)		
La-betaKTx1 (LaIT2)	MAKHLIVMFLVIMVISSLVDCAKKPFVQRVKNAASKAYNKLKGLAMQSQYGCPIISNMCEDHCRRKKMEGQCDLLDCVCS	Previously identified
La-betaKTx2 (LaIT3)	MQAQFTVLLLLVLVTLCSCGGILREKYFHKAADALTSNIPIPVVKDVLKSAANQMIRKIGKVQQACAFNKDLAGWCEKSCQEAE GKKGYCHGTKCKCGKPIDYRK	Previously identified
La-betaKTx3	MDTKLSILVFLCVVVIASCSWISEKRIQKALDEKLPKGFIQGAAKAIVHKFAKNQYGCLADMDVKGSCDRHCQETESTNGVCHG TKCKCGIGRVY	SCl1 (L0G8Z0, *Urodacus yaschenkoi*)
Potassium channel toxin-like peptide (delta)		
La-deltaKTx1	MASQFLLFCIVLIAVNPLVYSKGGCRLPPETGLCYAYFERYYYDPSSNNCKMFVYGGCGGNSNNFVSKKACLARCAN	BmKTT-2 (P0DJ50, *Mesobuthus martensii*)
Potassium channel toxin-like peptide (kappa)		
La-kappaKTx1	MKPSTSAYALLLVLTFGIITSGVFAVPMDEENTFEVEKRGNSCMEVCLQHEGNVAECEKACNKG	HeTx203 (P0DJ34, *Heterometrus petersii*)
La-kappaKTx2	MKPSTSAFILLLVLTFGIITSGVSAIPMDEENTFEEQKRDSACVEVCLHHEGNVAECEEACKKS	HeTx203 (P0DJ34, *Heterometrus petersii*)
La-kappaKTx3	MKLLPLLVILIICALMANEAFCDQGARERSENLEDTRDLVQKPCRIVCSENMRKCIRRCTLGR	HelaTx1 (P0DJ41, *Heterometrus laoticus*)
La-kappaKTx4	MKPSSFAIALILVLFLGFTNAVSGEYAESISGDRMERAERAGCRIRCLQFTDDFEKCRKLCG	PcavC10 (AEX09227, *Pandinus cavimanus*)
La-kappaKTx5	MKLLPLLLVILIVCALLPNEAFCDQSAVERSESLEEVSREIVKRSCKRVCSGTRRTKKCMQKCKSQPGR	HelaTx1 (P0DJ41, *Heterometrus laoticus*)
DDH peptide		
LaDDH1 (LaIT1)	MNFATKVFFLLLAVAVIAIVAGEEDDSWFEQNEESDTERDFPLSKEYETCVRPRKCQPPLKCNKAQICVDPKKGW	Previously identified
LaDDH2	MNCAIKVSFLLLAITVIFSVAGGEGDNSFEQREENDTERDLPLSKKHESCVRPRKCRPPLRCNKAHICVESKKGWRIPVISWVS KTKIMFSE	LaIT1 (P0C5F2, *Liocheles australasiae*)
Sodium channel toxin-like peptide		
LaLAP1	MNTGEHFASLIIFLLLLGENPCLGDGGWPMRIDGNYYLCYYEEKPAELYCKRACKLHKASQSSCSYHWKYMSWYCYCEDLKK GYTRDRGLKKGGGHQFSES	LVP1-alpha (P0CI45, *Lychas mucronatus*)
LaLAP2	MLMVIYIATLIPLILQGECRAKDDHPRNFEGNCYRCKYPDKSGYCEAICKMHKAETGYCSRSNLFCYCKGIEDKYVSARDFLEP	BmLVP1-alpha (Q6WJF5, *Mesobuthus martensii*)
La1-like peptide		
La1	MGGTLKHLLLVCLIVVCSSSLCLGFGESCIAGRFIVPLGQQVTDQRDCALYKCVNYNKKFALETKRCATVNLKSGCKTVPGGA GAAFPSCCPMVTCKG	Previously identified
La1-1	MKHLHAAVLLVCLSICALPSLTLGAGESCKVGSLVIPVGQKKFDKPSCAEYECSTEFNRVLLKAITCATLNTKGKGCKSVPGKS PNSFPDCCPTILCRGEQWNH	La1-like protein 15 (L0GB04, *Urodacus yaschenkoi*)

La1-2	MKIVCTLVPLVFACIANAHVLTERTCRTHTGVILKNGEEWADPNHCSIYRCTIYSGEAELQGLTCATYRVPPNCEFVRGRGKFY PSCCPTVMCKP	OcyC11 (C5J895, *Opisthacanthus cayaporum*)
La1-3	MALRFVLAFLLPCLVLGSNEPAKFISYKNDMLGPLVEGKCKVGNDKMIEQGGTWYRDDYCEKIYCLRTGTLGHMEVWGCAPV APLNPNCTVVHHSGLYPNCCSGEIVCEENPDTKKSDVEMAEIIRALLQSERK	MeVP-7 (ABR21061, *Mesobuthus eupeus*)
La1-4	MFNIVTITLLWSCTCIALCHSYGETCQAGELTIPLDNEKQDPEACTLYKCTMFAGRVVLNTLTCAPQEPRRGCRNVDSPVELPF PDCCPLVVCNVQPLGSK	HsVx1 (K7WMX6, *Heterometrus spinifer*)
La1-5	MRTLVPAVFLLALIVAAMASHKDPYHRNCPIGNKDLSDGEEWADQQRCVKYKCQIRGPDAALLLTRCPSVGIYPPDKCREIAG KGNFPDCCPKLQCD	MmKTx1-like (XP_02321219, *Centruroides sculpturatus*)
La1-6	MSRADKAFVGTLIAIFLVCSLTNAYSSLKRQKYGSGAPCVDHLGSNRKFDDVWYDNESCEEHTCIKYRGIPHVQIYGCIAAVAS PGCELVKGSGSYPDCCEEEIC	TxLP9 (ABY26691, *Lychas mucronatus*)
La1-7	MKTWEARFYIFLVALVTITFVESYVYLVPQDPGAVVCIGKDRVSHKPGDVWYDDSKCERLTCGHSSGGLVIDGAGCGSISVQD GCKLVPGVGSYPSCCPSPVC	Toxin-like protein 14 (L0GCW8, *Urodacus yaschenkoi*)
Serine protease inhibitor-like peptide		
LaAPI1	MKSWTYRLCFLILFLVCVNCTRPVLSPEECTRPGEFFTTAGSYCPLTCDNYKNPPIICSLIGIVGCECSPELVRDERSGNCVDTS DCTED	SjAPI (P0DM55, *Scorpiops jendeki*)
LaAPI2	MKGFAFLVLTFVVVFGDKEQECEDPNAEFRRCNTACPITCANMDNPPNICTLQCVIGCACKEGYYKNDDGLCVHPEGC	SjAPI (P0DM55, *Scorpiops jendeki*)
LaAPI3	MKALLLLLSFVVIHSAKSQEDLGGDEHAQSCLLPNEVWDKCGPSCPPSCIGVIEPGTLCSTECTPGCFCREGLVRTKRGSCIPP KACRNERETKL	SjAPI-2 (P0DM56, *Scorpiops jendeki*)
LaAPI4	MKGSKPLCFLYLVVLLWTSVKCTRWASSSEECTRPGEAFTSCGTDCPITCANYENPPEFCNYMCVIGCECTSGLVRDEGSGN CVNPSQCGP	HtPi1 (AOF40217, *Hadogenes troglodytes*)
LaAPI5	MKRNLVLLAFVLLLFGSIFEKCSAQRGGADRRCRRPGEVFMNCGPSCPLTCDNYQNPPTVCTLQCVIGCFCRRGLVRDTRRG GCVRPSQCRR	HtPi1 (AOF40217, *Hadogenes troglodytes*)
LaAPI6	MKRNLALSALFVLLFCSALDNCEAQRGGYDRTCQLEGEVFTRCGTACPLTCDNYKNPPEVCTLQCIIGCVCGRGWVKDTRRG RCVRPSQCRR	HtPi2 (AOF40218, *Hadogenes troglodytes*)
LaAPI7	MKSNLVLFAFVLLLFCSVLEICTAQRRVVDRSCRGAGEVFTRCGTACPLTCDNYRNPPRFCTRQCIIGCACRRGWVKATRRG GCVRPSQCRR	HtPi2 (AOF40218, *Hadogenes troglodytes*)
LaAPI8	MAKTLAFGVVLCMFVLAQSAPQYPGTFGCDEDKQFVRCLPPCPVTCKSILNTTPCTLLLPRCTPGCGCRGGKILDNAGKCVFP ADCRRG	CtAPI (P0DM57, *Chaerilus tricostatus*)

Table S2 Sequences of proteins identified by the transcriptome analysis of the venom gland

Name	Sequence	Similar protein (Accession number, species)
Phospholipase A_2		
$LaPLA_2$-1	MVFIFLAVLSGLVTLSHSTAVQREMHVHFEPLPGQRDSWPVARAALVNLATKSETGREFSDCRMLNSIDEIAREGAVLSRYEIK RVSKEEMRSLEKRCSRSSGIHQRLIFPGTKWCGAGDKAANYSDLGSAAETDKCCRAHDHCDNIAAGETKYGLENSYWFTKLN CKCEESFRNCLAEAKKKETTAKFVKFSYFNVYGPKCFVLDCDKRRFEMSRKCVAHWKESRRG	Hemilipin (A0A1L4BJ46, *Hemiscorpius lepturus*)
$LaPLA_2$-2	MMVSSLLIVLIVTSAVQCYVIDNMNDEEPLVTFYREKDGRRTVEVIEVNDNKKGPKIVGCVEYGDGYIADMVLNLSRNILVRDVN RQQMDEVVNRCRERETRDLRNEVINLFKSPAETSRNAFQSLMIFPGTKWCGAGNVSENYDDLGTARATDMCCRDHDHCND SIERFGSKHGLENTDFYTKSNCDCDNKLYSCLEASKDVTSDLVGYLYFNIMQTQCFKKYYPEVKCLKKTGILFFMQSCQEYEF NRNEPKKYEFFDAKYYEPQAASLEQIMSYFYSSSSQ	Phaiodactylipin (Q6PXP0, *Anuroctonus phaiodactylus*)
$LaPLA_2$-3	MSLQTLAVLLLSFIQPLPTAVIELPHENKLTGYYQNERSPYMLIIGQTGKVIHCHQYEDKNEADRVLAALKLEDVERVTKEQMDK LIKFCTEEEHMGHPKEQVKMLIYPGTKWCGMGNNAASENELGTEKEADSCCREHDHCSDSIPAFSIKYNLTNYSPFTKSNCNC DREFRQCLTKAGTKASEIISGLYFDLLKMECFGRTSCSSNDACTEGWQWKRSTSF	Hemilipin (A0A1L4BJ46, *Hemiscorpius lepturus*)
Serine protease		
LaSP1	MKYASLIASVFLLTQTEACTETGEHNQIRFPEEEEESCRTPNGRRGVCVPLNACPEFRNADDRYIRQSICWFDRNTPVVCCPS NEQPVTMPTHPTRPTRPTRPTRPTQPPTCPPVSPVTRCPVVGPNRRKPSILPEECGKSTVPLSRIVGGRKSDLGAWPWMVAV YLTRAGLNRGTDCGGSLISDKHVLTAAHCVYDKRRKTVMSASQLMVRVGEHTLNDDNDGASPIDVPVSNIMAHENFERKTFK NDIAILVLANTVQFSQFIRPICLPYDESSEANFTGRSAFVAGWGETEYEGQFNPELSEIQIPILNNEVCRQKYKRNIPITAEYLCA GVSDGTKDSCRGDSGGPLMLPEKDNRFYLIGVVSFGKRCATYGYPGVYTRMTMYLDWLASKLS	Serine protease 1 (AMO02563, *Tityus serrulatus*)
LaSP2	MWARFGLTFLFCYYLQNTFIGAEAQACGTRNMTLEFKIVGGTVAARGEWPWQVSVQLTHPQFGKIGHWCGGVLVGQQWVA TAAHCIINPLFSLPQPVFWKVLLGDHHIKKTEGSEVVIGVSRVYYNPWYHGYQNDIALLKLSEPVKLSSYIQPICLPTSNDGFQD MTCTATGWGKTDFNMKASDTLQKVDVKVLDNSICANAYLTQFKIPITPSHLCAGDTAGGKGTCLGDSGGPLQCLMPNGKWYL AGLTSFGSGCAKPGFPDVYTRVTYYVDWIKQNQLLPW	Serine protease 1 (AMO02563, *Tityus serrulatus*)
LaSP3	MNLIGLVALACVTLSSRVEARFLHDPSEESGRVFRGRFANQQEFPWMVHLQISKGNNMASVCGGSVISKNWVLTAAHCVCRN ATLKTYADVNGITGRIGHVNKSSATAVKFSQLIVHNSYDEDFNADIALLKFKDSLKKYDANVNRICLADKGKTFPNRQPVIQMG WGRFDNSSAGTSPSLKTTSVGYILHRADCIKEQESYADPGQICVSNFKGEKICGGDSGGPLVVANGAEKLDIGIVSYDYYSFCV AGSDSPAMYTEISYYADWIKTKTNDKEICWKK	Chymotrypsin-like protease-3 (ABR21040, *Mesobuthus eupeus*)
LaSP4	MLWSTRDSAIFLLFIITLNSRFNYSSNQGFFLNRRSDGNTCTRNGEVHQCQFFLFCLLGGGTSMGSCSGRVLTTCCAKPNLRR SRPSNFRQRALNNKDRASCGQTAWKPQSRIVGGQDALYGEFPWQAHIKIIQQQCGGVIVSPYFVVTAAHCVYRARLHQITVVL GAYDIHDQSFQLQPALFLRVDEKRLHPNFKFSPSHPDRYDVALLRLHRRVQYQENILPICLPPYKWNFRGWRAVVIGWGKTDP ALRNRYGTRLLQKVEVPIISNDECEYWHKSRGIKLKIYPEMICAGYEHGKKDACVGDSGGPLMVNMRGKWTLVGITSAGFGCA QWRQPGIYHSVSSTVDWINANIR	Serine protease 3 (AMO02565, *Tityus serrulatus*)

LaSP5	MAISLLVKYAILWISLAVLSSAQSALPDRRPNSFVFPETVRDSPRCRTADGSPGSCLKASECRDVNFQRGTLPLLCYWEDNQPI VCCADRSAAISSENLSEPQTGCGKSERKPTTRSPTIAGGWISQPSAWPWMVAILTSNLGEKFLCGGTLVSQKYVLTAAHCFRR NGVDQRRIPVARFLVRVGSTENNQGTAYRIRRIMIHEDYRVNQHYNDIAVIEVNDPISLSSSVRPICLPSSELQGRSVVGREVVV VGWGDQSFGGIRDNKLREVNISVIDRDTCNEAYNELSSRSIPNGITSQFMCAGDPEGGKDACQADSGGPLMMFSPSQWSIVG VVSFGYGCAHKGFPGVYTQVSSYLNWIKDKTDL	HLClotting-factor1 (API81376, *Hemiscorpius lepturus*)
LaSP6	MAMKYFVLFISTNALLAASFLPAKEENRIFKGREANEGEFPWMVFIRLTDELNCSGFLVSHNYVVTAAHCMIRSVTDMKGVVGS VDREQDNMLEFEKFVIHPEYNESTFHGDVALLKLKRPLEFTSLIKPICIGKKKSFINPGNEVLQMGWGRDRNDSAVVSKILKVTN VGSVMSEEHCHTFLGMINFTSIGRICVKNGEVEGVCEGDSGGPLVYKDSEHGNVAVGLSSFGFYLNCSVTNENPEIFTSTAYF SNWIAENVEDSVCVIG	Chymotrypsin-like protease-1 (ABR21038, *Mesobuthus eupeus*)
LaSP7	MFKSFELIYLAIACIGSGSVIFAKNCDDCILITSCPGAVYLAVHAKNAKTEDLIKHSLCSLEKVNGLPKVCCSEFPPAPQLDNHPN LELLPKDCGEIEGSRIVGGEVAKLYEFPWMVLISYDTRIGREFLCGGSLISPLYVLTAAHCVHGRKIAGVRIGDYDWRSKIDCEK DTNLCESYYQDIGVSERLPHPDYQGPPVVRNDIALLRLRRPVNLTVKNAGVICLPVTKELRERRLDTEQVTVAGWGITENNTAS SVLLKVNLPVHSGEMCRAYYGRNSKEDTTKNILCAGVLGKDSCKGDSGGPLMLEGNYDNVFKFIQYGIVSYGPSQCGSNFPG VYTDVSSFMKWILDNIKP	HLClotting-factor2 (API81377, *Hemiscorpius lepturus*)
LaSP8	MKCFAFLIFSSQLFPAVPFQVEEKTRIFGGREANDGEFPWMVFIRLSAEWNCGGFLISPSYVLTAAHCVKGSSVTDMRGVVGS VDREQQDMLEFKKYQIHPEYGPKRRWNADLALLELRTPLGFTDLIKPICIGKKTSFTRPGNAVLQMGWGRDREDSAVVTKKLK VTEVGNLMSRCDCHRFFESINIVDIQLNGRLCVKNREVEGVCEGDAGSPLVYQYAESSHVAIGIVSAGFYVNCSVTNENPEIYT DLAYYSDWIIRTVDEPICIIH	chymotrypsin-like protease-1 (ABR21038, *Mesobuthus eupeus*)
Alpha-amylase		
La-alpha-amylase	MIVDCLLLWFWVSVVHCSYHEPNTQAGRSVLVHLFEWRWKDIAEECETFLGPYGFGGVQVSPANENGIIWEPHWNSVIRRP WFERYQPVSYKIATRSGNESEFRDMVRRCNNAGVRIYVDAVINHMTGDIGRGKGTGGSDFDPGALQYYGVPYGPSDFNGRD QCPSGSGDIENYQDKYQVRNCRLSGLADLNLSKEYVRDKIVEYLNFLIDIGVAGFRVDASKHMWPGDLKIIYEKLKNLNTEYFP VHRRPFIYQEVIDLGGGEAVKAEEYTDLGRVTEFRFGKHLGDVIRKNYNQRLKYLKNFGQDWGMVSGSNALTFIDNHDNQRG HGAGGFGSILTFFESRMYKMAVAFMLAWPYGLPRVMSSYNWPRYVENGKDKNDWIGPPHDDEYNTKPVVRNPDLTCGNGW VCEHRWHQIYSMVKFHNVAGFEPVDFWWDNDYQQIAFGRKGKGFLAINNENHNLDQTLPTGLPPGTYCDVISGKLEGDKCT GRSVKVEQDGKAKIFIDNSWEDPMLAIHVEAKLNDVLHNGNTGRNRNG	Pancreatic alpha-amylase-like (XP_023225708, *Centruroides sculpturatus*)
Cysteine-rich venom protein		
LaCRVP1	MMHFILTCILFPQVIHLNGVVAKEIKYIFYSESDYCVLGERDCNTSIYAEDFLWPHNSNRAYMASGDVPELSASNMLKLEWDQN LANIAQRAAEQCIDDYPPTFHGKVECRQPDSSAVINFSWKQSQNLENVAKRIEERIFEWTEYIKYRYNDSSLHFYRGTGIFEDL WAKVVQATTWKVGCGLADSDAVNHDMEHHYTEVISCLYENTKLQPGDEIYKLGSPCSACPLGTRCIPGSNLCEVEPGNCPIP GNAKQDCAKNTERNTNRETLWKCDLKKYGEECEIVPSCSLLWSVDQQGNFKSISVTRDCVSANVFIKRIRIGKPSCFTFQYIKE GNRNEPSDTVVMGMVFNLESGDIIQVAKGEDATTWTDVMIDIPWTGVDIQVGVVARSFNDNFRKQQILMKDFHVSDSACLRD TEHFKPNLH	HLAllergen3 (API81354, *Hemiscorpius lepturus*)
LaCRVP2	MFIPLALLTAIRLSLRDRCCYAEEFLANPSVVWETETVCPTTPTYAGSVKSKRNANPGTPIPLSDQDKMEIVNEHNKYRGQVSP PAANMRYMIYNEDVAYAAQLWANGCQFRHGRPQGSKFQKLVGQNIYKGPSGSYSFYMRLWHNEVRDFDMKSQQCNATMC GHYLQMVWAESFLIGCGQSHCRGKRRSYNLFVCHYYKAFESFHPYKVGRPCSMCDAEAGGFCYNNTCVSKEDCEREGWPC	HLAllergen1 (API81352, *Hemiscorpius lepturus*)

	ECNLKCHNCGILDESTCTCKCPPGWDYQDCSQNCSDSSDYCGKEDGFEGWVGCEMLEDGYVKTHYCRKMCETCEVITEGN KEKTCCGGELCDKGYVLDNSNCNCRRLCPGPECYFTYEDNGVAYRVTNSLLTVIIAGIYALWQGQEPPKG	
LaCRVP3	MEKNVLWLIVLHLSLIWVKSQRDVGIAKFTGNLSSAWVHKLRVDYRLKLRRVARLASMMNDSDRLEVIRLHNLYRSMVAPPAA DMEYMAWDDRLASLSQQWAEECKWRHGNPRHDFPTGLGQNLYRGSVRSVALAIRLWYEEHVDYRYHNLSCNPKKQCGHY TQMVWSKTHLVGCGVKNGCWKGFKHYIVCNYWPSHYKGEKPFEIGRPCSLCNSTKSGLCWNKSCVSRSQCEEYKLDCSCD LICHNCGKLDREKCTCICKDGWKGVDCSEPCVDTMEDCSLYHCIYHEYSRNHPCLKSCNVCKPVNADDLRNSCCDGVSCPH GYVLDLADIPCECKPLCPGPKCGSLLHGPYFTLMLFVLLLLKCSAL	HLAllergen2 (API81353, *Hemiscorpius lepturus*)
LaCRVP4	MEWTFVCSIMLSFAPLTWGKTQFETERINVSYSWVHYGGMSRRARATSMMKDPDKLEITRLHNSYRSMVSPSAVDMEYMEW DDRIASVSQQWADRCTWKHSKDTAIADFPDGLGENLYRGWSSSPAYAMNLWYSEYTHYDIQNTSCKPGKVCGHYTQMVWS KSRLVGCGVKEGCWRDYRYYIVCNYSPPGNYKGEKPYQIGRSCSHCHSGSGLCSNKSCVDRSQCDKYNLDCSCDLVCLNC GELNRHSCTCKCKEGWKGVDCSEPCQDAQENCLVSACSYYERWPRGSHPCEKTCNVCKSVTPDSVQNTCCDGVTCPPGEI LDLSERPCECKKLCSGPKCQSSLHEPHSLFLLVAMSIFLRCYAR	HLAllergen2 (API81353, *Hemiscorpius lepturus*)
LaCRVP5	MAAVAVALVVLWITVTSGSANNDTCAERYTNITPEHTMCKSTNENCTFVRSSGEVFEEQLLRTHNLIRNSIRKYVGKKYHLATN MKLMVWDDELYEMARLHSLQCVEEPDCDLCHQIGDFPVEQNFAVKTFKSSEVDGNGPIRRLQAVIEEWAAELRSYDCDVVKV FRNTEGLPTNWTNIFRATTMLVGCASMTFLTDEPGTFKEVYVCNYGPANLTEGEEIYRTGRKPCSECEDDEGCDTEFKHLCFP ADMEEENIPEETELHNTFARNMVTKKFPRQRRRYLWNTDVARSYLDRRTENWRTTGGGVTGSYPDYRTGNWGNTGEGVTR SSHTYQCKRKSETSQLSESSILVMEQQGF	HLAllergen7 (API81358, *Hemiscorpius lepturus*)
LaCRVP6	MEIFLAIPCLLTAFLPSISNERCSDSFGNYDQALFREGTRDESVQISWWTTTPISKETGDKTRTSEEKECPELYQRYSRNHTFC QVSNCDIIRKGVTEEDKNIILEFHNSLRSKLASGMEKRYCSLPSAANMMQIEWDDELAAVAEAHAELCVYGHDRGKRAVESFS VGQNLMLYGGDIKRWGDADGWYKEEVCFYSPEKNSPFRSGIYGHFTQVTWATTWKIGCGFASYRKNGKVEALYTCNYGPS GNVREGRHYIVGEPCSQCPPNTECSTEDPGLCKSKTCDGPQMRRPPSEDFILFCDFSHEDSEECNKVKVNGSREFSTRHIYT GNYKTVVLNAGESITIDLGKAQNDGGICTFVYSRFGPNNAKDAPGSVMEIKYESPNILPVPPRTIGPYGPTFFMAGTHMSYHGE LQNTLTLRALEGAEPQYFDIKMWGIRKGDCLMSLESD	HLAllergen2 (API81353, *Hemiscorpius lepturus*)
LaCRVP7	MEIFLAIPCLLTAFLPSISNERCSDSFGNYDQALFREGTRDESVQISWWTTTPISKETGDKTRTSEEKECPELYQRYSRNHTFC QVSNCDIIRKGVTEEDKNIILEFHNSLRSKLASGMEKRYCSLPSAANMMQIEWDDELAAVAEAHAELCVYGHDRGKRAVESFS VGQNLMLYGGDIKRWGDADGWYKEEVCFYSPEKNSPFRSGIYGHFTQVTWATTWKIGCGFASYRKNGKVEALYTCNYGPS GNVREGRHYIVGEPCSQCPPNTECSTEYPGLCKSKTCDGPQMRRPPSEDFILFCDFSHEDSEECNKVKVNGSREFSTRHIYT GNYKTVVLNAGESITIDLGKAQNDGGICTFVYSRFGPNNAKDAPGSVMEIKYESPNILPVPPRTIGPYGPTFFMAGTHMSYHGE LQNTLTLRALEGAEPQYFDIKMWGIRKGDCLMSLESD	HLAllergen2 (API81353, *Hemiscorpius lepturus*)
LaCRVP8	MDLLFAISCLLAAFLPCICNKRCSNIYGKYDQGTLSKGARDETFLSSWRNVIPINIERGEKTQFSERKECPELYQRYSADHTFCK KSTCSIIQKGVTEDEKNTILKFHNSLRSKLASGKEARYSKTPLPSAANMMQLEWDNELAAVAEAHAELCLYDHDANEQRAVEN FPVGQNLMQYNGDIRKWGDADMWYKDEVCFYSPQYNSPFNSGDFGHFSQVTWATTWKIGCGFTSYRKDGSEKALYTCNYG PAGNVPGGRHYIVGKPCSQCPPNTECSTEYPGLCKSKTSDGPQMKRPSSEDFILFCDFSQADPEECKKVKVSGSREFSTRHI YTGDYKTVVLNAGESISIDVGKAQNTGGICPFVYARFGPNNAKDPVGSVVEIQFSAPNIVPMPPSTVTPLGSFFVAGTHMMYD GELQSVLTLKAEQGAKPQYFDVKAWGVRKGSCGTSLGPK	HLAllergen2 (API81353, *Hemiscorpius lepturus*)
LaCRVP9	MEVTSRPVTEANPVIGQPLLTWKWTHLYICARKRIRAIINLHFLTVETANMNLTLVVSILCLALFTCQVVYSQRCPEIYQRFSEDH TYCKHSTCQVIKSGVTDQDKKIILDMHNSYRNKVALGQEDTPQRQPPAANMLQMEWDDELARIAQAHANLCKFEHDSPDQRQ	HLAllergen1 (API81352, *Hemiscorpius lepturus*)

	VENFNVGQNLFISMMTQVIDWNKTAMWYTWEIKHYYPQYREPFASGPYLHLSQMIWADTWKVGCGVAVYYDNNERRDKVLY TCNYGPGGNQIGQKVYTAGKPCSQCPKNTQCSSEYKGLCKSRTPDGPQQDTTRNPDDFLLYCDFSNNEDRACKNVQISGLR QFETRKIYGGEYKIAILKGGESITFKLGKAKDSRGICPFIYGSFGPNRAGDAKQSAVSIGFAASGLIFGDPIKIDYGSSDFWPIGM HMQYDQEMESTIKLEAYPGAPPQYFKIKAFGIGKGKCPKF	
LaCRVP10	MDFLLAISCLLAAFLPCNSNHLCSELYGEYDEDSDEDYPSSEWEEIPTDEERGEKSQISERKECPELYQRYSTNHTFCKESTC DIIRRGITEEDKNTILEFHNSLRSKLASGKEARYSQTPLPSAANMMQMEWDDELAAVAQAHAELCIFNHEPDDQREVENFPVG QNLMQFDYDIKSWESAEIWYKEEVCFYSEEYNSPFNSGIFGHFSQLTWATSWKVGCGFASYRVNGSEKGLHTCNYGPPGNV AGGRHYIVGEPCSRCPPNTECSTEYPGLCKSKTSDGPQMKRPSSEDFILFCDFSQEDPEECDGVKVNGSREFSTRHLYIGDY KTVVLNAGESISIDLGKAQNTGGICPFVYARFGSNNAKDPAGSVLEIQFDAPNMWSMPPIRVNPYGDSFQVAGVQMMYDGEL QSILTFKAEEGAKPQYFDVKAWGIRRGSCETSLDPEDEN	HLAllergen2 (API81353, *Hemiscorpius lepturus*)
Insulin-like growth factor binding protein		
LaVP1	MGKFLLIALFLFGVTVSALGLSCRPCGTYQCPPLPRCPVGVVKDACFCCQVCAKGLNERCGGPWNISGRCGRGLKCFKQAQ DAIGVCRKV	Venom insulin-like growth factor binding protein-1 (ABR21044, *Mesobuthus eupeus*)
LaVP2	MALRFCFITLLLLGVILGAMTLRCRQCGTYECPPAPENCPVGKVKDICNCCDECGKNLGEECGGAWDMYGKCGKGLRCFKEP VEGDPFNAKGTCR	Venom insulin-like growth factor binding protein-1 (ABR21044, *Mesobuthus eupeus*)
LaVP3	MWFRLVLFCVLVTSIYSLSCPCHELDLEEFCGPPPKDCPLGLTLDACACCQVCFLTEGEVCGGPWDVNGKCGAGLTCVKPPG QSDFVFDQSDGVCKKQ	Venom insulin-like growth factor binding protein-1 (ABR21044, *Mesobuthus eupeus*)

Table S3　Comparison of inhibition of HCV infection between pretreatment and post-entry treatment with $LaPLA_2$-1

$LaPLA_2$-1 (ng/ml)	Inhibition (%)	
	Pretreatment	Post-entry treatment
100	>99.9	8 ± 6
10	92 ± 1	<0.1
1	16 ± 23	<0.1
0.1	<0.1	n.t.[a]

[a] n.t., not tested.

Table S4 Cytotoxicity of $LaPLA_2$-1.

$LaPLA_2$-1 (ng/ml)	Cell viability (%)
1000	79
100	100
10	100
1	100

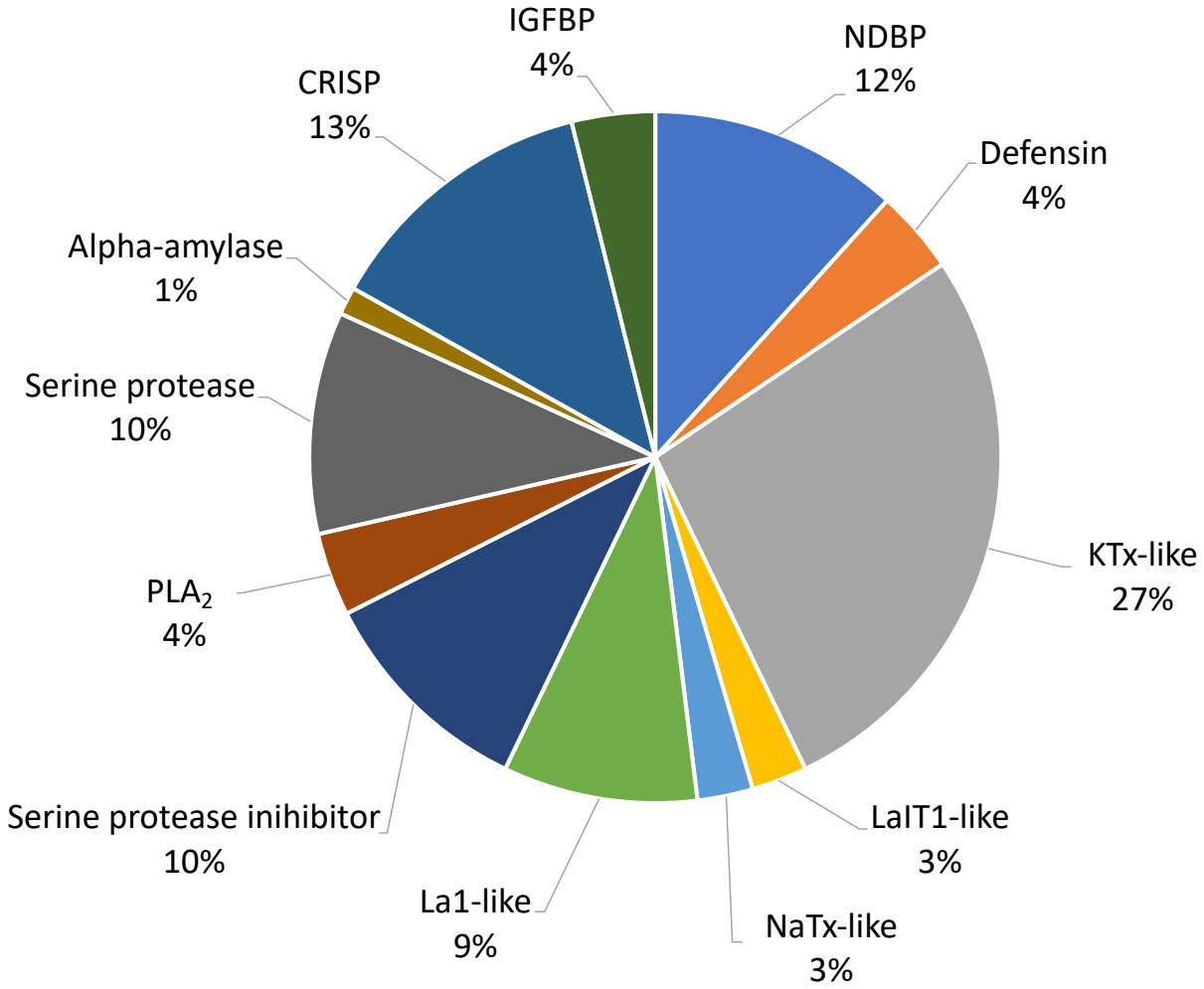

Figure 1. Proportion of the components identified by venom gland transcriptome analysis

(a) NDBP

```
LaNDBP2-1   MNIKVVLVVCLITLLVTEQVEGGWFSRIKSLAKKAWKSNLAKDLRKMAGKAARNYAAKVLNVSPEEQVQL--DNLLRYL--D    100(%Id)
LaNDBP2-2   MNGKVLVVCLIVAMLVMEPAEAGIWGWIKKTAKKVWNSDIANQLKNKALNAAKNYVAEKVGATPAEAGQMPFDEFMDIYY-S    37
Vejovine    MNAKTLFVVFLIGMLVTEQVEAGIWSSIKNLASKAWNSDIGQSLRNKAAGAINKFVADKIGVTPSQAASMTLDEIVDAMYYD    42

LaNDBP3-1   MQYKTFLVIFLAYLLVTEEAQAFWWALAKGAAKLLPSVVGAFTKR-KREIEKIFDPYQSSLDLEMERFLRQLQ    100
LaNDBP3-2   MKLRTLMVIFLAYLIVTDEAEAFWGFLAKAAAKLLPSLFSSKKEKEKREIEHFFDPYEKELDSELDQLLYELQ    56
Heterin2    MQYKTFLVIFLAYLLVTEEALAFWGALAKGALKLIPSLVSSFTKKDKRALKNIFDPYQKNLDLELERLLSQLQ    75

LaNDBP4-1   MKIQLAILVITVVLMQMLVQTEAGFWGKLWEGVKNVIGKRGLRNKDQLDD-LFDSDLSDADAKLLREMFK       100
LaNDBP4-2   MKIQLAILVITVVLMQMLVQTEASFWGKLWEGVKSAIGKRGLRDVDQMDD-LFDSDLSDADAKLLREIFK       90
LaNDBP4-3   MKNQFVILLLAIVFLQLFSQSEA-FLSALWGVAKSLFGKRGLKNLDQLDD-LFDGEVSEADLDFLKELMR       48
LaNDBP4-4   MKAQIVLLVITVVLMQMFAQSEAGFWGKLWEGVKSAIGKRGLRNLDQFDDELFDSDLSDADAKLLKEIFK       80
OcyC2       MKTQFAILMIAVVLMQMLVQTEGGILGKIWEGVKSLIGKRGLKKLDQLDD-TFDSDLSDADVKLLREMFK       78

LaNDBP4-5   MRLVSLTPLFLILLIAVDYCQSFP-FLLSLIPSAISALKKLGKRSTDFQRQLDFQRRYLNSDLDFDLDELEEFLDQLPDY      100
VpAmp1.0    MKLINLVPVFFVLIIVVDYCHSLPFFLLSLIPSAISAIKKIGKRSVESQRYVDLNRR----DLEQDLQELQDFLDQISEH      58
```

(b) Invertebrate defensin

```
LaDefensin1  MKAIATLLLLILVAFSVLEFGIVDAGFGCPFNRYQCHSHCQSIKRRGGFCAGTFRTTCTCYKSK------       100
LaDefensin2  MKAIATLLLLLVAFSILEDGIVDAGFGCPFNRYQCHSHCRSIGRRGGYCSGPFRFTCTCYK--------       85
LaDefensin3  MKAIATILLLLAAFSILEFGIVGAGFGCPHNRYQCHSHCQSIGRNGGYCGGTFRTTCTCYNSGGPTSSP       74
AbDef-1      MKTIVLLFVLALVFCTLEMGVVEAGFGCPFNQGRCHRHCRSIGRRGGYCRGIFKQTCACYRK-------       55
```

Figure 2. Multiple sequence alignment of NDBPs (a) and invertebrate defensin peptides (b) identified in this study and comparison with their similar peptides. Mature regions that were previously reported are underlined. Cys residues in the reported and putative mature regions are shaded in yellow. %Id represents the percentage of sequence identity.

(a) Alpha-KTx

```
La-alphaKTx1   MNAKFIYILLLTAVMFALYEASVPPNIPCQVTNQCPKPCREATGRPNSKCINGRCKCYG------      100(%Id)
La-alphaKTx2   MNTKFVFLLLVISTLMPTFDAS-AEDISCSSSKECYDPCEEETGCSSAKCVGGWCKCYGCRG---      41
La-alphaKTx3   MNRNFVFLLLLIVTLMPMLDAA-TEDINCDNWRDCLKPCKDETGCPNSKCEEGNCLCYGCNRLTV      42
Urotoxin       MNAKLIYLLLVVTTMMLTFDTTQAGDIKCSGTRQCWGPCKKQTTCTNSKCMNGKCKCYGCVG---      48

La-alphaKTx4   MNKKFIFLLLVVTTLMPMFDA-ATEAISCSNPNDCREPCKKQTGCSGGKCMNRKCKCHRCNG         100
OcKTx2         MNAKFILLLLVVTTTMLLPDTQGAEVIKCRTPKDCADPCRKQTGCPHGKCMNRTCRCNRC-G         61

La-alphaKTx5   MNAKLVCIVLL-TAVMFAPDEASLPPIRIPCYVSKDCRKPCLYLTGTPRSKCINRRCKCYG----      100
Hemitoxin      MNTKFIFLFLVVTTTMLLFDTTAVEA--IKCTLSKDCYSPCKKETGCPRAKCINRNCKCYGCSG-      48

La-alphaKTx6   MNKPFCAIFLVVLIMFAVSVLPAES--TGGCPV-DSLCKSYCKSNKFGTEGKCDGTS---CKCAIG      100
La-alphaKTx7   MNKLACYILICVMVSCLFKVPVAEGI-SAGCPLTAKLCTIYCKKHRFGREGKCIGPTRFRCKCYV-      44
LmKTx8         MNKVCFVVVLVLFVALAAYVSPIEGVPTGGCPLSDSLCAKYCKSHKFGKTGRCTGPNKMKCKCLV-      49

La-alphaKTx8   --MRLVIILLLMTTLVLAVGAPLGGAKCSSSTQCTRPCRYAGGTHGKCMNGRCRCY--G           100
St20           MKMSIVIILLLFTCLIATNGA--SGTKCSGSPECVKFCRTKGCRNGKCMNRSCKCYLCS           49

La-alphaKTx9   MELKYLLVLLAVTCLVSCQDNSLLPSGSCSRTGICMESCAPFLYQPKYHRRCPAGYVCCTLIY       100
Kbot55         ----------------------AGSMDSCSETGVCMKACSERIRQVENDNKCPAGECICTT--       39
```

(b) Beta-KTx

```
La-betaKTx1(LaIT2)  MAKHLIVMFLVIMVISSLVDCAKKPFVQRVKNAASKAYNKLKGLAMQSQYGCPIISNMCEDHCRRKKMEGQCDLLDCVCS

La-betaKTx2(LaIT3)  MQAQFTVLLLLVLVTLCSCGGILREKYFHKAADALTSNIPIPVVKDVLKSAANQMIRKIGKVQQACAFNKDLAGWCEKSCQEAEGKKGYCHGTKCKCGKPIDYRK

La-betaKTx3    MDTKLSILVFLCVVVIASCSWISEKRIQKALDEKLPKGFIQGAAKAIVHKFAKNQYGCLADMDVKGSCDRHCQETESTNGVCHGTKCKCGIGRVY          100
SCl1           MNTKFTVLIFLGVI-VVSYGWITEKKIQKVLDEKLPNGFIKGAAKAVVHKLAKSEYGCMMDISWNKDCQRHCQSTEQKDGICHGMKCKCGKPRSY          60
```

(c) Kappa-KTx

```
La-kappaKTx1   MKPSTSAYALLLVLTFGIIT--SGVFAVPMDEENTFEVEKRGNSCMEVCLQHEGNVAECEKACNKG          100
La-kappaKTx2   MKPSTSAFILLLVLTFGIIT--SGVSAIPMDEENTFEEQKRDSACVEVCLHHEGNVAECEEACKKS          78
HeTx203        MKTSGTVYVFLLLLAFGIFTDISSACSEQMDDEDSYEVEKRGNACIEVCLQHTGNPAECDKACDK-          59

La-kappaKTx3   MKLLPLL-VILIICALMANEAFCDQGARER-SENLED-TRDLVQKPCRIVCSENMR--KCIRRCTL--G--R    100
La-kappaKTx5   MKLLPLLLVILIVCALLPNEAFCDQSAVER-SESLEEVSREIVKRSCKRVCSGTRRTKKCMQKCKSQPG--R    58
HelaTX1        MKLLPLLFVILIVCAILPDEASCDQSELERKEENFKDESREIVKRSCKKECSGSRRTKKCMQKCNREHGHGR    47

La-kappaKTx4   MKPSSFAIALILVLFLGFTNAVSGEYAESI---SGDRME---RAERAGCRIRCLQFTDDFEKCRKLCG--      100
PcavC10        MK-SSLVVASLLILLFVSINDYTSVYAQSTHEESGSKMERRARIKRAGCVIKCLQFTPELEKCRKLCGLK      50
```

(d) Delta-KTx

```
La-deltaKTx1   MASQFLLFCIVLIAVNPLVYSKG-GCRLPPETGLCYAYFERYYYDPSSNNCKMFVYGGCGGNSNNFVSKKACLARCAN---     100
BmKTT-2        MMNVITVVGIILSVVCTISDAEGVDCTLPSDTGRCKAYFIRYFYNQKAGECQKFVYGGCEGNSNNFLTKSDCCKQCSPGKC     42
```

Figure 3. Multiple sequence alignment of KTx peptides identified in this study and comparison with their similar peptides. Mature regions that were previously reported are underlined. Cys residues in the reported and putative mature regions are shaded in yellow. %Id represents the percentage of sequence identity.

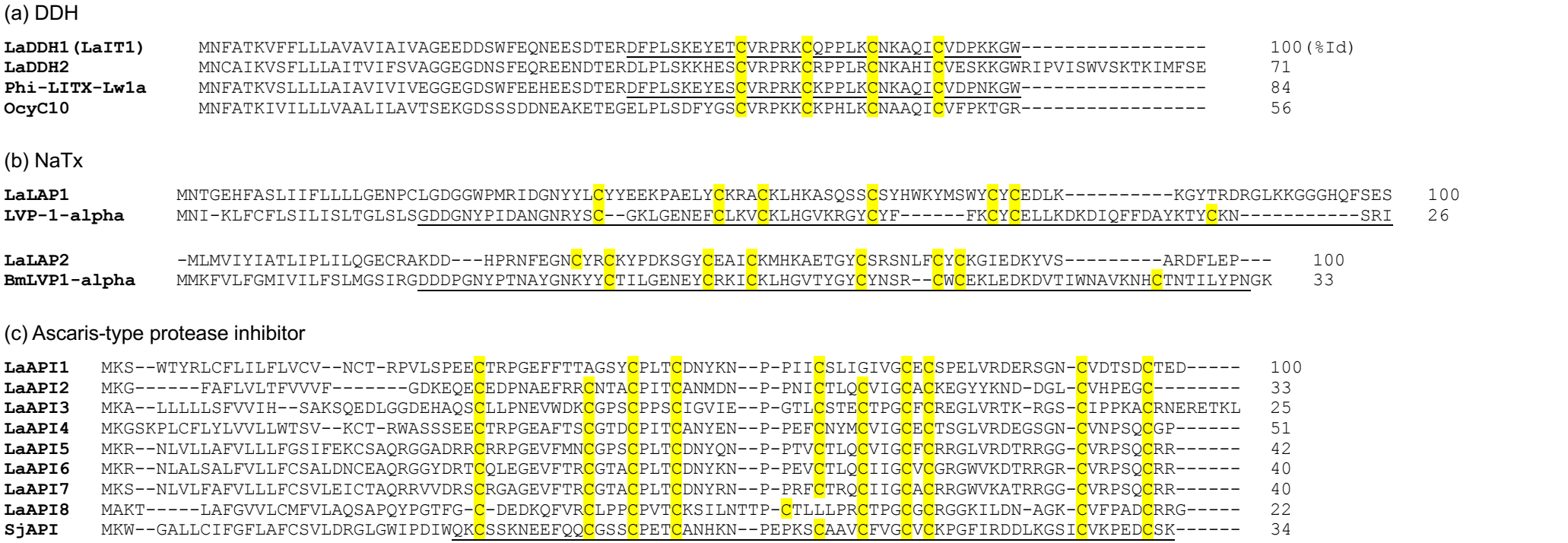

Figure 4. Multiple sequence alignment of DDH (a), NaTx (b), and Ascaris-type protease inhibitor peptides (c) identified in this study and comparison with their similar peptides. Mature regions that were previously reported are underlined. Cys residues in the reported and putative mature regions are shaded in yellow. %Id represents the percentage of sequence identity.

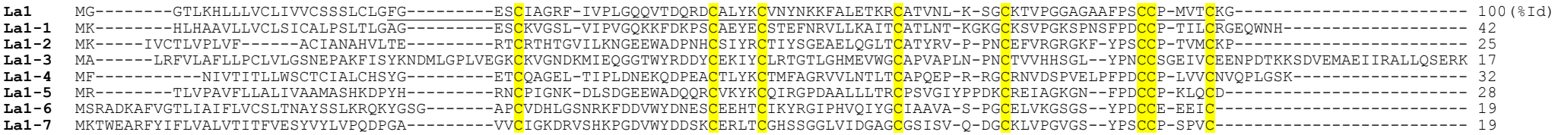

Figure 5. Multiple sequence alignment of La1-like peptides identified in this study. Mature regions that were previously reported are underlined. Cys residues in the reported and putative mature regions are shaded in yellow. %Id represents the percentage of sequence identity.

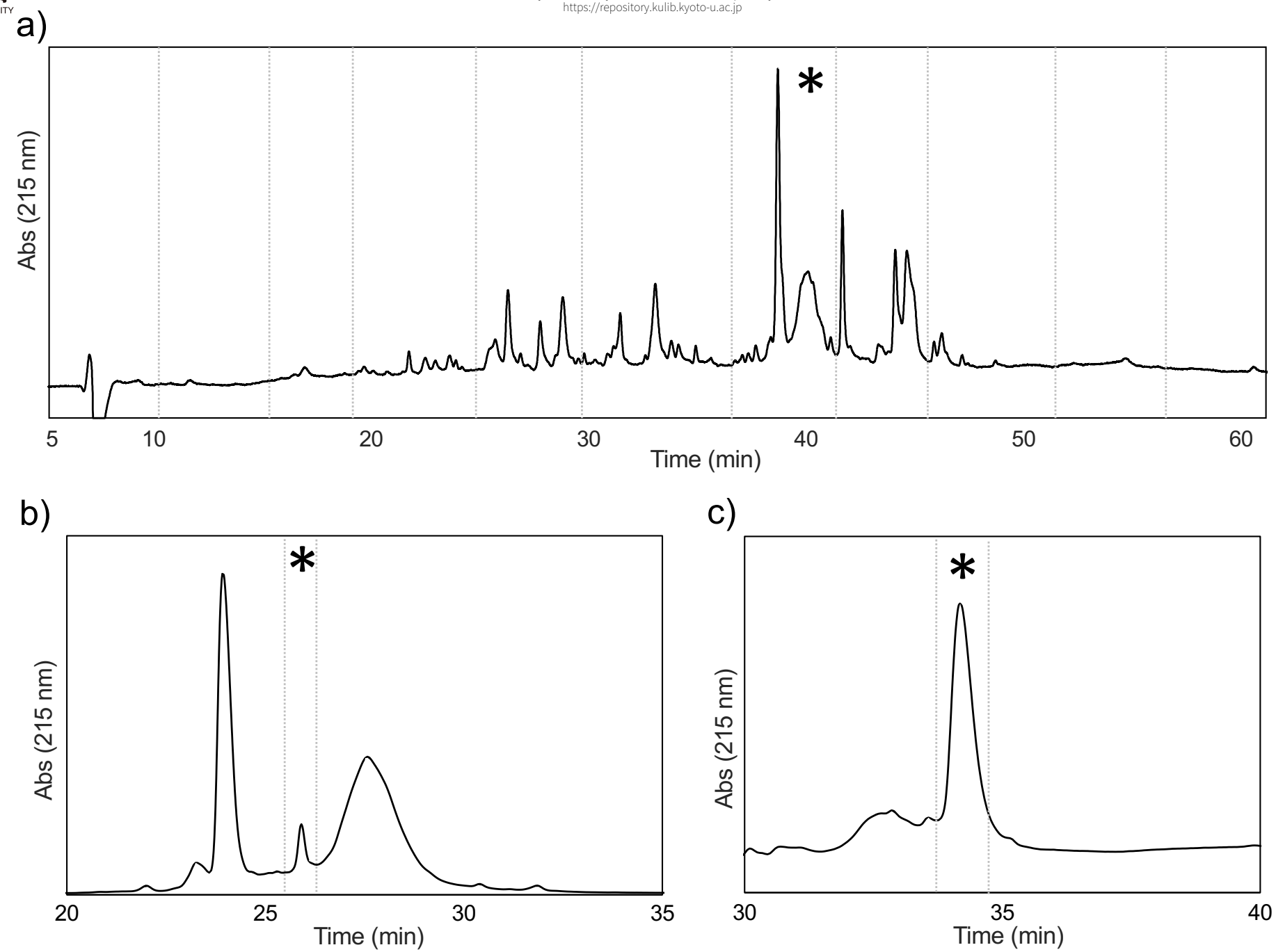

Figure 6. HPLC purification of an anti-HCV component from the venom. (a) First separation using a C4 column. (b) Second purification using a C18 column. (c) Final purification using a C18 column. Asterisks indicate fractions showing anti-HCV activity.

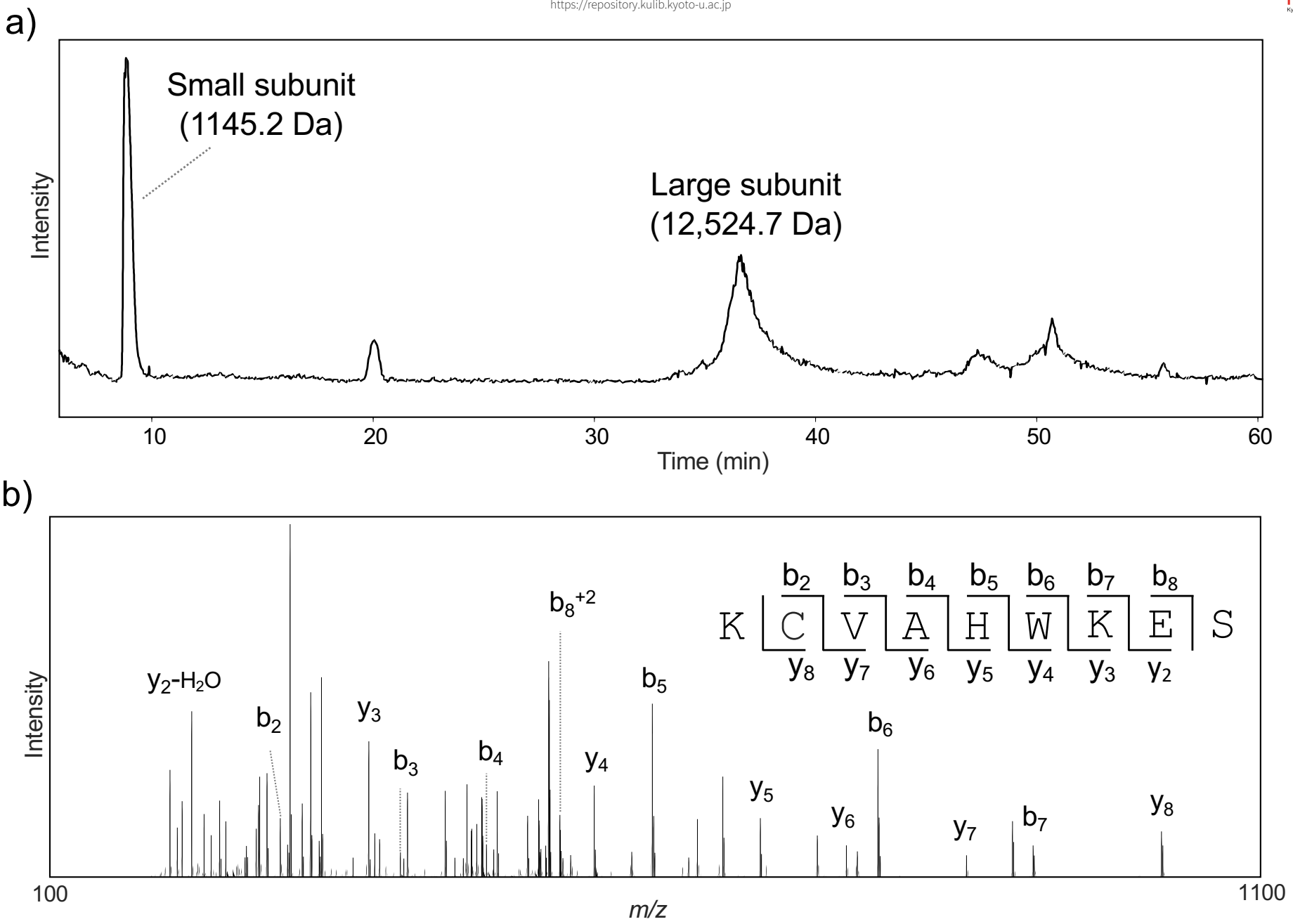

Figure 7. LC/MS analysis of the anti-HCV component after reduction/alkylation reactions. (a) TIC chromatogram of the reaction mixture. (b) Product ion spectrum of the small subunit.

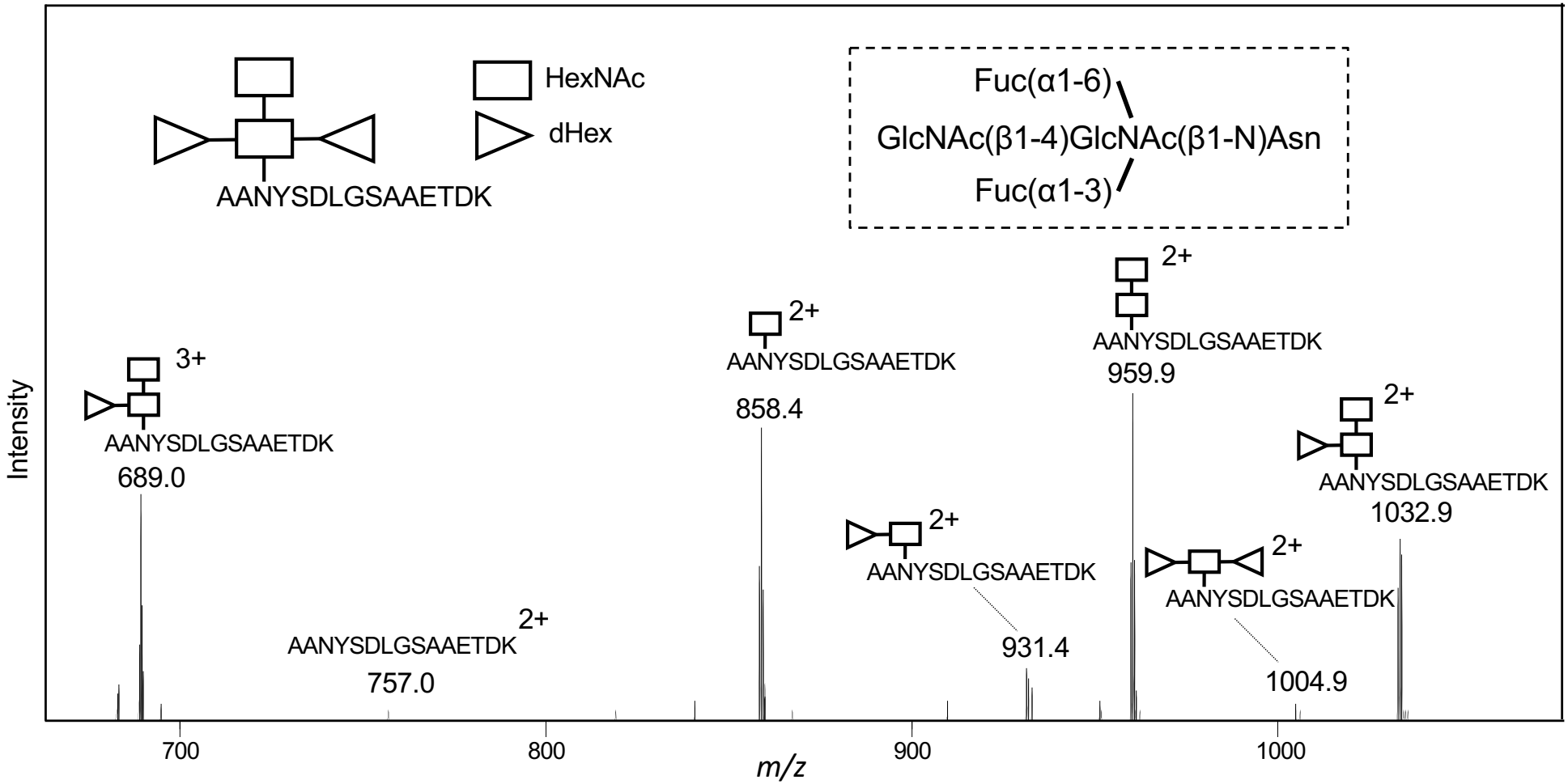

Figure 8. Product ion spectrum of the Lys-C digest containing N-glycosylation and its estimated glycan structure (dashed square).

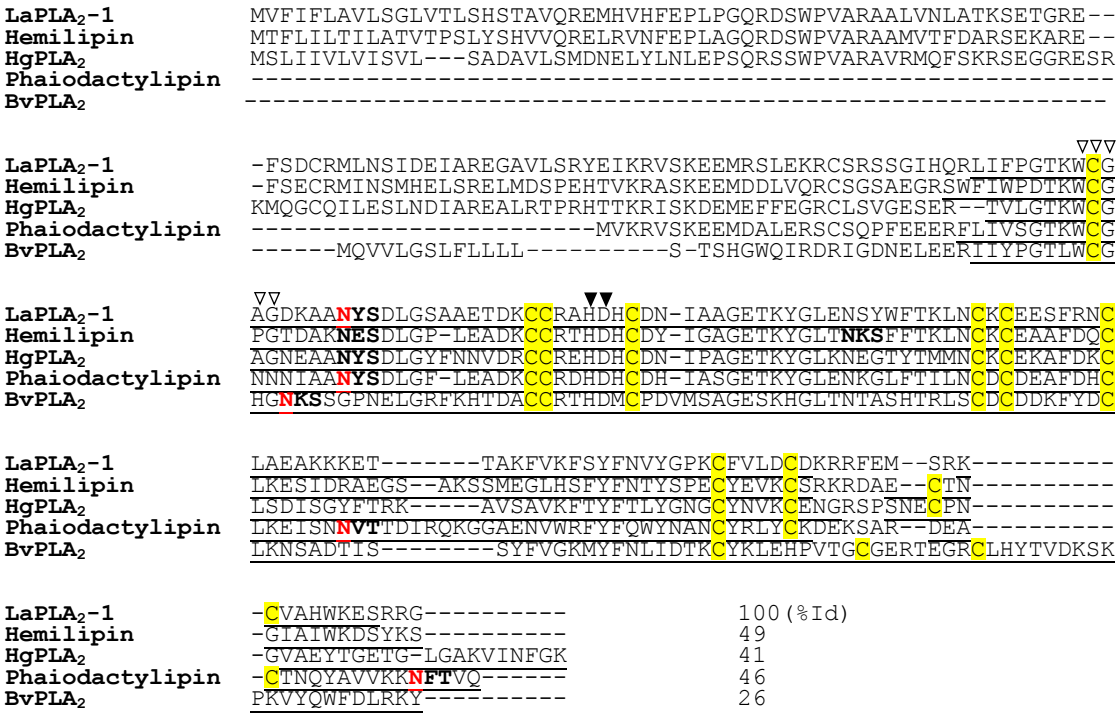

Figure 9. Multiple sequence alignment of LaPLA$_2$-1 and its related PLA$_2$ molecules. Mature regions that were experimentally determined are underlined. Cys residues in the mature regions are shaded in yellow. Open triangles indicate the Ca^{2+}-binding motif. Closed triangles indicate the catalytic dyad. N-glycosylation motif is shown in bold. The N-glycosylation sites that were experimentally confirmed are shown in red. %Id represents the percentage of sequence identity.

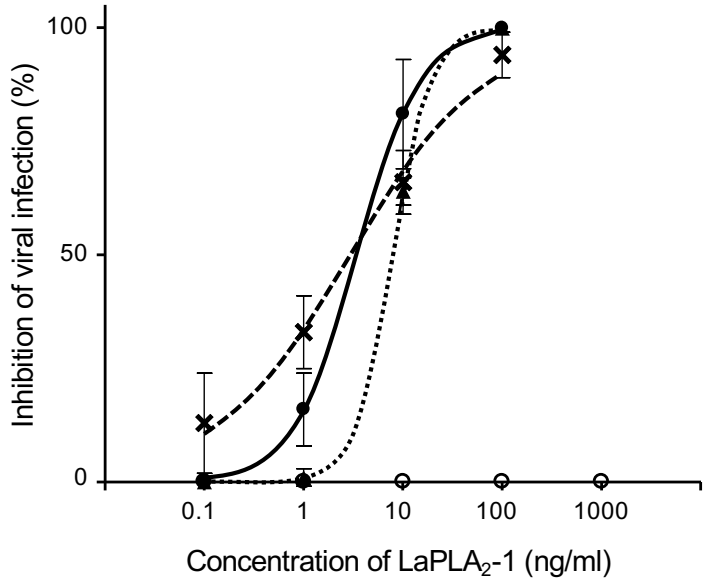

Figure 10 Representative concentration-inhibition curves for $LaPLA_2$-1 on HCV (●), DENV (×), JEV (▲), and HSV-1 (○) infection. Each point represents means ± SEM.

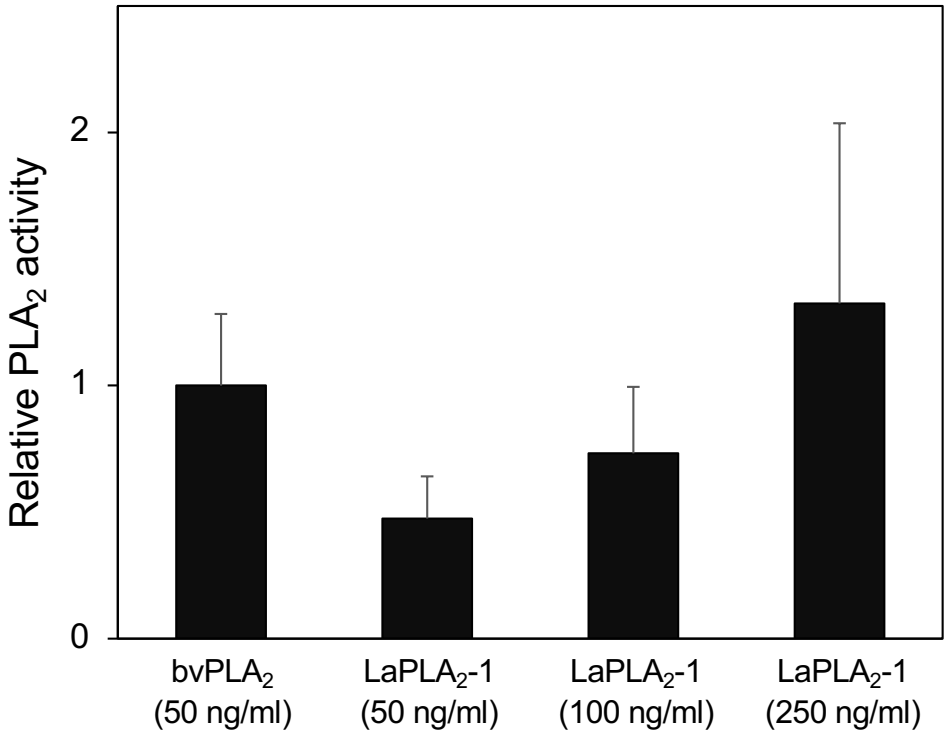

Figure 11. Comparison of PLA_2 activity between $bvPLA_2$ and $LaPLA_2$-1

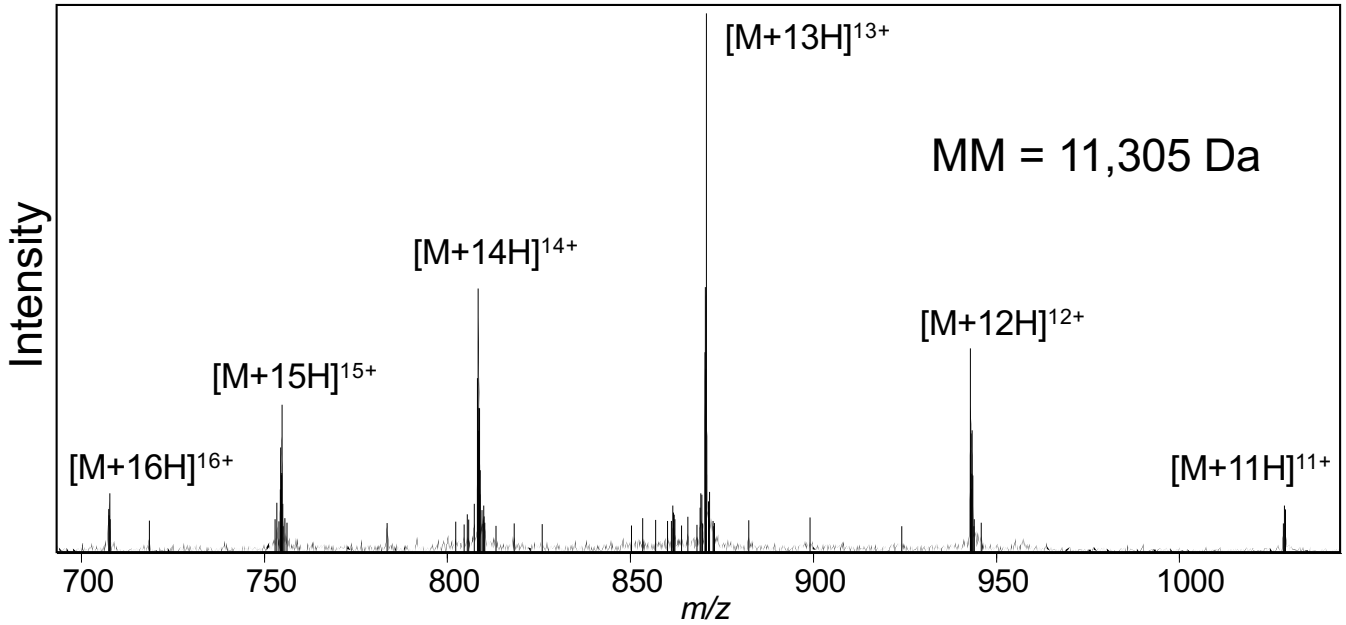

Figure S1 Mass spectrum of $LaPLA_2$-1 after deglycosylation