

Venomic and Transcriptomic Analysis of Centipede *Scolopendra subspinipes dehaani*

Zi-Chao Liu,^{†,§,#} Rong Zhang,^{‡,#} Feng Zhao,^{†,§,#} Zhong-Ming Chen,[†] Hao-Wen Liu,[‡] Yan-Jie Wang,[†] Ping Jiang,[†] Yong Zhang,[†] Ying Wu,[‡] Jiu-Ping Ding,^{*,‡} Wen-Hui Lee,^{*,†} and Yun Zhang^{*,†}

[†]Key Laboratory of Animal Models and Human Disease Mechanisms of Chinese Academy of Sciences & Yunnan Province, Kunming Institute of Zoology, the Chinese Academy of Sciences, Kunming, Yunnan 650223, China

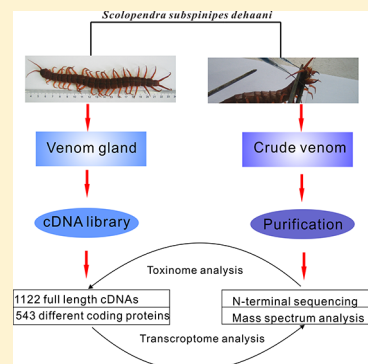
[‡]Key Laboratory of Molecular Biophysics, Huazhong University of Science and Technology, the Ministry of Education, Wuhan, Hubei 430074, China

[§]Graduate School of the Chinese Academy of Sciences, Beijing 100049, China

S Supporting Information

ABSTRACT: Centipedes have venom glands in their first pair of limbs, and their venoms contain a large number of components with different biochemical and pharmacological properties. However, information about the compositions and functions of their venoms is largely unknown. In this study, *Scolopendra subspinipes dehaani* venoms were systematically investigated by transcriptomic and proteomic analysis coupled with biological function assays. After random screening approximately 1500 independent clones, 1122 full length cDNA sequences, which encode 543 different proteins, were cloned from a constructed cDNA library using a pair of venom glands from a single centipede species. Neurotoxins, ion channel acting components and venom allergens were the main fractions of the crude venom as revealed by transcriptomic analysis. Meanwhile, 40 proteins/peptides were purified and characterized from crude venom of *S. subspinipes dehaani*. The N-terminal amino acid sequencing and mass spectrum results of 29 out of these 40 proteins or peptides matched well with their corresponding cDNAs. The purified proteins/peptides showed different pharmacological properties, including the following: (1) platelet aggregating activity; (2) anticoagulant activity; (3) phospholipase A₂ activity; (4) trypsin inhibiting activity; (5) voltage-gated potassium channel activities; (6) voltage-gated sodium channel activities; (7) voltage-gated calcium channel activities. Most of them showed no significant similarity to other protein sequences deposited in the known public database. This work provides the largest number of protein or peptide candidates with medical-pharmaceutical significance and reveals the toxin nature of centipede *S. subspinipes dehaani* venom.

KEYWORDS: centipede, proteomic, *Scolopendra*, transcriptomic, venom



■ INTRODUCTION

Centipedes are regarded as the oldest and largest terrestrial arthropods belonging to the class Chilopoda of the subphylum Myriapoda. They comprise approximately 3300–3500 species distributed in every continent except Antarctica.¹ Because they lack the waxy cuticle of insects and easily lose water through their skin, centipedes usually inhabit leaf, bark, wood and soil in forests.² Centipedes mainly prey on insects, earthworms, snails and even small vertebrates. Their first pairs of limbs stretch forward from the body to cover the remainder of the mouth. These limbs, or maxillipeds, end in sharp claws and connect with venom glands. The venom is used not only to paralyze and kill prey but also for defense against predators. Human injuries caused by centipede stings are also frequently reported. Symptoms of the stings include intense local pain, redness, swelling, superficial necrosis, chills, fever, and weakness. Occasionally, serious secondary infections can even lead to human death.^{3–6} However, knowledge on centipede venom is limited at present, and centipedes were regarded as a neglected group of venomous animals.¹ Thus, understanding centipede

venom compositions and functions will help people to reveal their biological importance and might provide novel leading compounds of potential pharmaceutical interest.

Up to now, several papers reported the composition and pharmacological properties of centipede venoms. Rates et al. found 62 and 65 proteins by two-dimensional chromatographic analysis from *Scolopendra viridicornis nigra* and *S. angulata* venoms, respectively.⁷ The molecular weights of these proteins range from 1.3 to 22.6 kDa. N-Terminal sequencing of 24 of these proteins yielded a total of 10 protein families, out of which as many as nine families showed no significant similarities with known protein families, reflecting that the venom composition and structure of centipedes are largely unknown. Serotonin, histamine, lipids, polysaccharides and polypeptides were also found from the crude extracts of centipede venom glands.⁸ However, more in depth analyses of their components were rarely studied. Gonzalez-Morales et al.

Received: September 17, 2012

Published: November 14, 2012

found a fraction with phospholipase A₂ (PLA₂) activity by HPLC from the venom of *S. viridis* and cloned the gene encoding this enzyme.⁹ The mature *S. viridis* PLA₂ has 119 amino acid residues with MW of 13 752 Da. Phylogenetic analyses showed that it is more similar to snake phospholipases than to insect or arachnid PLA₂ sequences. Two antimicrobial peptides exhibiting broad-spectrum antimicrobial activity were characterized from *S. subspinipes mutilans*. They also showed moderate hemolytic activity against both human and rabbit red blood cells.¹⁰ A serine protease has also been reported from the venom of *S. subspinipes mutilans*.¹¹

Many toxins from arthropods such as spiders and scorpions have been proved to be powerful tools for the study of voltage-gated ion channels and have potential applications as novel pharmaceutical drugs.^{12–15} Nearly all victims of centipede stings experienced burning pain, hinting that the centipede venoms contain neurotoxins or ion channel related components. More recently, Yang et al. reported neurotoxin-like peptides belonging to 10 groups from the venoms of *S. subspinipes mutilans*, revealing the diversity of neurotoxins existing in centipede venoms.¹⁶ Voltage-gated sodium, potassium, and calcium channel activities were characterized by some of these neurotoxins and showed potential insecticidal abilities. Interestingly, although voltage-gated ion channel regulation activities of centipede neurotoxins were similar to other venom neurotoxins from poisonous animals, like snakes, spiders, scorpions, marine cone snails and sea anemones, the primary structures of most centipede-venom neurotoxins were quite different from other known venom neurotoxins. However, the paper by Yang et al. focuses their investigation on insecticidal neurotoxins by peptidomics combined with transcriptomic analysis; the venom nature of centipedes are still largely unknown at present.

The *S. subspinipes dehaani* is a big (usually 6–20 cm in length) and very venomous centipede species distributed widely in the hilly areas of southern China. In this study, we reported the general composition of venoms from *S. subspinipes dehaani* by large-scale identification and analysis of proteins/peptides using proteomic and transcriptomic strategies coupled with pharmacological investigations.

■ EXPERIMENTAL SECTION

Animals and Ethics

Adult specimens of *S. subspinipes dehaani* ($n = 400$; weight ranges 8–15 g/centipede) were captured in Guangxi province of China. Animal care and handling were conducted in accordance with the requirements of the Ethics Committee of Kunming Institute of Zoology, the Chinese Academy of Sciences and Huazhong University of Science and Technology, respectively.

Collection of Centipede Venoms

The venom was obtained by electrical stimulation as follows: A piece of glass with a thickness of 5 mm was placed between the two maxillipeds of a centipede. The venom was extracted onto the glass with 3 V electrical stimulations and collected with a micropipet.¹⁰ The collected solutions were diluted with Tris-HCl buffer (20 mM Tris-HCl, pH 7.6, containing 0.1 M NaCl) and centrifuged. The supernatants were lyophilized and stored at $-20\text{ }^{\circ}\text{C}$ until use. To obtain 300 mg of lyophilized crude venom, approximately 400 *S. subspinipes dehaani* individuals were used, and each animal was milked 2–3 times with a time interval of 20 days.

Isolation and Purification of Venom Components

Venom sample (300 mg) of *S. subspinipes dehaani* was solubilized in 3 mL of the same Tris-HCl buffer mentioned above. The solution was then loaded on a Sephadex G-50 (Superfine, GE Healthcare) gel filtration column with a flow rate of 12 mL/h. Fractions were collected every 20 min. The absorbance of each tube was monitored at 280 nm. Five peaks (named P1–5) were obtained. P1 was pooled and further purified by ÄKTA Mono Q anionic exchange column (1 mL). P2, P3 and P5 were separately purified by C₁₈ reverse-phase high performance liquid chromatography column (RP-HPLC, Hypersil BDS C₁₈, 4.0 × 250 mm, Elite, China), respectively. The absorbance of isolated fractions was monitored at 215 nm. P4 was first desalted by RP-HPLC C₄ column. Pass-through peak was collected and directly purified by RP-HPLC C₁₈ column as mentioned above. Elution peak was pooled and lyophilized and then dissolved in buffer, applied to ÄKTA Resource Q anionic exchange column, and eluted by NaCl gradient (0–0.6M). Fractions obtained from Resource Q column were further separately purified by RP-HPLC C₁₈ column.

Structural Analysis of Proteins/Peptides

Mass spectrum and N-terminal sequencing of purified proteins/peptides were carried out according to our previous reported methods.¹⁷ Briefly, The observed molecular weights and the purity of samples were determined on an Autoflex Speed TOF/TOF mass spectrometer (Bruker Daltonik GmbH) in linear mode. 2,5-Dihydroxybenzoic acid, sinapinic acid and α -cyano-4-hydroxycinnamic acid were used as the matrix for samples with MW >20 000, 5000–20 000, and 500–4000 Da, respectively. All procedures were carried out according to manufacturer's standard protocols, and the data were analyzed by the software package provided by the manufacturer.

The complete sequences of proteins/peptides were determined by Edman degradation on PPSQ-31A protein sequencer (Shimadzu, Japan) according to the standard glass fiber disk (GFD) or PVDF method instructed by manufacturer's protocol. Normally, samples originating from peak 1–3 of gel filtration step were transferred to PVDF membrane and determined by PVDF method. Liquid samples originating from peak 4 and 5 of gel filtration step were determined by GFD method.

Construction of a Venom Gland cDNA Library

S. subspinipes dehaani venom gland cDNA library was constructed mainly according to our previously reported methods.^{17,18} Briefly, the two venom glands from one centipede species were dissected and immediately homogenized in liquid nitrogen. The total RNA was extracted with TRIzol Reagent (Invitrogen, USA). Isolation of mRNA from total RNA was performed with Oligotex mRNA Mini Kit according to the standard protocol (Qiagen, USA), and 0.1 μg of mRNA was used for library construction. The synthesis of cDNA was performed using a SMART PCR cDNA synthesis kit (Clontech, Palo Alto, CA, USA). The first strand of cDNA was synthesized using the SMARTScribe MMLV Reverse Transcription system with the primer pair CDS III/3' PCR primer, 5'-ATTCTAGAGGCCGAGGCGGCCGACATG-D (T)30N-1N-3' (N = A, G, C, or T; N-1 = A, G, or C) and SMART IV Oligonucleotide, 5'-AAGCAGTGGTATCAACGCAGAGTGGCCATTACGGCCGGG-3'. The second strand was produced using Advantage polymerase and primer pair CDS III/3' PCR primer and 5' PCR primer 5'-

AAGCAGTGGTATCAACGCAGAGT-3. The PCR was performed as follows: 2 min at 94 °C followed by 30 cycles of 30 s at 92 °C, 30 s at 55 °C, and 60 s at 72 °C, and then a final extension step of 72 °C for 10 min. The PCR products were ligated into pMD 19-T vector (TaKaRa Biotechnology (Dalian) Co., Ltd., China) and then transformed into *E. coli* DH5 α competent cells. A specific centipede venom gland cDNA library was constructed. Clones were randomly chosen to carry out DNA sequencing on an Applied Biosystems DNA sequencer (ABI 3730XL, USA).

Bioinformatic Analysis

All cDNA sequences were submitted into the GenBank database with accession number of KC144021 to KC145130 in NCBI (<http://www.ncbi.nlm.nih.gov/>). The signal peptide was predicted with the SignalP 4.0 program (<http://www.cbs.dtu.dk/service/SignalP/>). The propeptide cleavage site was confirmed by the determined N-terminal amino acid sequences of mature toxins. All cDNA sequences were aligned by the Basic Local Alignment Search Tool (cutoff E-value of 1×10^{-6}) in the priority order of Nr, Swissprot, and Nt current released database (<http://www.ncbi.nlm.nih.gov/>). All precursor sequences were aligned using the ClustalW2 (<http://www.ebi.ac.uk/Tools/clustalw2/index.html>).

Platelet Aggregation Assay

Human platelet-rich plasma (PRP) was provided by Kunming Blood Center of Yunnan province. Washed human platelet suspensions were prepared according to our previous described methods.¹⁹ Briefly, platelet pellets obtained from centrifugation of PRP were resuspended in buffer C (113 mM NaCl, 4.3 mM K₂HPO₄, 24.4 mM NaH₂SO₄, 5.5 mM glucose, pH 6.5) and then centrifuged at 500g for 10 min. The platelets were washed again with buffer C. Finally, the platelets were resuspended in buffer D (20 mM Hepes, 140 mM NaCl, 4 mM KCl, 5.5 mM glucose, pH 7.4). The platelet concentration was adjusted to 3×10^8 platelets/mL for both PRP and washed platelet suspensions. Before aggregation analysis, platelets were incubated with buffer D containing 2 mM CaCl₂ and 2 mM MgCl₂ at 37 °C for 2 min. Platelets aggregation was measured by light transmission in an aggregometer (Pricel, Beijing, China) with continuous stirring at 1100 rpm at 37 °C.

Hemolytic Assay

Hemolytic assays were investigated using human, mouse and rabbit red blood cells in liquid medium as reported.¹⁷ Briefly, serial dilutions of the samples were incubated with washed red blood cells (2%) at 37 °C for 30 min, the cells were centrifuged, and the absorbance of the supernatant was measured at 540 nm. Maximum hemolysis was determined by adding 1% Triton X-100 to the cell samples.

Anticoagulant Activity Assay

Human platelet-poor plasma (PPP) was provided by Kunming Blood Center of Yunnan province. Recalcification time was determined as described by Lee et al. with the following modification.²⁰ PPP (100 μ L) and certain amount of sample in 100 μ L 0.9% NaCl were incubated at 37 °C for 2 min. Then 100 μ L of 0.03 M CaCl₂ was mixed simultaneously, and the clotting time was recorded. Accordingly, the anticoagulant activity was expressed by elongating the plasma recalcification time. The plasma aliquots incubated only with 0.1 mL of 0.9% NaCl were served as controls.

PLA₂ Activity Assay

Quantitative PLA₂ activity of sample was determined mainly according to previously described pH-stat method with some modification.²¹ Briefly, diluted egg yolk (1:20, with saline) was used as substrate. Liberation of acid was measured at pH 8.0 at 37 °C by titration with 0.01 N NaOH in a total volume of 15 mL under a constant stream of argon. One unit of activity was defined as the amount of enzyme that liberates 1 μ mole of free fatty acid/min under the above conditions. Accordingly, specific PLA₂ activity of crude venom and purified enzyme were calculated as μ mole of free fatty acid/min/mg protein.

Trypsin Inhibiting Activity Assay

The trypsin inhibiting activity was performed mainly as described previously.²² The sample on the hydrolysis of synthetic chromogenic substrate S-2238 (H-D-Phe-Pip-Arg-pNA, Kabi Vitrum, Stockholm, Sweden) was assayed at 25 °C in 20 mM Tris-HCl, pH 7.8 buffer. The reaction was initiated by the addition of the substrate with a final concentration of 0.02 mM. The formation of *p*-nitroaniline was monitored continuously at 405 nm for 2 min. The effect of inhibitor was estimated by setting the initial velocity obtained in the presence of enzyme alone (without inhibitor) as 100%. The slopes (K_m/V_{max}) of lines obtained from the Lineweaver–Burk representation ($1/V$ vs $1/[S]$) of saturation curves at different inhibitor concentrations were plotted against the concentrations of inhibitor. The inhibitory constant (K_i) of the inhibitor was determined from the intercept point of the X-axis.²³

Cytotoxic Activity Assay

HeLa and HCT-116 cell lines were donated by the Chinese Type Culture Collection (Kunming Institute of Zoology, the Chinese Academy of Sciences, China). Cytotoxicity was measured by the MTT method.²⁴ Briefly, 5000 cells were plated in 96-well plates and incubated for 24 h to allow the cells to attach. Then the cells were treated with 10, 20, 40, and 80 μ g/mL final concentration of peptides. The same volume of sterile-deionized water served as a negative control. After incubation at 37 °C for 24 h, MTT (5 mg/mL in PBS) was added to each well and incubated for 4 h. Then, the MTT solution was removed and 100 μ L of DMSO was added to each well to dissolve the purple formazan crystals. The absorbance at 570 nm was then measured using an ELISA assay reader. The percentage of cell viability was calculated as viability (%) = absorbance of sample/absorbance of control $\times$ 100.

DRG Neuron Isolation

Male Wistar rats of 3–5 weeks were sacrificed by decapitation. The lumbar segments of the vertebrate column were dissected, and the lumbar L4, L5, and L6 DRG, together with the dorsal, ventral roots, and attached spinal nerves were taken out from the outside of the spinal column. These 6 DRGs were transferred into iced Dulbecco's modified Eagle's medium (DMEM (Gibco, Grand Island, NY, USA) 13.5 g/L, NaCl 2.15 g/L, HEPES 2.0 g/L, pH 7.4, 320 mOsm) immediately. After the removal of attached nerves and surrounding connective tissues, DRGs were minced with iridectomy scissors and incubated with enzymes, including 1 mL of collagenase (2 mg/mL, type I (Sigma-Aldrich, St. Louis, MO, USA), 1 mL of trypsin (0.5 mg/mL, Type IX, Sigma) and 50 μ L of DNase (4 mg/mL) in calcium-free buffer with 4 mg/mL bovine serum albumin (BSA) in a 37 °C shaking bath (170 r/min) for 35–40 min with gently mechanical trituration every 10 min. The addition of 8 mL of preincubated DMEM, including 20% fetal

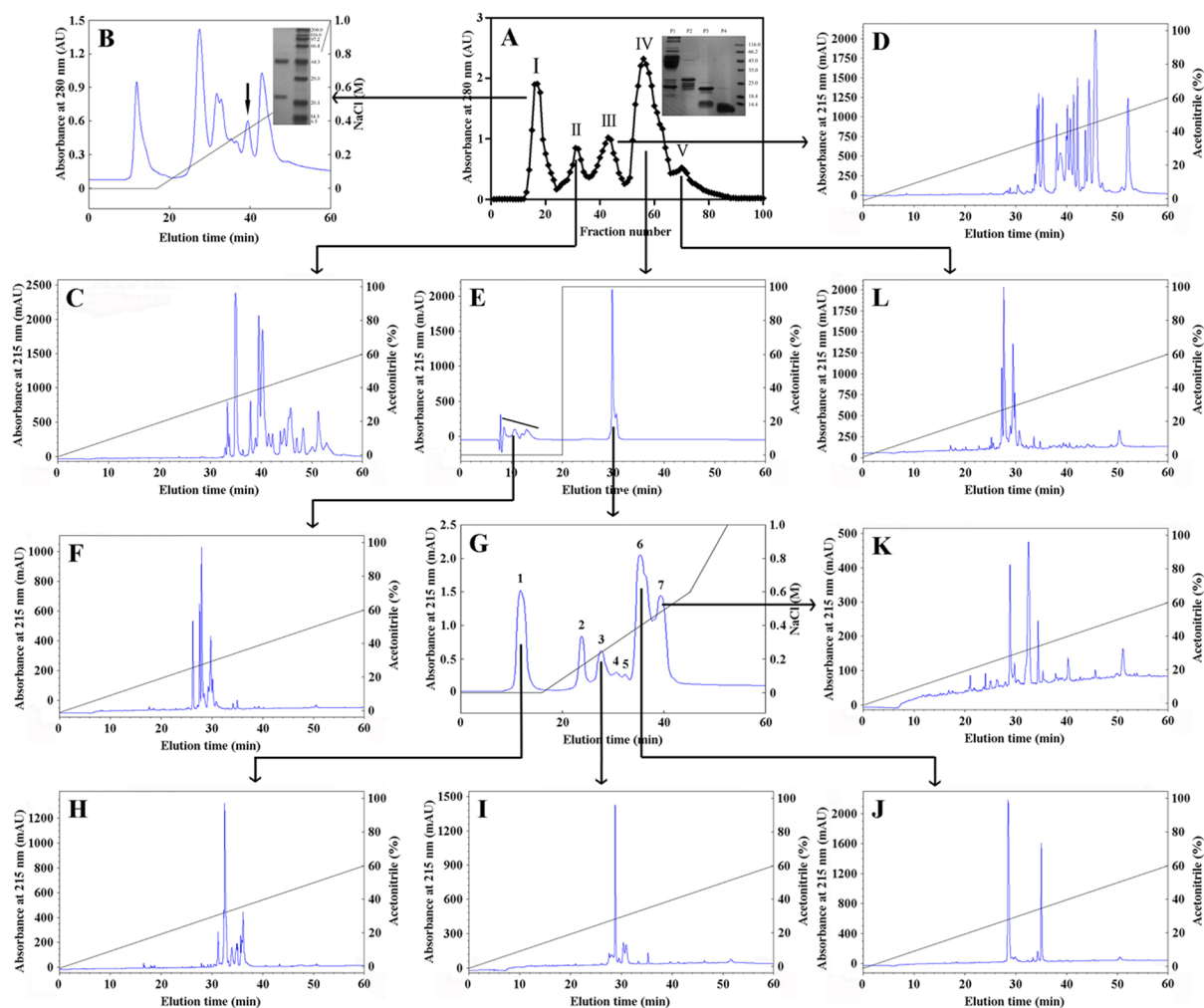


Figure 1. Purification of proteins/peptides from the crude venom of *S. subspinipes dehaani*. (A) Lyophilized venom sample of *S. subspinipes dehaani* was loaded on a Sephadex G-50 gel filtration column. Inset was SDS-PAGE profile of peak 1–4 under reducing conditions. (B) Peak 1 from (A) was applied to ÄKTA Mono Q anionic exchange column. The elution was performed at a flow rate of 1 mL/min with the indicated NaCl gradient. Peak VII with platelet aggregating activity was indicated by an arrow. Inset was SDS-PAGE profile of peak VII under reducing conditions. (C) Purification of peak II from (A) by a RP-HPLC C_{18} column. The elution was performed at a flow rate of 1 mL/min with the indicated gradients of acetonitrile. (D) Purification of peak III from (A) by a RP-HPLC C_{18} column. (E) Desalting of peak IV from (A) by a RP-HPLC C_4 column. After sample loading and baseline backs to normal, 100% acetonitrile were used immediately to elute the bind peptides. (F) Purification of collected pass-through of (E) by a RP-HPLC C_{18} column. (G) Purification of eluted peptides of (E) by ÄKTA Resource Q anionic exchange column. The elution was performed at a flow rate of 1 mL/min with the indicated NaCl gradient. (H–K) Purification profiles of peaks I, III, VI and VII from (G) by a RP-HPLC C_{18} column, respectively. (L) Purification of peak V from (A) by a RP-HPLC C_{18} column.

bovine serum, was used to stop the enzymatic digestion. The isolated neurons were plated on 0.5 mg/mL of poly-lysine coated glass coverslips and maintained in a 37 °C humidified incubator with 5% CO₂ for at least 2 h before use. The medium neurons with a diameter of 25–35 μ m were used in the experiments.

Electrophysiology

All currents were recorded in whole-cell patch configuration at room temperature (22–25 °C). Pipette resistance was typically 2–4 M Ω . For K⁺ and Na⁺ current recordings, extracellular solution contained the following: 140 mM NaCl, 3 mM KCl, 1 mM CaCl₂, 1 mM MgCl₂, 10 mM HEPES adjusted to pH 7.3 with NaOH. Intracellular solution contained the following: 140 mM KCl, 10 mM NaCl, 1 mM MgCl₂, 10 mM HEPES, 5 mM EGTA, adjusted to pH 7.3 with KOH. For Ca²⁺ current recording, extracellular solution: 135 mM NMDG, 20 mM BaCl₂, 2 mM MgCl₂, 10 mM HEPES; the pipet solution: 150

mM CsCl, 1 mM MgCl₂, 1 mM EGTA and 10 mM HEPES. The toxin solution was made by adding SSD609 or SSD559 into the extracellular solution. All chemicals were obtained from Sigma-Aldrich. The experiments were performed using a PC2C patch clamp amplifier with its software (InBio). The currents were typically digitized at 100 kHz. Macroscopic records were filtered at 5 kHz. Patch clamp recording data were analyzed with Clampfit (Axon Instruments, Foster city, CA) and Sigmaplot (SPSS Science, Chicago, IL).

Statistical Analysis

All the data were analyzed with student's *t* test for variance. Experimental values are expressed as means $\pm$ SEM. The level of statistical significance was set at the level of *P* < 0.05.

Table 1. Purification and Characterization of Proteins/Peptides from the Venom of *S. subspinipes dehaani*

position	full or N-terminal sequence	clone ID	GenBank accession No.	theoretical mass	MS Figure	determined mass	function
Mono Q profile of gel filtration peak 4 (Figure 1B)							
α -subunit	YVRVKEKQS...	SSD14	KC144034				Platelet aggregation and Hemolytic
β -subunit	TAHISVVS...						
HPLC profile of gel filtration peak (Figure 1C)							
33.462	SIGFELVTPLPGKPDECPVPV...	SSD377	KC144377	9321.21	S1	18644.6	
35.044	EGSDSKAIPTCREASFCALQKNPIDSNDL...	SSD16	KC144036	22557.12	S2	22544.4	
37.985	EMDLKDIIIDTCHQNAACAHIQM...	SSD282	KC144287	22410.39	S3	22407.8	K+ channel inhibitor
39.623	EMDLEDIIIDTCHQKAVCAHIQ...	SSD92	KC144104	22465.54	S4	22462.1	K ⁺ channel inhibitor
40.357	ELTTEDELATCNSDSICGYL...	SSD431	KC144430	22215.73	S5	22213.7	Anticoagulation
44.626	CDMKVRGLDANMKKMILDL...	SSD346	KC144347	20749.50	S6	20722	Ca ²⁺ channel inhibitor
47.071	TAHISVVSRAQDAVVVTTTINYW...	SSD20	KC144040	20694.54	S7	20667.7	K ⁺ channel inhibitor
48.345	CDMKVRGLDANMKKMILDLHNKKRQIVANG...	SSD552	KC144549	20558.3	S8	20485.6	
51.34	SGKDDLAGVCKAGKGVLVK...				S9	21622.5	
HPLC profile of gel filtration peak (Figure 1D)							
30.425	NTGEAETTFAPVNNKFTRED...	SSD996	KC144986	6117.16	S10	6117.8	
33.756	IKCIKCGESGLFGTEDCVT...	SSD202	KC144209	8556.42	S11	8557	
34.874	NNLPDASDIAYEEITTQAGT...				S12	12565	K ⁺ channel inhibitor
35.314	EGSDSKAIPTCREASFCALQKNP...	SSD16	KC144036	22557.12	S13	22559.3	
38.022	LKVEVFRQTVADNVLVGSYCAISN...				S14	9443.9	
38.801	LKIEDLPEPESYKKAQLAVKDANGDKRAEGL...	SSD449	KC144448	8726.10	S15	8726.6	Ca ²⁺ channel inhibitor
39.954	LKIEDLPEPESYKKAQLAVKDANGDKRA...	SSD856	KC144849	8668.06	S16	8668.2	K ⁺ channel inhibitor
40.133	SLANFLSMTLTAARKRPKEYDGYGNYCG...				S17	13081.2	
40.691	LKVEDLPEPESYKRAQLAVKDA...	SSD1020	KC145009	8286.77	S18	8286.6	
41.546	SLANFLSMTLTAARKSPQEYDGY...	SSD387	KC144387	13426.81	S19	13367.3	
42.176	SLANFLSMTLTAARKSPQEYDGYGNYCG...	SSD387	KC144387	13426.81	S20	13440.8	Phospholipase A ₂
43.689	SLANFFFMTLTAARKSPQEYDGYGNYCG...	SSD1043	KC145030	13520.93	S21	13511.8	
44.441	CDMKVRGLDANMKKMILDLHNKKRQIVANG...	SSD558	KC144555	20704.36	S22	20700.6	
45.725	CDMKVRGLDANMKKMILDLHNKKRQIV...	SSD976	KC144967	20646.33	S23	20643.3	Ca ²⁺ channel inhibitor
52.197	WGCQMSERGLDKMKMKNKILKHHNELR...	SSD43	KC144061	21405.07	S24	21412.7	Trypsin inhibitor
HPLC profile of gel filtration peak (Figure 1F)							
27.956	ADNKCDSRTRRIYQRCN...				S25	5522.4	K ⁺ channel inhibitor
HPLC profile of gel filtration peak (Figure 1H)							
29.749	LEDNLVLPCPGFSCPKGYVCDRASQKCRQGT				S26	3469.7	
35.652	TDDKPIGKCGEKQSK...	SSD1084	KC145071	5730.58	S27	5711.6	
36.169	NLIYECKWADSIRLKDKNPTHEFCKK...	SSD559	KC144556	8513.60	S28	8556.2	K ⁺ channel inhibitor
HPLC profile of gel filtration peak (Figure 1I)							
28.725	ADDKCEDSLRREIACTKCRDRVRTDD...	SSD609	KC144606	5624.30	S29	5624.5	K ⁺ channel inhibitor
30.966	EVIKRDIPYKKRKFYPYKSECLK...	SSD1052	KC145039	6027.98	S30	6028.1	Ca ²⁺ channel inhibitor
HPLC profile of gel filtration peak (Figure 1J)							
28.544	EDNLVLSCAEFPCEGYICDTASQKCRPGTD	SSD410	KC144409	3358.72	S31	3358.6	
34.392	EVIRDSVIHDEEKFAQRS...				S32	7500.3	
35.094	EVIRDSVIHDEEKFANRSY...	SSD1014	KC145003	6317.06	S33	6317.2	
HPLC profile of gel filtration peak (Figure 1K)							
29.208	EVIRKEIPYKKRKFYPYKSE...	SSD893	KC144884	6101.08	S34	6102	
32.976	DDNLVLSCAEFPCEEYIC...				S35	3415.8	
HPLC profile of gel filtration peak (Figure 1L)							
27.672	EVTVEPLRHSNKNPTESECKKACA...	SSD800	KC144793	6034.71	S36	6035.7	Na ⁺ channel inhibitor
28.944	EQIFPCPGFPCPKGYFCDKGSQKCR...	SSD219	KC144226	3234.67	S37	3233.7	K ⁺ channel inhibitor
29.458	NLIYECRWADSIRLKDKNPT...				S38	8569.2	
30.681	ADNKCDSRTRRIYQQCNCN...				S39	5499	

RESULTS AND DISCUSSION

Venomic Profiling of *S. subspinipes dehaani*

The pooled venom of *S. subspinipes dehaani* was separated into 5 peaks (named P1 through P5) by gel filtration on a Sephadex G-50 column (Figure 1A). SDS-PAGE analysis indicated that the main MW ranges of pooled P1–P4 were >35, 20–30, 15–20, and <14 kDa, respectively (Figure 1A inset). P1 was further purified by Mono Q anionic exchange column, and 8 fractions

presented in the elution peaks (Figure 1B). P2, P3 and P5 from gel filtration step were further separately isolated by C₁₈ RP-HPLC column, and chromatographs of HPLC isolation were indicated in Figure 1C,D,L, respectively. As the main components of crude venom, P4 was initially loaded onto a RP-HPLC C₁₈ column. However, the results indicated that the isolation efficacy was not good enough, as many peaks were not well isolated because of the richness of the peptides contained in the pooled samples (figure not shown). Thus, to achieve

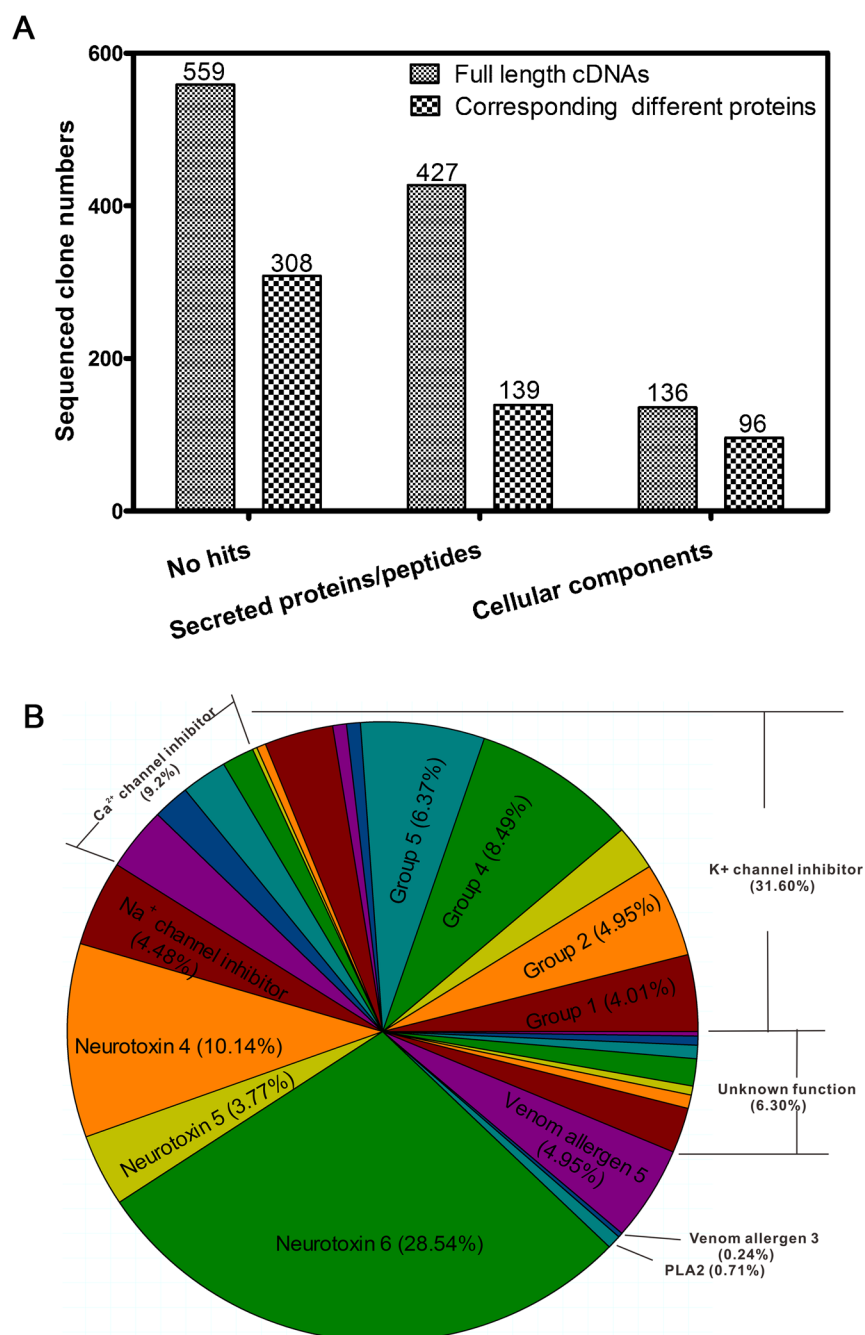


Figure 2. The venom gland full length cDNA transcriptome profile of *S. subspinipes dehaani*. (A) Overview of venom gland cDNAs and their coding molecules. (B) Relative abundance of different components in the Secreted proteins/peptides category of (A). No Hits, cDNA sequences could not be confidently identified in GenBank (Score < 10^{-6}). Secreted proteins/peptides, cDNA sequences or their corresponding mature proteins/peptides could be extracted from significant BLAST matches or purified from crude venom.

better isolation, P4 from gel filtration step was first desalted by a RP-HPLC C₄ column (Figure 1E). The pass-through of RP-HPLC C₄ column was collected and loaded onto a RP-HPLC C₁₈ column (Figure 1F). The elution peak of RP-HPLC C₄ column was first purified by a Resource Q column (Figure 1G) followed by RP-HPLC C₁₈ fractionation for the isolated peaks (Figure 1H–K), respectively.

Totally, 55 proteins/peptides were purified by the above-mentioned isolation procedures, and a few molecules appeared more than once in our isolation steps; e.g., 27.565 min of Figure 1F, 31.167 min of Figure 1H, 28.725 min of Figure 1I, and 27.364 min of Figure 1L were identified to be the same

molecule. Finally, we characterized 40 well isolated proteins/peptides, their purification procedures, N-terminal amino acid sequence determination, theoretical and observed molecular mass, together with biological function assay results are summarized in Table 1. Mass spectra are provided in the Supporting Information, Figures S1–S39. Surprisingly, ion channel activities were widely distributed in purified fractions. Nine different fractions were characterized as voltage-gated potassium channel inhibitors, in which 6 molecules or their isoforms could be confirmed by proteomic and transcriptomic analysis. The remaining 2 voltage-gated potassium channel inhibitors could not find their corresponding cDNAs. There

Table 2. Best Hit of the Putative Mature Proteins/Peptides Identified in the Venom Gland Transcriptome of *S. subspinipes dehaani*^a

transcripts	GenBank accession No.	length	no. of transcripts	gene description	hit accession	E-value	theoretical pI	MW
K ⁺ channel inhibitor activity								
Group 1								
SSD282	KC144287	202	7	K ⁺ channel inhibitor by this work			5.99	23495.64
SSD115	KC144124	193	2	Putative K ⁺ channel inhibitor			5.23	22611.80
SSD140	KC144149	193	1	Putative K ⁺ channel inhibitor			5.23	22551.70
SSD222	KC144229	202	1	Putative K ⁺ channel inhibitor			6.05	23608.79
SSD367	KC144367	202	1	Putative K ⁺ channel inhibitor			5.72	23496.58
SSD443	KC144442	156	1	Putative K ⁺ channel inhibitor			5.25	18249.67
SSD456	KC144455	166	1	Putative K ⁺ channel inhibitor			5.81	19499.17
SSD835	KC144828	156	2	Putative K ⁺ channel inhibitor			5.25	18221.61
SSD957	KC144948	202	1	Putative K ⁺ channel inhibitor			5.82	23946.30
Group 2								
SSD92	KC144104	203	5	K ⁺ channel inhibitor by this work			5.59	23552.79
SSD3	KC144023	203	1	Putative K ⁺ channel inhibitor			5.62	23703.07
SSD60	KC144075	203	3	Putative K ⁺ channel inhibitor			5.62	23737.09
SSD163	KC144172	203	2	Putative K ⁺ channel inhibitor			5.62	23737.09
SSD185	KC144193	203	1	Putative K ⁺ channel inhibitor			5.47	23481.67
SSD332	KC144335	203	1	Putative K ⁺ channel inhibitor			5.47	23555.81
SSD417	KC144416	200	1	Putative K ⁺ channel inhibitor			5.28	23059.23
SSD507	KC144505	203	1	Putative K ⁺ channel inhibitor			5.76	23707.06
SSD521	KC144519	202	2	Putative K ⁺ channel inhibitor			5.99	23555.68
SSD573	KC144570	193	1	Putative K ⁺ channel inhibitor			5.23	22523.64
SSD599	KC144596	193	1	Putative K ⁺ channel inhibitor			5.33	22294.39
SSD809	KC144802	203	1	Putative K ⁺ channel inhibitor			5.67	23561.80
SSD1115	KC145102	203	1	Putative K ⁺ channel inhibitor			5.59	23535.76
Group 3								
SSD20	KC144040	195	1	K ⁺ channel inhibitor by this work	gil270012458	2.20×10^{-057}	6.97	21157.06
SSD5	KC144025	151	1	Putative K ⁺ channel inhibitor	gil346471281	5.40×10^{-045}	9.04	16302.63
SSD824	KC144817	151	1	Putative K ⁺ channel inhibitor	gil346471281	1.70×10^{-045}	9.30	16347.78
SSD973	KC144964	86	1	Putative K ⁺ channel inhibitor	gil346471281	4.80×10^{-018}	9.30	9578.04
SSD248	KC144254	105	4	Putative K ⁺ channel inhibitor	gil346471281	9.30×10^{-026}	9.69	11425.15
SSD616	KC144613	105	1	Putative K ⁺ channel inhibitor	gil346471281	5.50×10^{-026}	9.69	11455.18
SSD748	KC144742	105	1	Putative K ⁺ channel inhibitor	gil346471281	1.20×10^{-025}	9.69	11515.28
Group 4								
SSD59	KC144074	51	4	Putative K ⁺ channel inhibitor 4	gil392295737	2.00×10^{-037}	7.81	5829.76
SSD154	KC144163	51	3	Putative K ⁺ channel inhibitor 4	gil392295737	5.90×10^{-036}	8.32	5851.81
SSD161	KC144170	51	1	Putative K ⁺ channel inhibitor 4	gil392295737	4.40×10^{-033}	8.60	5798.83
SSD190	KC144198	51	10	Putative K ⁺ channel inhibitor 4	gil392295737	1.40×10^{-033}	8.60	5782.79
SSD296	KC144300	51	8	Putative K ⁺ channel inhibitor 4	gil392295737	3.40×10^{-034}	8.62	5752.76
SSD300	KC144304	50	1	Putative K ⁺ channel inhibitor 4	gil392295737	1.60×10^{-017}	9.97	5809.98
SSD318	KC144321	51	3	Putative K ⁺ channel inhibitor 4	gil392295737	2.30×10^{-035}	8.62	5752.76
SSD325	KC144328	50	1	Putative K ⁺ channel inhibitor 4	gil392295737	9.80×10^{-019}	9.97	5809.98
SSD379	KC144379	51	1	Putative K ⁺ channel inhibitor 4	gil392295737	4.00×10^{-033}	8.31	5810.80
SSD390	KC144390	51	1	Putative K ⁺ channel inhibitor 4	gil392295737	6.30×10^{-033}	8.59	5831.86
SSD797	KC144790	51	1	Putative K ⁺ channel inhibitor 4	gil392295737	3.70×10^{-028}	8.33	5685.67
SSD841	KC144834	51	1	Putative K ⁺ channel inhibitor 4	gil392295737	4.70×10^{-036}	8.31	5753.75
SSD971	KC144962	51	1	Putative K ⁺ channel inhibitor 4	gil392295737	1.10×10^{-033}	8.60	5798.83
Group 5								
SSD609	KC144606	47	1	K ⁺ channel inhibitor by this work			5.69	5690.40
SSD44	KC144062	47		kappa-slptx-ssm1a neurotoxin	gil392295717	1.90×10^{-012}	5.69	5630.30
SSD101	KC144112	47	16	kappa-slptx-ssm1a neurotoxin	gil392295717	4.40×10^{-014}	5.69	5750.50
SSD188	KC144196	47	1	kappa-slptx-ssm1a neurotoxin	gil392295717	2.90×10^{-014}	5.69	5690.40
SSD577	KC144574	46	1	kappa-slptx-ssm1a neurotoxin	gil392295717	2.80×10^{-013}	5.69	5467.13
SSD603	KC144600	47	1	kappa-slptx-ssm1a neurotoxin	gil392295717	9.60×10^{-013}	5.69	5630.30
SSD619	KC144616	47	2	kappa-slptx-ssm1a neurotoxin	gil392295717	1.20×10^{-013}	5.69	5640.34
SSD662	KC144659	47	2	kappa-slptx-ssm1a neurotoxin	gil392295717	3.50×10^{-014}	5.69	5648.34
SSD955	KC144946	47	1	kappa-slptx-ssm1a neurotoxin	gil392295717	9.60×10^{-015}	5.69	5658.36
SSD827	KC144820	19	1	kappa-SLPTX-Ssm2a neurotoxin	gil392295719	8.00×10^{-011}	4.41	2095.38
Group 6								

Table 2. continued

transcripts	GenBank accession No.	length	no. of transcripts	gene description	hit accession	E-value	theoretical pI	MW
SSD335	KC144338	50	1	Putative K ⁺ channel inhibitor 8	gil392295745	1.40×10^{-032}	7.55	5742.59
SSD1084	KC145071	50	1	Putative K ⁺ channel inhibitor 8	gil392295745	1.20×10^{-032}	7.53	5736.58
SSD1124	KC145111	50	1	Putative K ⁺ channel inhibitor 8	gil392295745	5.20×10^{-034}	7.55	5743.53
Group 7								
SSD162	KC144171	47	3	Putative K ⁺ channel inhibitor 7	gil392295743	5.00×10^{-009}	8.32	5528.38
Group 8								
<u>SSD219</u>	KC144226	41	13	K ⁺ channel inhibitor by this work			7.94	4587.28
SSD299	KC144303	41	1	Putative K ⁺ channel inhibitor			7.94	4587.28
SSD34	KC144052	41	1	Putative K ⁺ channel inhibitor			6.38	4544.25
Group 9								
<u>SSD856</u>	KC144849	78	1	K ⁺ channel inhibitor by this work			8.27	8674.06
SSD730	KC144727	70	1	Putative K ⁺ channel inhibitor			9.35	7956.40
Group 10								
<u>SSD559</u>	KC144556	74	1	K ⁺ channel inhibitor by this work			7.60	8663.73
Ca ²⁺ channel inhibitor								
Group 1								
<u>SSD976</u>	KC144967	187	3	Ca ²⁺ channel inhibitor by this work	gil321449768	3.00×10^{-035}	9.06	20652.33
SSD331	KC144334	187	3	Putative Ca ²⁺ channel inhibitor	gil321449768	4.10×10^{-035}	9.06	20652.33
SSD558	KC144555	187	1	Putative Ca ²⁺ channel inhibitor	gil321476851	3.50×10^{-034}	8.95	20710.36
Group 2								
SSD376	KC144376	54	7	omega-slptx-ssm2a neurotoxin	gil392295725	4.10×10^{-037}	5.16	6306.14
SSD1074	KC145061	54	1	omega-slptx-ssm2a neurotoxin	gil392295725	3.30×10^{-036}	5.16	6290.08
SSD1076	KC145063	54	1	omega-slptx-ssm2a neurotoxin	gil392295725	7.80×10^{-036}	5.16	6276.07
SSD1121	KC145108	54	1	omega-slptx-ssm2a neurotoxin	gil392295725	2.80×10^{-036}	5.16	6280.06
Group 3								
<u>SSD1052</u>	KC145039	53	7	Ca ²⁺ channel inhibitor by this work			9.19	6031.98
SSD386	KC144386	53	1	Scolopendra 5848.48 Da toxin	gil212288533	5.00×10^{-006}	9.19	5997.96
Group 4								
<u>SSD449</u>	KC144448	78	13	Ca ²⁺ channel inhibitor by this work			7.74	8732.10
SSD980	KC144971	78	1	Putative Ca ²⁺ channel inhibitor			7.74	8732.10
Na ⁺ channel inhibitor								
<u>SSD800</u>	KC144793	53	11	Na ⁺ channel inhibitor by this work			6.31	6038.71
SSD1123	KC145110	62	2	Putative Na ⁺ channel inhibitor			7.90	7216.19
SSD591	KC144588	53	1	Putative Na ⁺ channel inhibitor			7.06	5980.68
SSD696	KC144693	53	4	Putative Na ⁺ channel inhibitor			6.31	6038.71
SSD364	KC144364	55	1	Putative Na ⁺ channel inhibitor			6.29	6285.02
Neurotoxin								
Neurotoxin 4								
SSD267	KC144272	54	2	neurotoxin 4	gil392295757	6.00×10^{-006}	9.65	6121.16
SSD686	KC144683	48	2	neurotoxin 4	gil392295757	1.00×10^{-008}	9.70	5403.22
SSD986	KC144976	54	1	neurotoxin 4	gil392295757	7.00×10^{-012}	9.65	6175.21
SSD996	KC144986	54	38	neurotoxin 4	gil392295757	4.00×10^{-011}	9.65	6121.16
Neurotoxin 5								
SSD41	KC144059	42	9	neurotoxin 5	gil392295759	5.50×10^{-011}	7.80	4832.44
SSD119	KC144128	42	2	neurotoxin 5	gil392295759	4.90×10^{-009}	7.84	4858.46
SSD233	KC144240	42	3	neurotoxin 5	gil392295759	4.00×10^{-009}	7.80	4840.50
SSD565	KC144562	42	2	neurotoxin 5	gil392295759	1.70×10^{-010}	7.80	4784.40
Neurotoxin 6								
SSD23	KC144042	40	1	neurotoxin 6	gil392295761	1.70×10^{-019}	8.20	4304.90
SSD24	KC144043	40	22	neurotoxin 6	gil392295761	8.10×10^{-020}	8.20	4304.90
SSD29	KC144048	42	47	neurotoxin 6	gil392295761	2.30×10^{-010}	7.80	4798.42
SSD81	KC144094	40	32	neurotoxin 6	gil392295761	7.80×10^{-020}	8.20	4304.90
SSD84	KC144097	40	2	neurotoxin 6	gil392295761	1.60×10^{-019}	8.26	4387.97
SSD268	KC144273	40	2	neurotoxin 6	gil392295761	1.40×10^{-020}	8.20	4304.90
SSD279	KC144284	40	1	neurotoxin 6	gil392295761	6.80×10^{-018}	8.61	4402.05
SSD383	KC144383	40	1	neurotoxin 6	gil392295761	1.10×10^{-018}	7.43	4376.96
SSD409	KC144408	40	2	neurotoxin 6	gil392295761	1.20×10^{-019}	8.54	4232.84

Table 2. continued

transcripts	GenBank accession No.	length	no. of transcripts	gene description	hit accession	E-value	theoretical pI	MW
SSD451	KC144450	40	1	neurotoxin 6	gil392295761	1.30×10^{-018}	8.62	4311.92
SSD560	KC144557	40	1	neurotoxin 6	gil392295761	3.30×10^{-018}	8.14	4244.86
SSD673	KC144670	17	2	neurotoxin 6	gil392295761	2.80×10^{-009}	4.78	1822.01
SSD729	KC144726	40	2	neurotoxin 6	gil392295761	9.60×10^{-019}	8.26	4387.97
SSD783	KC144776	40	1	neurotoxin 6	gil392295761	5.80×10^{-021}	8.20	4330.94
SSD939	KC144930	40	1	neurotoxin 6	gil392295761	3.60×10^{-018}	8.71	4339.98
SSD1015	KC145004	42	1	neurotoxin 6	gil392295761	3.60×10^{-019}	7.80	4798.42
SSD1037	KC145025	44	1	neurotoxin 6	gil392295761	8.90×10^{-010}	9.62	4751.54
SSD418	KC144417	20	1	neurotoxin 6	gil392295761	1.00×10^{-010}	3.91	2229.43
Phospholipase A2								
<u>SSD387</u>	KC144387	119	1	this work			4.49	14505.09
<u>SSD1043</u>	KC145030	119	2	phospholipase A2 precursor	gil306755927	3.50×10^{-050}	4.63	14325.88
Serine protease								
SSD1134	KC145121	242	1	Trypsin-like protease	gblACE80257.1	1.00×10^{-046}	6.13	27199.02
SSD1135	KC145122	130	1	Trypsin-like protease	embl CAA75310.1	3.00×10^{-023}	5.01	14435.29
Venom allergen								
Venom allergen 3								
SSD552	KC144549	187	1	venom allergen 3	gil321476851	2.80×10^{-026}	8.80	20564.30
Venom allergen 5								
SSD43	KC144061	187	1	trypsin inhibitor by this work	gil321476851	2.10×10^{-034}	9.06	20626.29
SSD7	KC144027	191	3	venom allergen 5	gil321449768	5.80×10^{-036}	8.82	21411.07
<u>SSD346</u>	KC144347	197	1	venom allergen 5	gil321449768	6.80×10^{-031}	9.00	21757.71
SSD362	KC144362	187	1	venom allergen 5	gil321449768	6.50×10^{-038}	9.06	20806.58
SSD415	KC144414	187	1	venom allergen 5	gil321449768	9.90×10^{-036}	9.06	20678.32
SSD608	KC144605	191	4	venom allergen 5	gil321476851	4.40×10^{-033}	8.82	21350.97
SSD706	KC144703	147	2	venom allergen 5	gil345498011	1.80×10^{-024}	8.54	16306.29
SSD768	KC144762	187	1	venom allergen 5	gil321476851	9.70×10^{-035}	9.06	20641.30
SSD897	KC144888	187	1	venom allergen 5	gil321449768	1.60×10^{-034}	9.06	20626.25
SSD969	KC144960	191	1	venom allergen 5	gil321476851	3.50×10^{-034}	8.66	21483.13
SSD1044	KC145031	187	1	venom allergen 5	gil321449768	5.20×10^{-035}	9.06	20638.30
SSD1080	KC145067	187	1	venom allergen 5	gil321449768	7.70×10^{-038}	9.06	20822.63
SSD1082	KC145069	187	1	venom allergen 5	gil321449768	8.70×10^{-036}	9.06	20852.65
SSD1090	KC145076	187	1	venom allergen 5	gil321449768	3.40×10^{-037}	9.06	20828.63
SSD1091	KC145077	187	1	venom allergen 5	gil321449768	6.80×10^{-035}	9.12	20909.69

^aThe underline transcripts were identified by both N-terminal amino acid sequencing and MS.

were 3 voltage-gated calcium channel inhibitors with their corresponding coding cDNAs identified. Only 1 peptide and its coding cDNA that acted as voltage-gated sodium channel inhibitor was characterized from crude venom. Platelet aggregating, anticoagulant, phospholipase A₂ and trypsin inhibiting activities were exerted by purified proteins/peptides from the crude venom of *S. subspinipes dehaani*.

Overview of *S. subspinipes dehaani* Venom Gland Transcriptome

The titer of the constructed *S. subspinipes dehaani* venom gland cDNA library was about 1.2×10^6 . 1122 full length cDNA sequences were obtained after random screening of 1500 independent clones and removal of nonvenom gland sequences (vector, adaptor) and short (<100 bp) or partial sequences. Finally, on the basis of the identity of deduced protein/peptide precursors, 543 full length cDNAs encoding different proteins/peptides were obtained after discarding repeated sequences (Figure 2A). On the basis of the protein blast and functional assay results, these proteins/peptides were classified into no hits (308, 56.72%), secreted proteins/peptides (139, 25.60%) and cellular components (96, 17.68%), respectively (Figure 2A). Within "no hits" category, 308 different proteins/peptides

were encoded by 559 cDNAs. BLAST search of these molecules could not retrieve similar molecules from known public database, revealing that 49.82% of ESTs from a constructed *S. subspinipes dehaani* venom gland are currently unknown. A phylogenetic tree was constructed using these present unknown protein molecules, which indicated that they could be grouped into 8 families; the constructed phylogenetic tree and members belonging to each family are provided in the Supporting Information, Figures S40–S48. Comparing with well characterized snake venoms, venom gland transcriptome survey results gave 7.2, 10.57, and 19.19% of ESTs (calculated both unidentified clusters and singletons) among three saw-scaled vipers of *Echis coloratus*, *E. pyramidum leakeyi* and *E. carinatus sochureki*, respectively.²⁵ The "no hits" molecules in these 3 snake species were less than 20%. Authors estimated that a minimum of 800 EST sequences were required to provide an accurate representation of the three viper gland ESTs, and they sequenced 1070, 1078, and 1157 clones for *Echis coloratus*, *E. pyramidum leakeyi* and *E. carinatus sochureki*, respectively.²⁵ Above results not only reflect the current status of centipede venom research, but also provide clues to search for novel structures of molecules with specific biological functions. Our results reinforce the concept pointed by Rate

Table 3. Best-Hit Homology Search Results for Allergen 3 (SSD552) and Allergen 5 (SSD1080) by NCBI Database Blasting

accession	description	species	query coverage (%)	E-value	max identity (%)
SSD552 (KC144549)					
EFX87811.1	hypothetical protein DAPPUDRAFT_96576	<i>Daphnia pulex</i>	96.00	2×10^{-36}	37.00
ABZ79691.1	SCP-like protein	<i>Diabrotica virgifera virgifera</i>	90.00	2×10^{-34}	40.00
XP_002403125.1	CRISP3: cysteine-rich secretory protein, putative	<i>Ixodes scapularis</i>	91.00	7×10^{-34}	34.00
EFX62063.1	hypothetical protein DAPPUDRAFT_302039	<i>Daphnia pulex</i>	94.00	1×10^{-33}	40.00
XP_001862170.1	venom allergen 3	<i>Culex quinquefasciatus</i>	85.00	3×10^{-31}	36.00
SSD1080 (KC145067)					
EFX62063.1	hypothetical protein DAPPUDRAFT_302039	<i>Daphnia pulex</i>	94.00	8×10^{-38}	41.00
EFX87811.1	hypothetical protein DAPPUDRAFT_96576	<i>Daphnia pulex</i>	97.00	8×10^{-34}	36.00
XP_002403125.1	CRISP3: cysteine-rich secretory protein, putative	<i>Ixodes scapularis</i>	87.00	7×10^{-33}	35.00
ABZ79691.1	SCP-like protein	<i>Diabrotica virgifera virgifera</i>	99.00	2×10^{-32}	35.00
XP_001845617.1	venom allergen 5	<i>Culex quinquefasciatus</i>	85.00	4×10^{-31}	39.00

et al. that centipede venoms are a neglected but important source of novel bioactive compounds.⁷

Notably, within “secreted proteins/peptides” category, 139 different toxin molecules were encoded by 427 sequenced independent clones, indicating that the predicted toxins were largely expressed by the animals. On the basis of the random nucleotide sequencing results of “secreted proteins/peptides” category, the relative abundance of neurotoxins (including neurotoxin 4, 5 and 6), K⁺ channel inhibitors (10 groups), Ca²⁺ channel inhibitors (4 groups), venom allergen (including venom allergen 3 and 5), and Na⁺ channel inhibitors (1 group with 5 similar isoforms) were calculated to be 42.16, 31.62, 9.13, 5.15, and 4.45%, respectively (Figure 2B). These results strongly indicated that neurotoxins and various ion channel acting components are main fractions existing in the venom. The enzymes were not widely distributed in *S. subspinipes dehaani* venom since only 2 PLA₂ (encoded by 3 cDNAs) and 2 serine proteases (encoded by 2 cDNAs) were identified in the present investigation. The unknown function components (relative abundance of 6.3%) within “secreted proteins/peptides” category implied that these molecules had lower similarity with known proteins/peptides deposited in different databases but could be purified from crude centipede venom. The importance of these molecules contained in the centipede venoms needs be further elucidated.

Taken together, *S. subspinipes dehaani* venom gland transcriptomic analysis reveals that neurotoxins and ion channel acting components are diversely distributed in the venoms and there are many unknown proteins/peptides with novel structures and biological activities that need to be further investigated.

Functional Annotation and Bioinformatic Analysis of Secreted Proteins/Peptides

Functional annotated “secreted proteins/peptides” group and bioinformatic analysis of these members were summarized in Table 2. K⁺ channel inhibitors were the most divergent groups identified by both functional assays and bioinformatic analysis. Totally, 10 groups were found in venom gland transcriptome. Nine purified molecules showed K⁺ channel inhibitory activities and greatly help us to annotate functions of their corresponding cDNAs. Six novel groups of K⁺ channel inhibitors, namely, groups 1, 2, 3, 8, 9 and 10, were first functionally annotated to their corresponding cDNAs. Except group 3, the primary structures of these novel K⁺ channel inhibitor groups were quite different from other known molecules since no obvious similarity was found for each member(s) of different groups by BLAST search. Interestingly, the venom of *S. subspinipes*

dehaani contained high MW components that acted as K⁺ channel inhibitors. Members belonging to group 1, 2, and 3 had a MW ranges between 20 and 22.5 kDa. The remaining K⁺ channel inhibitor groups were peptide inhibitors, and their MW were below 9 kDa.

Four groups of voltage-gated calcium channel inhibitors were identified by present investigation (Table 2). Members belonging to group 1, 2, and 4 were purified from crude venom, and their calcium channel inhibitory activities were identified. Group 3 was classified on the basis of the BLAST results, and members of this group showed a high similarity with omega-slptx-ssm2a from *S. subspinipes mutilans*.²⁰

Only one group of peptides having voltage-gated sodium channel activity was identified. This group contained 5 members. Notably, SSD800, SSD591, SSD696 and SSD364 of this group all coded for an identical mature peptide, and the difference among these molecules existed in the signal peptide segments.

Recently, in depth analysis of ion channel acting components from the venoms of *S. subspinipes mutilans* was carried out. TTX-S Na_v channel inhibitor family of μ -SLPTX-Ssm1a; K_v channel inhibitor families of κ -SLPTX-Ssm1a, -Ssm2a, and -Ssm3a; Ca_v channel activation family of ω -SLPTX-Ssm1a and Ca_v channels inhibitory family of ω -SLPTX-Ssm2a were characterized.²⁰ However, no identical ion channel acting molecules was found between *S. subspinipes dehaani* and *S. subspinipes mutilans* species, indicating that rich diversity of ion channel acting molecules existed in centipede venoms.

Neurotoxins were the most abundant toxins as revealed by transcriptomic analysis of a constructed *S. subspinipes dehaani* venom gland cDNA library (Table 2). Since neurotoxin activity assays were not carried out in the present research, the classification of neurotoxin came from BLAST search results. Three groups of neurotoxins, namely, neurotoxin 4, 5 and 6, were found in *S. subspinipes dehaani*. Neurotoxin 4 group had only 4 different members, but their corresponding cDNAs were very high; totally, 43 cDNAs coding for these 4 molecules and SSD996 repeated for 38 times were observed. Similar patterns were also found in both neurotoxin 5 and neurotoxin 6 groups, indicating that neurotoxins play an important role in centipede venoms. Wide distribution of neurotoxins and ion channel acting components might be closely related with pain, a consistent symptom of centipede bites.^{4,19}

Allergens were also abundant toxins in *S. subspinipes dehaani* venom. Venom allergen 3 group and venom allergen 5 group were classified by BLAST search, suggesting that these molecules were similar to other known molecules. Only 1

Table 4. Best-Hit Homology Search Results for Two Predicated Serine Proteases by NCBI Database Blasting

accession	description	species	query coverage (%)	E-value	max identity (%)
SSD1134 (KC145121)					
ADB66714.1	trypsin 3	<i>Panulirus argus</i>	88.00	4×10^{-58}	44.00
ACE80257.1	trypsin	<i>Marsupenaeus japonicus</i>	88.00	1×10^{-57}	45.00
CAA75311.1	trypsin	<i>Litopenaeus vannamei</i>	88.00	6×10^{-57}	45.00
CAA75310.1	trypsin	<i>Litopenaeus vannamei</i>	88.00	3×10^{-56}	44.00
AEZ67461.1	trypsinogen 1	<i>Litopenaeus vannamei</i>	88.00	3×10^{-56}	44.00
SSD1135 (KC145122)					
CAA75310.1	trypsin	<i>Litopenaeus vannamei</i>	83.00	3×10^{-27}	50.00
AEZ68613.1	trypsinogen 2	<i>Litopenaeus vannamei</i>	83.00	5×10^{-27}	50.00
AAD00320.1	plasminogen activator sPA	<i>Scolopendra subspinipes</i>	91.00	2×10^{-26}	49.00
ADB66714.1	trypsin 3	<i>Panulirus argus</i>	83.00	2×10^{-26}	48.00
ACE80257.1	trypsin	<i>Marsupenaeus japonicus</i>	83.00	2×10^{-26}	49.00

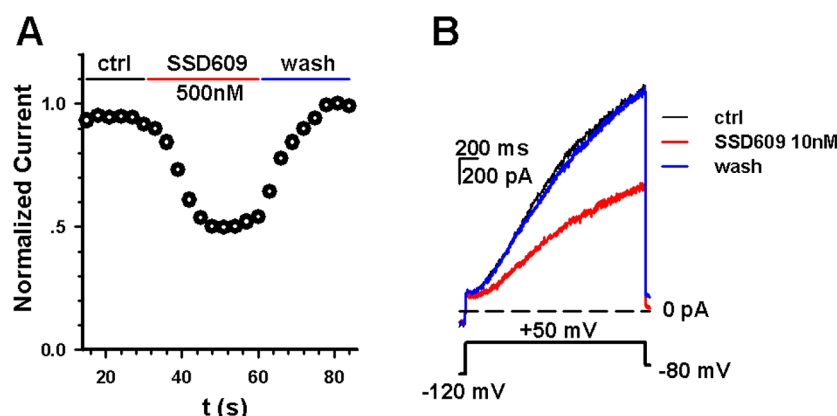


Figure 3. The effects of SSD609 on KCNQ channel. (A) Action of SSD609 on whole-cell patches from HEK293 cells transfected with cDNA encoding KCNQ1. The representative currents were activated by a voltage step from -180 to 50 mV for 2 s, in the presence and absence of 500 nM SSD609. A voltage protocol is plotted at the bottom of the current trace. (B) The peak currents from the above patches are plotted as a function of elapsed time. Currents were elicited with repetitive voltage steps to 50 mV with a time interval of 3 s, before, during, and after the application of 500 nM SSD609 as indicated by the horizontal bars.

and 15 members belonging to venom allergen 3 group and venom allergen 5 group were observed, respectively. Most *S. subspinipes dehaani* venom allergens had a MW between 20 and 22 kDa and composed of 187 or 191 amino acids for their mature proteins (Table 2). The only allergen 3 group member (SSD552) and a representative allergen 5 group member (SSD1080) were used to do blast search against NCBI database, and the results were shown in Table 3. Both SSD552 and SSD1080 showed a similar pattern with their best-hit molecules with an identity around 35–40% and sequence coverage of 85–97%, respectively. Interestingly, SSD43-coding mature protein, a member of venom allergen 5 group, had trypsin inhibitory activity.

At present, most of the identified allergens were from stinging insects of the Hymenoptera order. The stinging insects that cause severe allergic reactions include bees (honeybees, bumblebees), vespids (yellow jackets, hornets, wasps), and fire ants (*Solenopsis* genus).^{26,27} Honeybee venom contains the major allergens Api m 1 (phospholipase A₂), Api m 2 (hyaluronidase), Api m 3 (acid phosphatase), Api m 4 (melittin), and Api m 7 (serine protease), along with at least one minor allergen, Api m 6, which shows 42% IgE responsiveness.²⁷ However, none of centipede venom allergens were confirmed by venom skin tests or a radioallergosorbent test using IgE antibodies. Our present data and other investigations strongly suggest that allergens seem to be an important component in centipede venoms; molecules

belonging to venom allergen 5 group were also found in venoms of *S. viridicornis nigra* and *S. angulata* species.⁷ Comparing with the determined N-terminal amino acid sequencing results of 4 members belonging to venom allergen 5 group of *S. viridicornis nigra* and *S. angulata*, *S. subspinipes dehaani* allergens showed no similarity with those from *S. viridicornis nigra* and *S. angulata*, indicating that allergens were diversified among centipede species. Nevertheless, members belonging to venom allergen 5 group seem largely expressed in different centipede species. Swelling deterioration after initial improvement of centipede victims could be due to infection or hypersensitive reactions.⁸ Allergens contained in the centipede venoms, together with other allergic molecules, might be contributing a lot in hypersensitive reactions and causing localized itchiness.

Serine protease seemed to be not largely expressed by venom gland of *S. subspinipes dehaani*. Only 2 different but similar molecules could be found by transcriptomic analysis, and each molecule was coded by a cDNA as indicated in Table 2. Blast search against NCBI database results of each protease (SSD1134 and SSD1135) were seen in Table 4. Both SSD1134 and SSD1135 showed best match trypsin-like enzymes from prawn species of *Litopenaeus vannamei*. Identity around 45 and 50% with sequence coverage of 88 and 93% were found for SSD1134 and SSD1135, respectively.

PLA₂ also showed a similar pattern as serine proteases in *S. subspinipes dehaani* venom. Three cDNAs coding for 2 similar

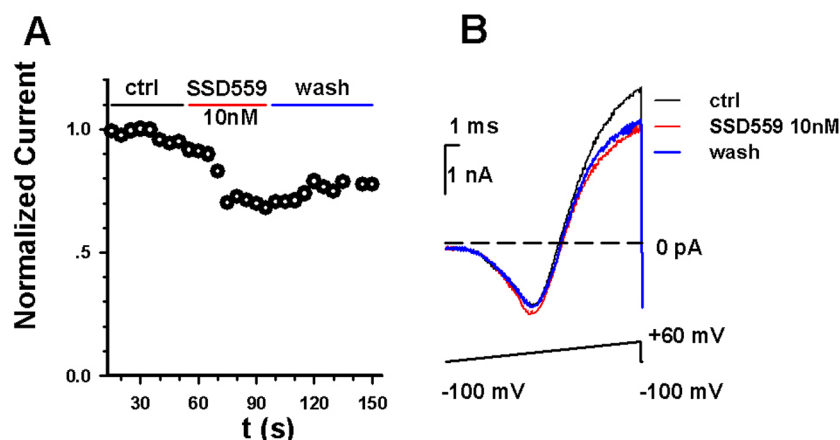


Figure 4. The effects of SSD559 on DRG neuron. (A) Traces were obtained from whole-cell patches from DRG neurons. The representative currents were activated by a ramp voltage potential from -100 to 60 mV for 20 ms, in the presence and absence of 10 nM SSD559. A voltage protocol is plotted at the bottom of the current trace. (B) The peak currents from the above patches are plotted as a function of elapsed time. Currents were elicited with repetitive voltage potential with a time interval of 5 s, before, during, and after the application of 10 nM SSD559 as indicated by the horizontal bars.

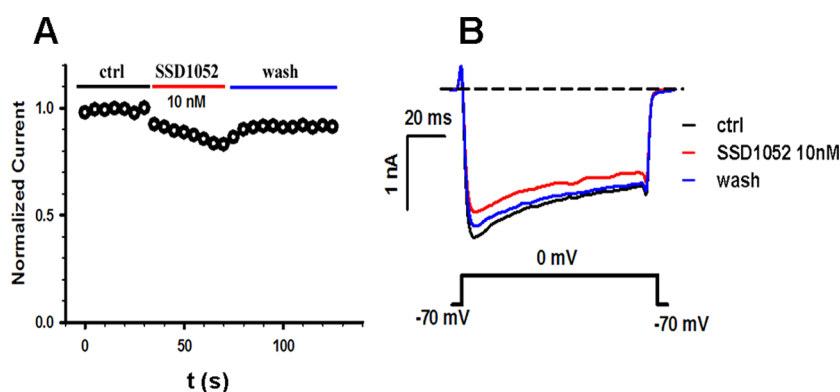


Figure 5. The effects of SSD1052 on DRG neuron. (A) Traces were obtained from whole-cell patches from DRG neurons. The representative currents were activated by a voltage step from -70 to 0 mV for 100 ms, in the presence and absence of 10 nM SSD1052 as indicated by the horizontal bars. (B) The peak currents are plotted as a function of elapsed time. Currents were elicited with repetitive voltage potential with a time interval of 5 s, before, during, and after the application of 10 nM SSD1052.

PLA₂ precursors were observed. The calculated pI values for the deduced mature PLA₂ were 4.49 and 4.63 for SSD387 and SSD1043, respectively. Thus, only two kinds of acidic PLA₂ molecules were found in our constructed *S. subspinipes dehaani* venom gland cDNA library from a single species. Our results indicated that enzymes were a minor component of centipede venom, at least for *S. subspinipes dehaani* species.

Functional Assay and Characterization of Purified Proteins/Peptides

KCNQ1 Channel Inhibitor

A peptide with N-terminal amino acid sequence of ADDKC-EDSLRREIACTKCRDRVRTDD and observed MW of 5624.5 Da was isolated (Figure 1I). The peptide was encoded by SSD609. Figure 3 illustrated that 500 nM purified SSD609-coding mature peptide reversibly blocked the currents encoded by KCNQ1 coexpressed with KCNE1 in HEK 293 cells. The estimated IC_{50} is lower than 500 nM, indicating that SSD609 mature peptide may be a good candidate of peptidyl inhibitors for KCNQ1 channels.

K_v Channel Inhibitor

A peptide with N-terminal amino acid sequence of NLIYE-CKWADSIRLKDKNP THEFCKK and observed MW of

8556.2 Da was isolated (Figure 1H). Although the determined N-terminal amino acid sequence matched well with SSD559, the deduced mature peptide of SSD559 had a theoretical mass of 8513.60 Da. Thus, the isolated peptide should be regarded as a homologue of SSD559. Figure 4 shows that 10 nM purified peptide irreversibly blocked the currents evoked by a ramp voltage command. The blocked outward current was obviously the potassium currents, as the inward sodium current was not affected at all. The estimated IC_{50} is around 10 nM, suggesting that the peptide may be a good candidate of K_v channels.

Calcium Channel Inhibitor

A peptide with N-terminal amino acid sequence of EVIKR-DIPYKKRKFPYKSECLK and observed MW of 6028.1 Da was isolated (Figure 1I). The peptide was encoded by SSD1052. Figure 5 indicated that 10 nM SSD1052-coding mature peptide reversibly blocked the 8.6% of calcium currents recorded from DRG neurons. Since the DRG cells almost contain all kinds of calcium channels including L-, N-, P/Q- and R-type calcium channels, of which each has about 25% of total calcium currents, the real target of SSD1052-coding mature peptide need to be further investigated in the future.

Platelet Aggregating Inducer

Peak 7 from Mono Q anionic exchange column composed of α - and β -chains as revealed by SDS-PAGE (Figure 1B, insert). Two distinct bands with apparent MW of 45 kDa (α -subunit) and 21 kDa (β -subunit) under reducing conditions and their N-terminal amino acid sequences were separately determined by PVDF protocol, respectively. The protein was encoded by SSD14, as both α - and β -chains of the purified could be found in the SSD14 encoding protein precursor. SSD14 coding mature protein dose-dependently induced human platelet aggregation on both PRP and washed platelet suspensions (Figure 6). At the concentration 3.2 $\mu\text{g/mL}$, purified SSD14-

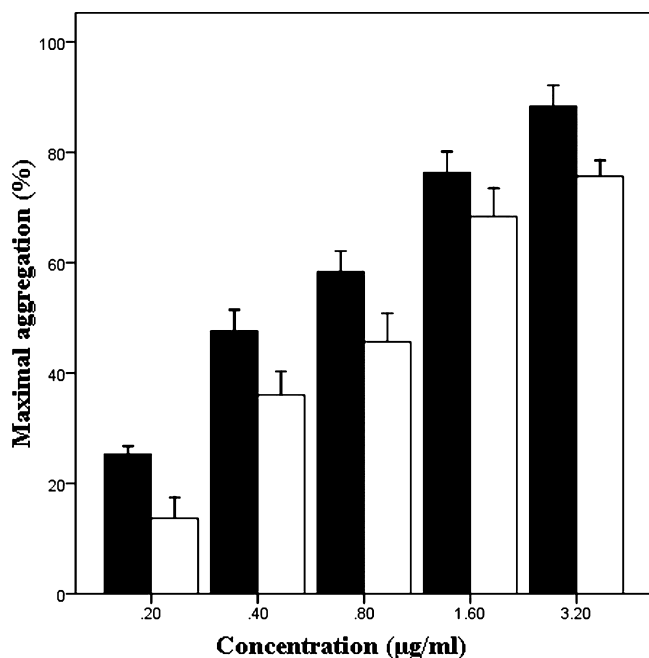


Figure 6. Dose-dependent platelet aggregation induced by SSD14. Human PRP (black bar) and washed platelet suspensions (white bar) were incubated with various concentrations of SSD14 (0.2, 0.4, 0.8, 1.6, 3.2 $\mu\text{g/mL}$) monitoring for 5 min with constant stirring at 37 $^{\circ}\text{C}$. Values are mean $\pm$ SEM of three experiments. A positive control is given for platelet aggregation, such as *stajnulxin*.

coding mature protein induced the maximal response. At the concentration 5 $\mu\text{g/mL}$, SSD14-coding mature protein showed obvious hemolytic activity against mouse and rabbit red blood cells (35 and 65%, respectively). Meanwhile, the molecule showed no hemolytic activity against human red blood cells, even at the concentration of 100 $\mu\text{g/mL}$.

Phylogenetic analysis indicated that SSD14 showed high similarity with human gamma-glutamyl transferase (GGT).²⁸ GGT is present in the cell membranes of many tissues, including the kidneys, liver, bile duct, pancreas, spleen, brain and heart.²⁹ It catalyzes the transfer of gamma-glutamyl moiety to target molecules, including amino acids, peptides, and water leaving the cysteine product to preserve intracellular homeostasis of oxidative stress.³⁰ High activity of GGT is specific for diseases of the liver, biliary tract, and pancreas. Its value is applied to the clinical diagnosis of various liver diseases. Many patients with liver disease also have higher GGT value and lower platelet amount. However, up to now, no report links these phenomena. The result of the present investigation

implies that certain members of GGT can both catalyze the transfer reaction and induce platelet aggregation.

Blood Coagulation Anticoagulator

The fraction eluted at 40.357 min from RT-HPLC C_{18} column (Figure 1C) showed anticoagulant activity. The protein contained in this fraction was encoded by SSD431 as both determined N-terminal amino acid sequence and observed mass of purified protein matched well with deduced mature protein of SSD431 (Table 1). Purified SSD431-coding mature protein prolonged recalcification time of PPP in vitro in a dose-dependent manner (Figure 7). The recalcification time was significantly prolonged when purified fraction was added above 0.4 $\mu\text{g/mL}$. At the concentration of 3.2 $\mu\text{g/mL}$, the recalcification time was longer than 120 min compared with the control (about 2.5 min).

Characterization of Purified PLA_2 s

The fractions eluted at 40.133, 41.546, 42.176, and 43.689 min on RT-HPLC C_{18} (Figure 1D) had very similar N-terminal amino acid sequences, but their observed MW were slightly different (Table 1). Blast search using determined N-terminal amino acid sequences against public database suggested that all four molecules belonged to PLA_2 family, indicating that the diversity of PLA_2 contained in *S. subspinipes dehaani* venom was rich. The specific PLA_2 activities of 40.133 min peak and 42.176 min peak were determined to be 840 and 110 $\mu\text{M}/\text{min}/\text{mg}$ under our assay conditions, respectively. However, the anticoagulant activity of 40.133 min peak was weaker than that of 42.176 min peak, suggesting that the anticoagulant activity was not directly related with PLA_2 activity for the isolated two enzymes. Meanwhile, both isolated enzymes had no obvious cytotoxic effects against used HeLa and HCT-116 cell lines at 80 $\mu\text{g/mL}$.

Only 3 cDNAs encoding for 2 different PLA_2 molecules were found in the present investigations (Table 2). The determined N-terminal amino acid sequences of 41.546 min peak and 42.176 min peak were identical. Although the determined N-terminal amino acid sequences of both peaks matched well with SSD387, the deduced mature protein of SSD387 had a theoretical mass of 13426.81 Da, which was slightly different than the observed MW of the isolated 2 proteins. Thus, the isolated 2 enzymes should be regarded as homologues of SSD387. Similar situations existed in 43.689 min peak, which was supposed to be a homologue of SSD 1043. However, coding cDNA for the 40.133 min peak could not be found in the present work. SSD387 of *S. subspinipes dehaani* showed higher identity (61%) with PLA_2 from *S. viridis*.⁹ Interestingly, isolated PLA_2 from *S. viridis* and 2 PLA_2 s from *S. subspinipes dehaani* (SSD387 and SSD 1043) were all acidic PLA_2 s, and the mature proteins were all composed of 119 amino acid residues. Biochemically isolated molecules that could not match with cloned sequences are very rare, and the reason might be one of the following: (1) There is a rich diversity of PLA_2 existing in the *S. subspinipes dehaani* venom. Purification was originated from pooled venoms of 400 individuals but the constructed cDNA library was from a single species. (2) The nucleotide sequencing depth was not high enough to cover all the cDNAs contained in the constructed cDNA library.

PLA_2 is ubiquitous in many venomous snakes such as cobras, rattlesnakes and pit viper.^{31,32} Besides digestive roles, snake venom PLA_2 s exert an amazing variety of toxic and pharmacological effects, which include neurotoxic, myotoxic, hemolytic, edematogenic, hyperalgesic, pro-inflammatory,

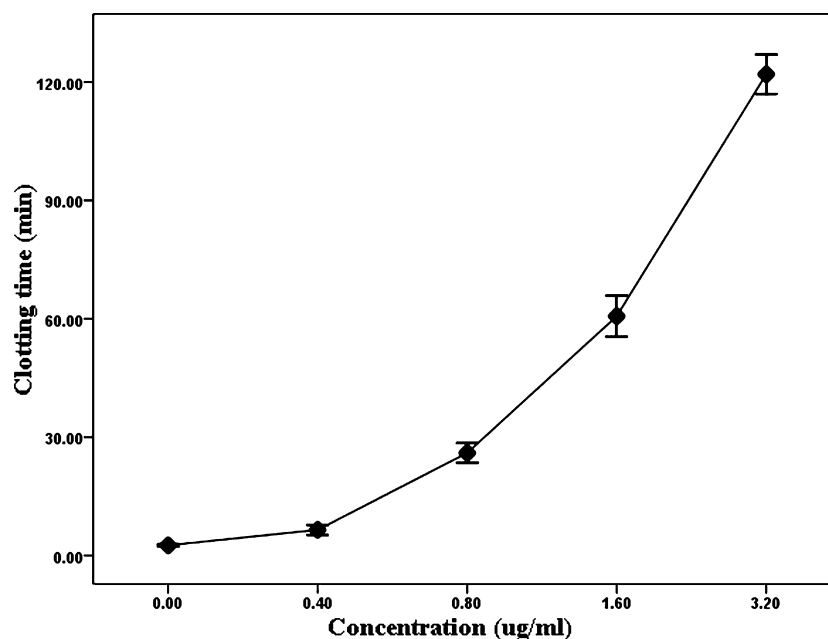


Figure 7. Dose-dependent anticoagulation activity of SSD431. Values are the means of three determinations.

hypotensive, platelet-aggregation inhibitory, anticoagulant, cytotoxic, and bactericidal activities.³³ In humans, PLA₂-generated inflammatory lipid mediators play an important role in skin inflammatory/allergic conditions, including skin permeability, maintaining the skin pH, membrane phospholipid remodeling and wound healing.³⁴ Skin inflammatory diseases by centipede stings are most commonly treated with corticosteroids, especially topical preparations. The delayed local hypersensitivity reaction of centipede sting was alleviated in clinic practice by topical steroids.⁶ However, these drugs have side effects, especially under long-term use. The control of PLA₂ activities provides a promising approach for the treatment of inflammatory/allergic skin diseases. Thus, investigation to determine the role of specific centipede PLA₂s in skin pathologies might be beneficial for treating centipede victims and develop more effective PLA₂ controlling therapeutic strategies.

Trypsin Inhibitor

The inhibitory effects of fractions on the hydrolysis of synthetic chromogenic substrates by serine protease, such as trypsin, thrombin, chymotrypsin and elastase, were investigated. The fraction eluted at 52.197 min from RT-HPLC C₁₈ column (Figure 1D) showed trypsin inhibiting activity (Figure 8). After incubation in boiling water for 10 min, the trypsin inhibitory activity of the molecule was not lost, revealing that the molecule is highly thermally stable. The purified trypsin inhibitor was determined to be encoded by SSD43, which belonged to venom allergen 5 group by blast search (Table 1).

CONCLUSION

Centipedes have adapted to eat a variety of different prey. The venoms stored in their maxillipeds are very powerful weapons to paralyze or kill prey. Centipede venom is a bank of toxins with different pharmacological properties. In this study, we demonstrated that neurotoxins, ion channel acting components and venom allergens were the main fractions contained in the centipede venom of *S. subspinipes dehaani*. Meanwhile, venom gland transcriptomic analysis revealed that there were a large

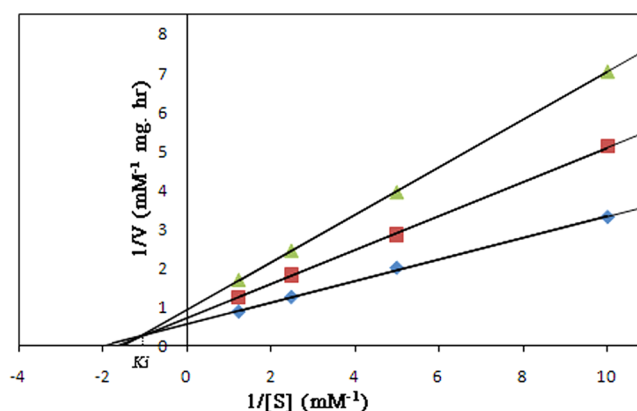


Figure 8. The inhibitory activities of SSD43 against trypsin. Lineweaver–Burk plots for trypsin activity on H-D-Phe-Pip-Arg-pNA in the absence (blue $\blacklozenge$) or in the presence of SSD43 at 100 nM (red $\blacksquare$) and 400 nM (yellow $\blacktriangle$), respectively. Values are the means of three determinations.

number of components with lower structural and functional similarities with other known proteins/peptides existing in the venom. Purified proteins/peptides showed different pharmacological properties, including (1) platelet aggregating activity; (2) anticoagulant activity; (3) PLA₂ activity; (4) trypsin inhibitory activity; (5) voltage-gated potassium channel inhibiting activities; (6) voltage-gated sodium channel inhibiting activities; and (7) voltage-gated calcium channel activities. This is the largest number of proteins/peptides reported from a centipede species. Most of them showed no significant similarity to other protein sequences deposited in the public database. The toxin nature of *S. subspinipes dehaani* revealed by the present work not only provides a large number of protein/peptide candidates with medical/pharmaceutical significance, but also helps us to understand the symptoms caused by centipede bite.

■ ASSOCIATED CONTENT

■ Supporting Information

Supplemental table and figures. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*(J.-P. Ding) Tel/Fax: +86-27-87792153. E-mail: jpdng@mail.hust.edu.cn. (W.-H. Lee) Tel: +86-871-5192476. Fax: +86-871-5191823. E-mail: leewh@mail.kiz.ac.cn. (Y. Zhang) Tel/Fax: +86-871-5198515. E-mail: zhangy@mail.kiz.ac.cn.

Author Contributions

[#]These authors contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by grants from National Basic Research Program of China (973 Program, 2010CB529800) and The National Natural Science Foundation of China (NSFC-Yunnan-Union-Funding U1132601, 31071926 and 31270835). Y206A41181 from Kunming Institute of Zoology and 2011CI139 from Yunnan Province are also acknowledged.

■ REFERENCES

- (1) Undheim, E. A.; King, G. F. On the venom system of centipedes (Chilopoda), a neglected group of venomous animals. *Toxicon* **2011**, *57* (4), 512–24.
- (2) Edgecombe, G. D.; Giribet, G. Evolutionary biology of centipedes (Myriapoda: Chilopoda). *Annu. Rev. Entomol.* **2007**, *52*, 151–70.
- (3) Uzel, A. P.; Steinmann, G.; Bertino, R.; Korsaga, A. Necrotizing fasciitis and cellulitis of the upper limb resulting from centipede bite: two case reports. *Chir. Main* **2009**, *28* (5), 322–5.
- (4) Knysak, I.; Martins, R.; Bertim, C. R. Epidemiological aspects of centipede (Scolopendromorpha: Chilopoda) bites registered in greater S. Paulo, SP, Brazil. *Rev. Saude Publica* **1998**, *32* (6), 514–8.
- (5) Mohri, S.; Sugiyama, A.; Saito, K.; Nakajima, H. Centipede bites in Japan. *Cutis* **1991**, *47* (3), 189–90.
- (6) Fung, H. T.; Lam, S. K.; Wong, O. F. Centipede bite victims: a review of patients presenting to two emergency departments in Hong Kong. *Hong Kong Med. J.* **2011**, *17* (5), 381–5.
- (7) Rates, B.; Bemquerer, M. P.; Richardson, M.; Borges, M. H.; Morales, R. A.; De Lima, M. E.; Pimenta, A. M. Venomic analyses of *Scolopendra viridicornis nigra* and *Scolopendra angulata* (Centipede, Scolopendromorpha): shedding light on venoms from a neglected group. *Toxicon* **2007**, *49* (6), 810–26.
- (8) Gomes, A.; Datta, A.; Sarangi, B.; Kar, P. K.; Lahiri, S. C. Isolation, purification & pharmacodynamics of a toxin from the venom of the centipede *Scolopendra subspinipes dehaani* Brandt. *Indian J. Exp. Biol.* **1983**, *21* (4), 203–7.
- (9) Gonzalez-Morales, L.; Diego-Garcia, E.; Segovia, L.; Gutierrez, M. d. C.; Possani, L. D. Venom from the centipede *Scolopendra viridis* Say: Purification, gene cloning and phylogenetic analysis of a phospholipase A₂. *Toxicon* **2009**, *54* (1), 8–15.
- (10) Peng, K.; Kong, Y.; Zhai, L.; Wu, X.; Jia, P.; Liu, J.; Yu, H. Two novel antimicrobial peptides from centipede venoms. *Toxicon* **2010**, *55* (2–3), 274–9.
- (11) You, W. K.; Sohn, Y. D.; Kim, K. Y.; Park, D. H.; Jang, Y.; Chung, K. H. Purification and molecular cloning of a novel serine protease from the centipede *Scolopendra subspinipes mutilans*. *Insect Biochem. Mol. Biol.* **2004**, *34* (3), 239–50.
- (12) Tang, X.; Zhang, Y.; Hu, W.; Xu, D.; Tao, H.; Yang, X.; Li, Y.; Jiang, L.; Liang, S. Molecular diversification of peptide toxins from the tarantula *Haplophelma hainanum* (*Ornithoctonus hainana*) venom based

on transcriptomic, peptidomic, and genomic analyses. *J. Proteome Res.* **2010**, *9* (5), 2550–64.

(13) Carbone, E.; Wanke, E.; Prestipino, G.; Possani, L. D.; Maelicke, A. Selective blockage of voltage-dependent K⁺ channels by a novel scorpion toxin. *Nature* **1982**, *296* (5852), 90–1.

(14) Mintz, I. M.; Venema, V. J.; Swiderek, K. M.; Lee, T. D.; Bean, B. P.; Adams, M. E. P-type calcium channels blocked by the spider toxin omega-Aga-IVA. *Nature* **1992**, *355* (6363), 827–9.

(15) Rogers, J. C.; Qu, Y. S.; Tanada, T. N.; Scheuer, T.; Catterall, W. A. Molecular determinants of high affinity binding of alpha-scorpion toxin and sea anemone toxin in the S3-S4 extracellular loop in domain IV of the Na⁺ channel alpha subunit. *J. Biol. Chem.* **1996**, *271* (27), 15950–62.

(16) Yang, S.; Liu, Z.; Xiao, Y.; Li, Y.; Rong, M.; Liang, S.; Zhang, Z.; Yu, H.; King, G. F.; Lai, R. Chemical punch packed in venoms makes centipedes excellent predators. *Mol. Cell. Proteomics* **2012**, *11* (9), 640–50.

(17) Yang, X.; Lee, W. H.; Zhang, Y. Extremely abundant antimicrobial peptides existed in the skins of nine kinds of Chinese odorous frogs. *J. Proteome Res.* **2012**, *11* (1), 306–19.

(18) Lee, W. H.; Li, Y.; Lai, R.; Li, S.; Zhang, Y.; Wang, W. Variety of antimicrobial peptides in the *Bombina maxima* toad and evidence of their rapid diversification. *Eur. J. Immunol.* **2005**, *35* (4), 1220–9.

(19) Zhang, Y.; Yu, G.; Wang, Y.; Xiang, Y.; Gao, Q.; Jiang, P.; Zhang, J.; Lee, W.; Zhang, Y. Activation of protease-activated receptor (PAR) 1 by frog trefoil factor (TFF) 2 and PAR4 by human TFF2. *Cell. Mol. Life Sci.* **2011**, *68* (22), 3771–80.

(20) Lee, W. H.; Zhang, Y.; Wang, W. Y.; Xiong, Y. L.; Gao, R. Isolation and properties of a blood coagulation factor X activator from the venom of king cobra (*Ophiophagus hannah*). *Toxicon* **1995**, *33* (10), 1263–76.

(21) Shiloah, J.; Klibansky, C.; de Vries, A.; Berger, A. Phospholipase B activity of a purified phospholipase A from *Vipera palestinae* venom. *J. Lipid Res.* **1973**, *14* (3), 267–78.

(22) Zhang, Y.; Wisner, A.; Xiong, Y. L.; Bon, C. A novel plasminogen activator from snake venom. Purification, characterization, and molecular cloning. *J. Biol. Chem.* **1995**, *270* (17), 10246–55.

(23) Dixon, M. The determination of enzyme inhibitor constants. *Biochem. J.* **1953**, *55* (1), 170–1.

(24) Liu, S.-B.; He, Y.-Y.; Zhang, Y.; Lee, W.-H.; Qian, J.-Q.; Lai, R.; Jin, Y. A novel non-lens betagamma-Crystallin and trefoil factor complex from amphibian skin and its functional implications. *PLoS ONE* **2008**, *3* (3), e1770.

(25) Casewell, N. R.; Harrison, R. A.; Wuster, W.; Wagstaff, S. C. Comparative venom gland transcriptome surveys of the saw-scaled vipers (Viperidae: Echis) reveal substantial intra-family gene diversity and novel venom transcripts. *BMC Genomics* **2009**, *10*, 564.

(26) Golden, D. B. Stinging insect allergy. *Am. Fam. Physician* **2003**, *67* (12), 2541–6.

(27) Hoffman, D. R. Hymenoptera venom allergens. *Clin. Rev. Allergy Immunol.* **2006**, *30* (2), 109–28.

(28) Courtay, C.; Oster, T.; Michelet, F.; Visvikis, A.; Diederich, M.; Wellman, M.; Siest, G. Gamma-glutamyltransferase: nucleotide sequence of the human pancreatic cDNA. Evidence for a ubiquitous gamma-glutamyltransferase polypeptide in human tissues. *Biochem. Pharmacol.* **1992**, *43* (12), 2527–33.

(29) Whitfield, J. B. Gamma glutamyl transferase. *Crit. Rev. Clin. Lab. Sci.* **2001**, *38* (4), 263–355.

(30) Goldberg, D. M. Structural, functional, and clinical aspects of gamma-glutamyltransferase. *CRC Crit. Rev. Clin. Lab. Sci.* **1980**, *12* (1), 1–58.

(31) Dennis, E. A.; Cao, J.; Hsu, Y.-H.; Magriotti, V.; Kokotos, G. Phospholipase A₂ enzymes: physical structure, biological function, disease implication, chemical inhibition, and therapeutic intervention. *Chem. Rev.* **2011**, *111* (10), 6130–85.

(32) Fernandez, J.; Lomonte, B.; Sanz, L.; Angulo, Y.; Maria Gutierrez, J.; Calvete, J. J. Snake venomomics of *Bothriechis nigroviridis* reveals extreme variability among palm pitviper venoms: different

evolutionary solutions for the same trophic purpose. *J. Proteome Res.* **2010**, 9 (8), 4234–41.

(33) Gutierrez, J. M.; Lomonte, B. Phospholipases A₂: Unveiling the secrets of a functionally versatile group of snake venom toxins. *Toxicon* **2012**, DOI: 10.1016/j.toxicon.2012.09.006.

(34) Dan, P.; Rosenblat, G.; Yedgar, S. Phospholipase A₂ activities in skin physiology and pathology. *Eur. J. Pharmacol.* **2012**, 691 (1–3), 1–8.

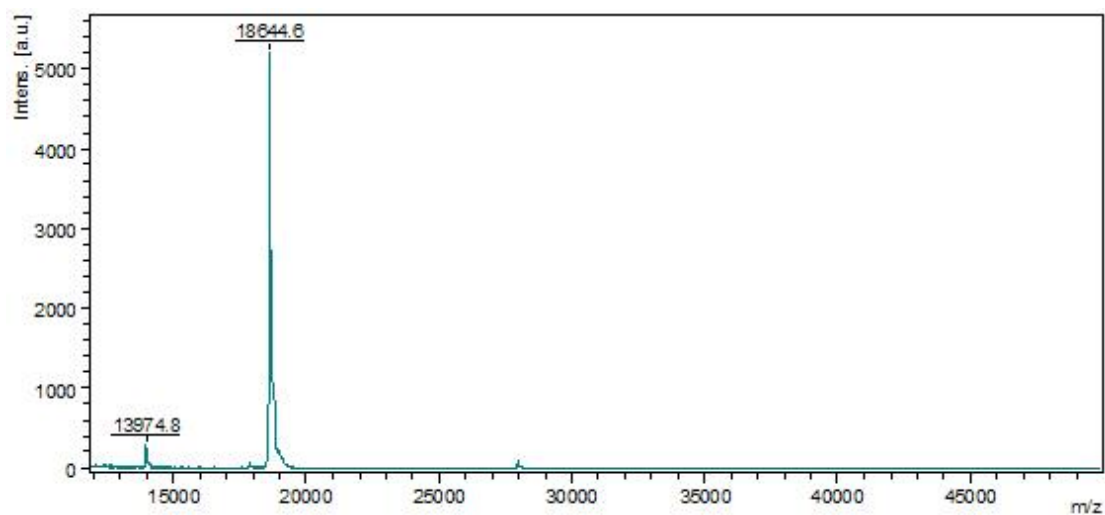


Figure S1. The mass spectrum of 33.462 min peak in Figure 1-C.

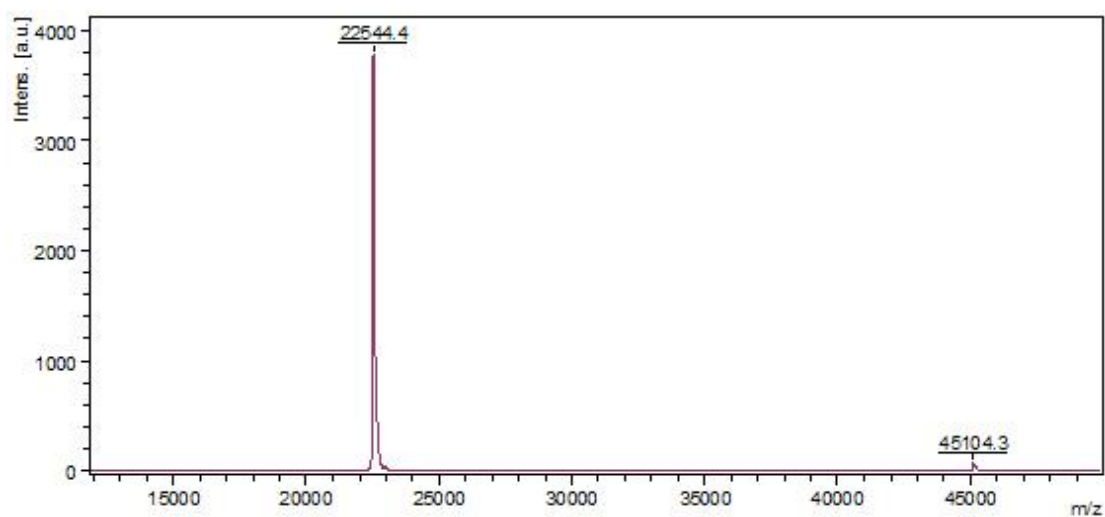


Figure S2. The mass spectrum of 35.044 min peak in Figure 1-C.

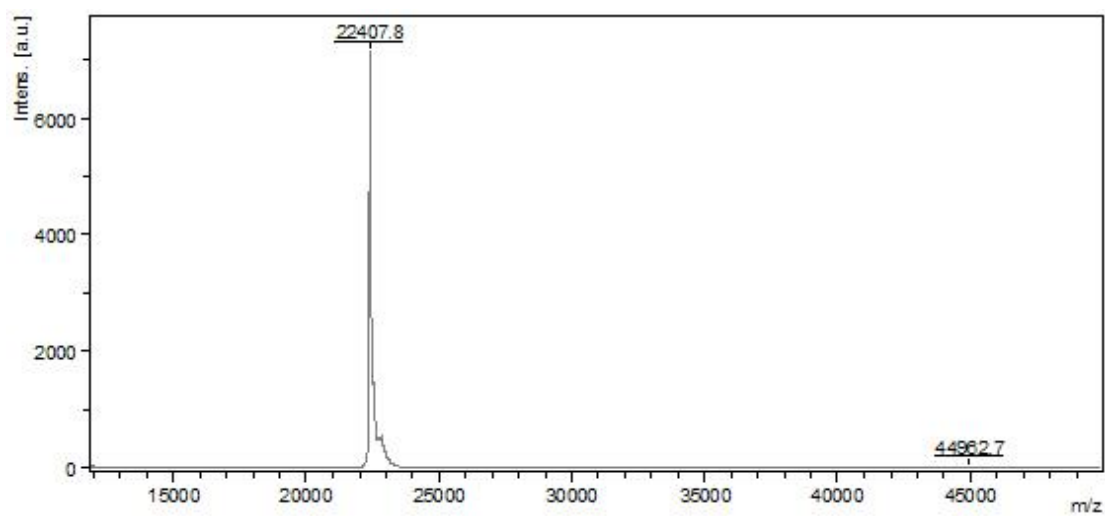


Figure S3. The mass spectrum of 37.985 min peak in Figure 1-C.

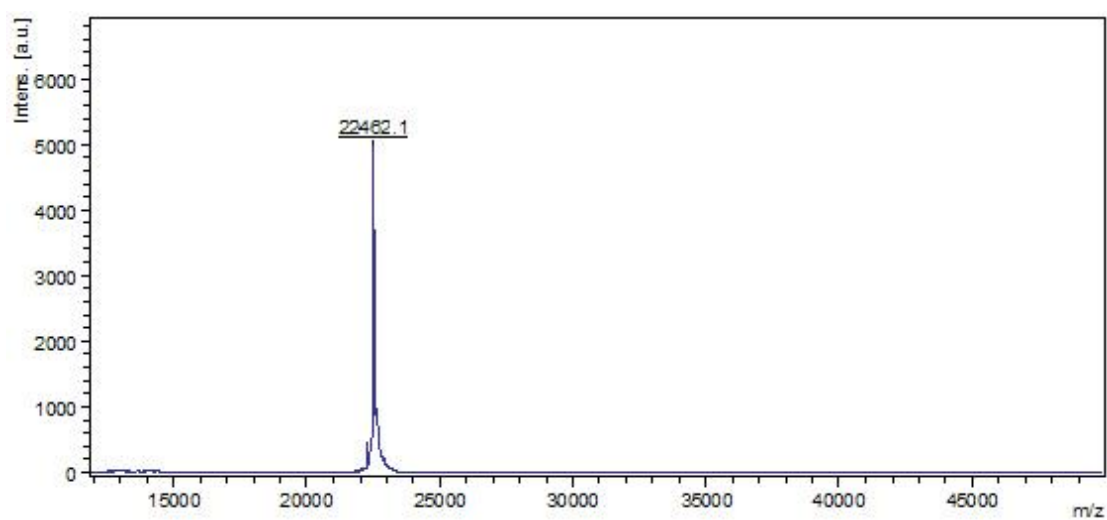


Figure S4. The mass spectrum of 39.623 min peak in Figure 1-C.

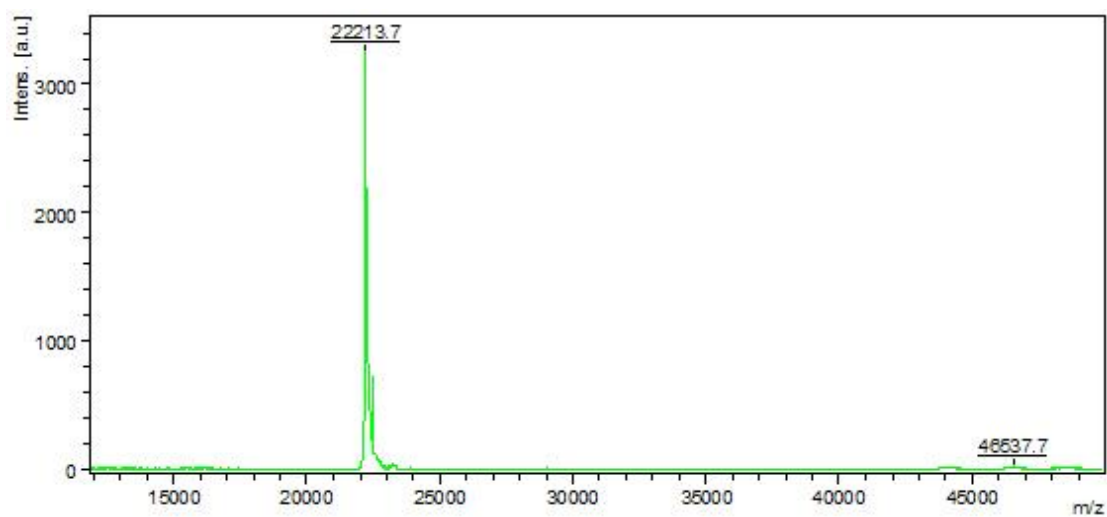


Figure S5. The mass spectrum of 40.357 min peak in Figure 1-C.

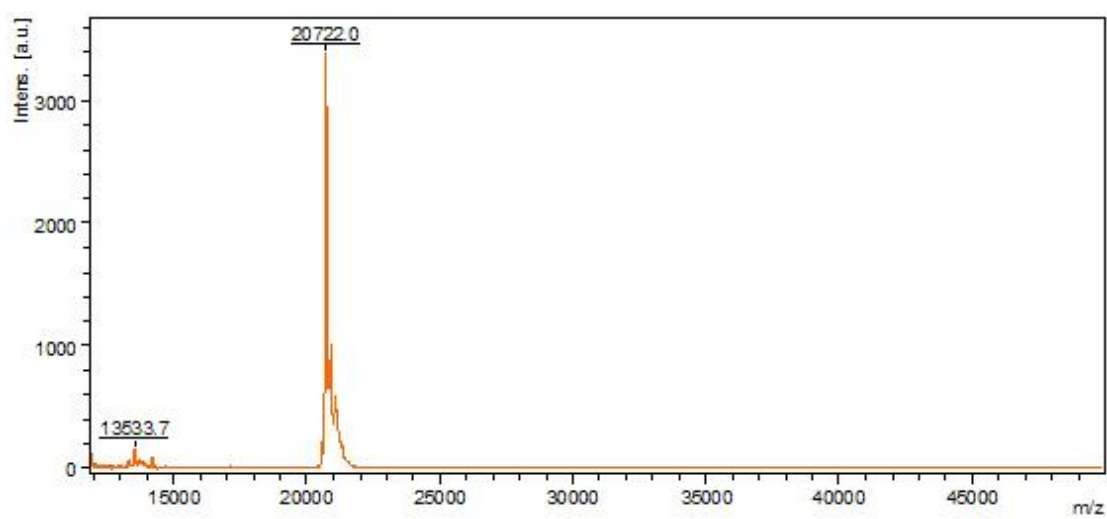


Figure S6. The mass spectrum of 44.626 min peak in Figure 1-C.

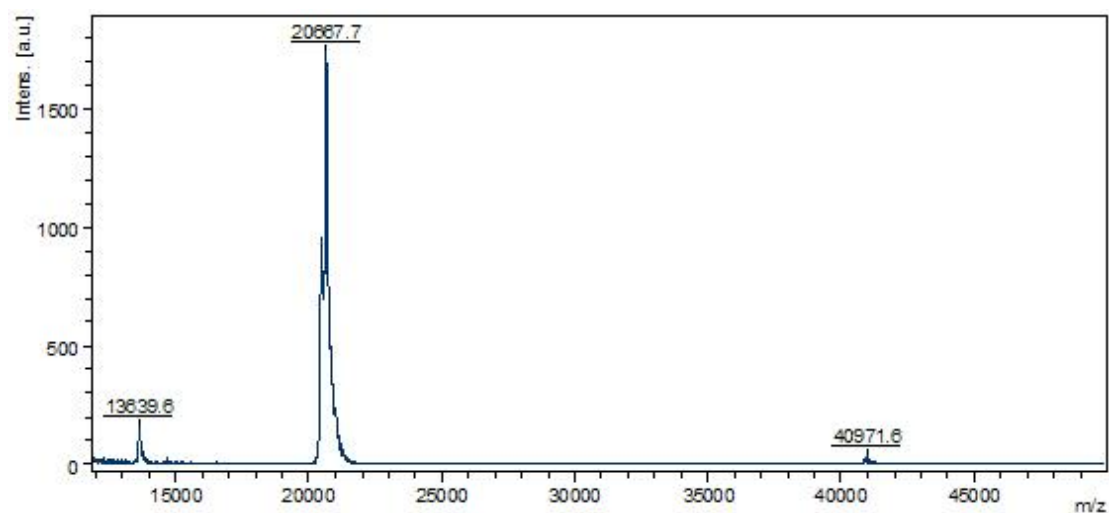


Figure S7. The mass spectrum of 47.071 min peak in Figure 1-C.

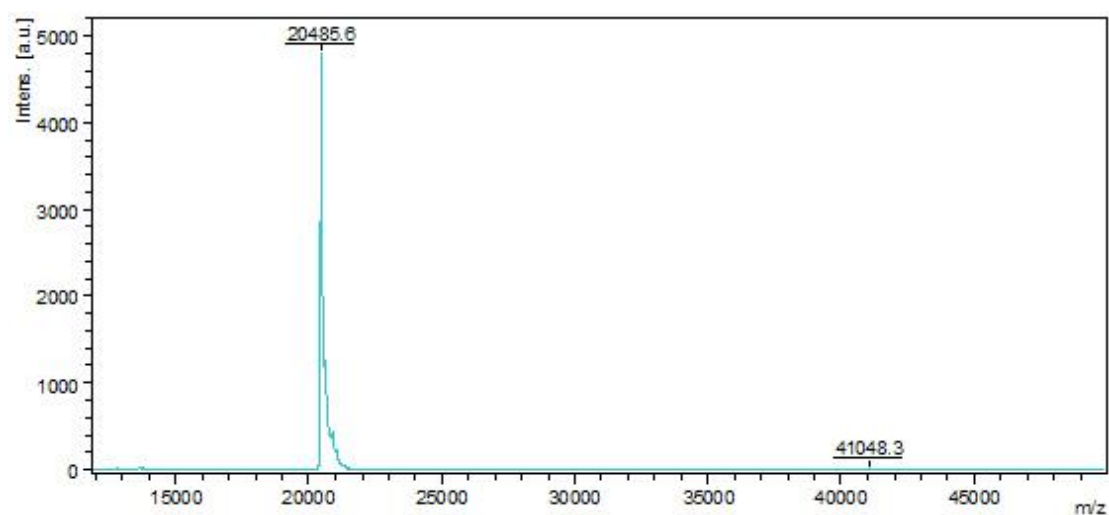


Figure S8. The mass spectrum of 48.345 min peak in Figure 1-C.

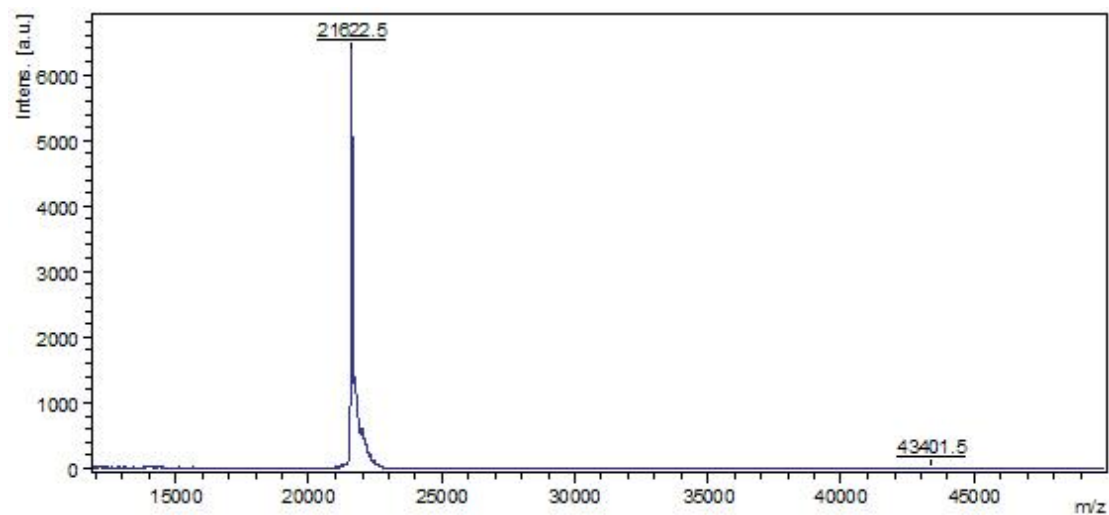


Figure S9. The mass spectrum of 51.340 min peak in Figure 1-C.

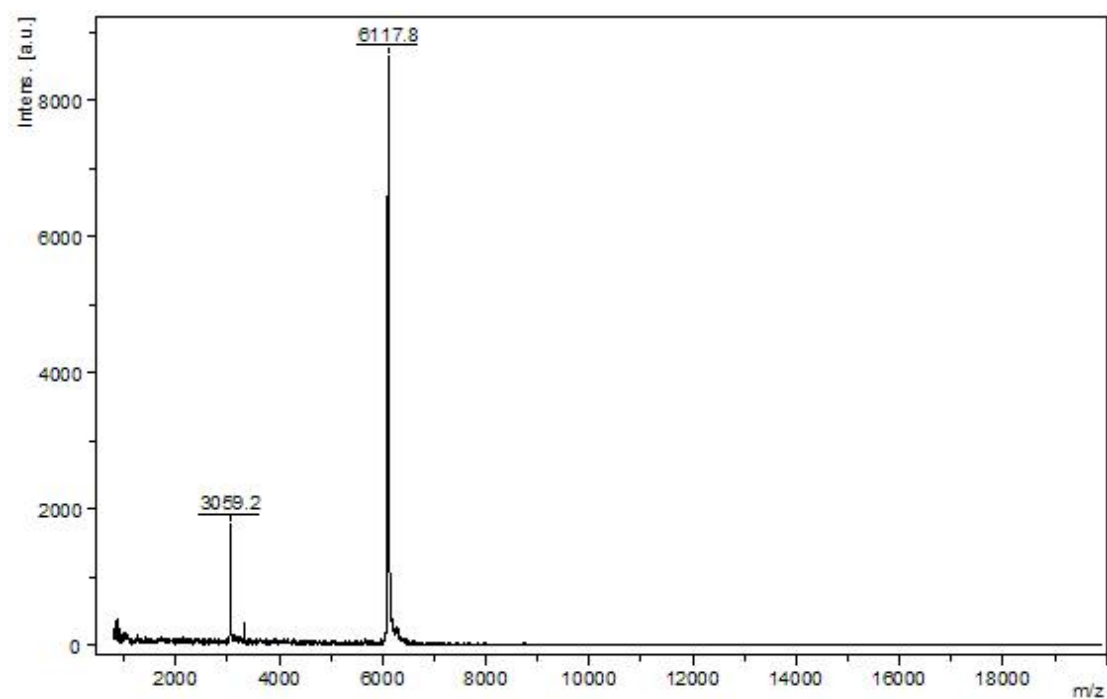


Figure S10. The mass spectrum of 30.425 min peak in Figure 1-D.

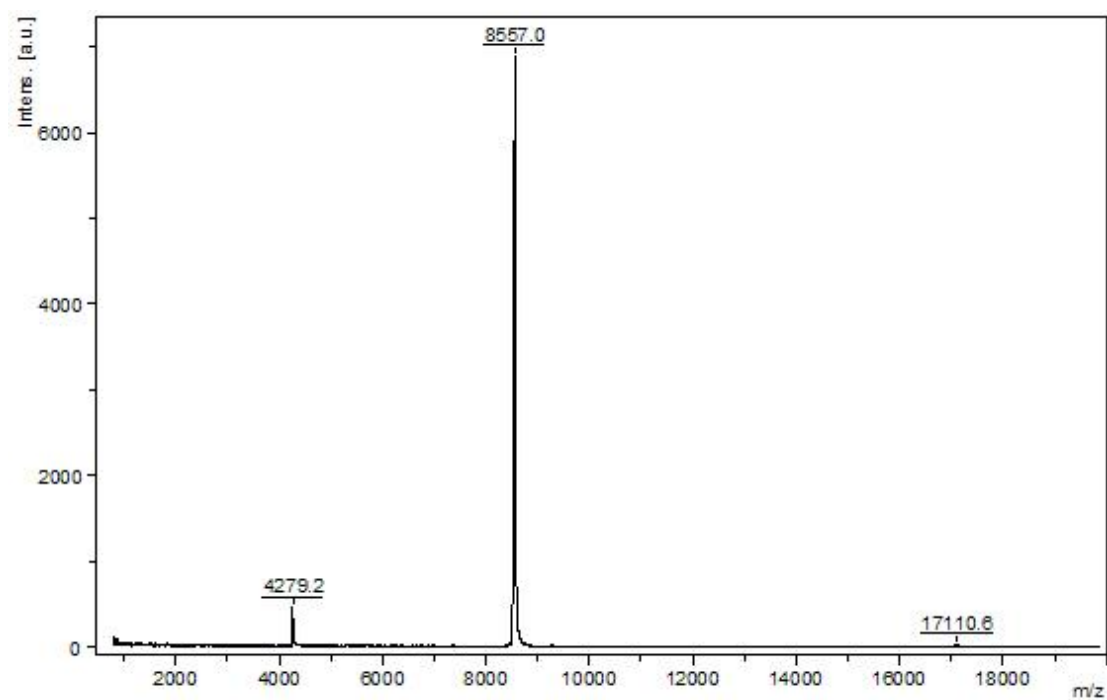


Figure S11. The mass spectrum of 33.756 min peak in Figure 1-D.

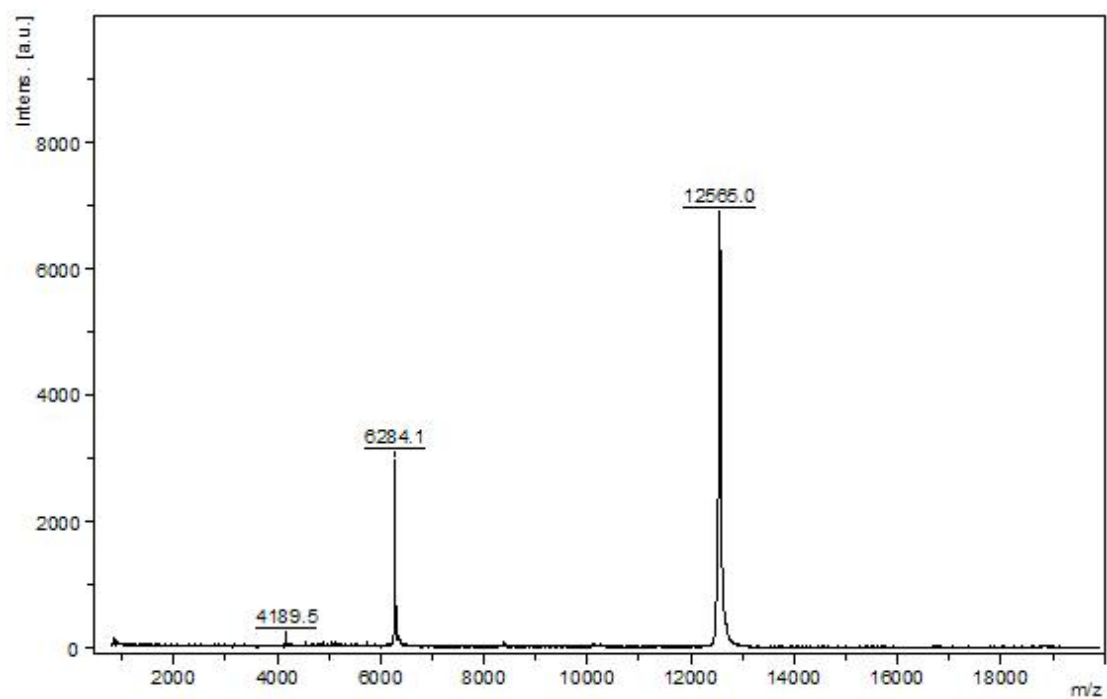


Figure S12. The mass spectrum of 34.874 min peak in Figure 1-D.

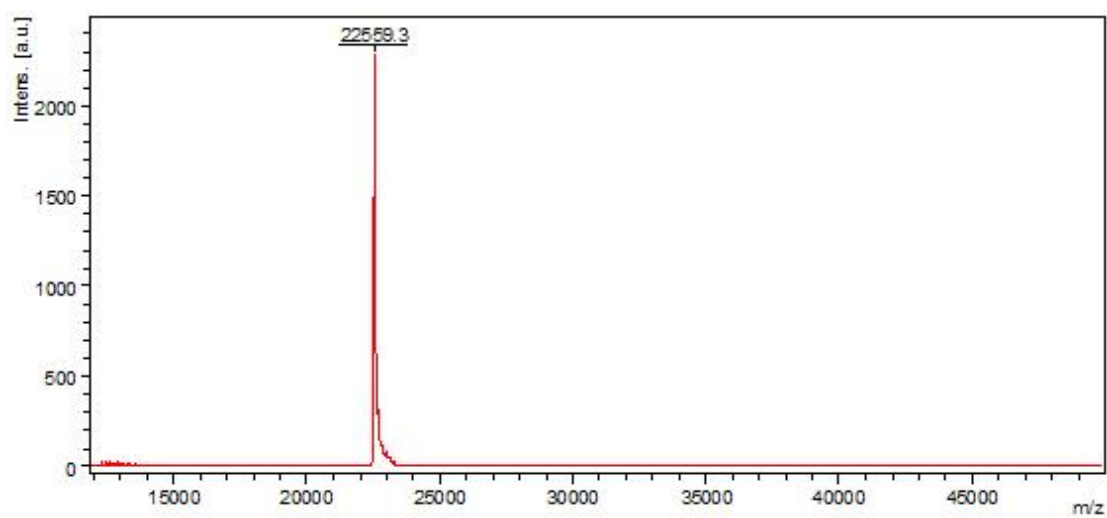


Figure S13. The mass spectrum of 35.314 min peak in Figure 1-D.

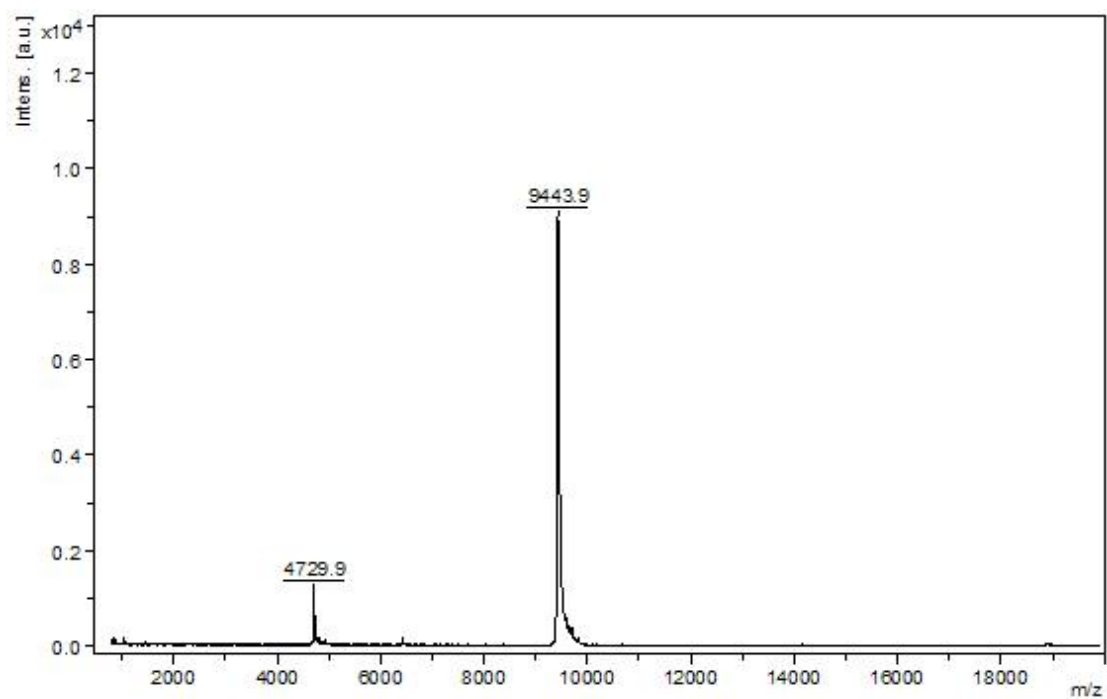


Figure S14. The mass spectrum of 38.022 min peak in Figure 1-D.

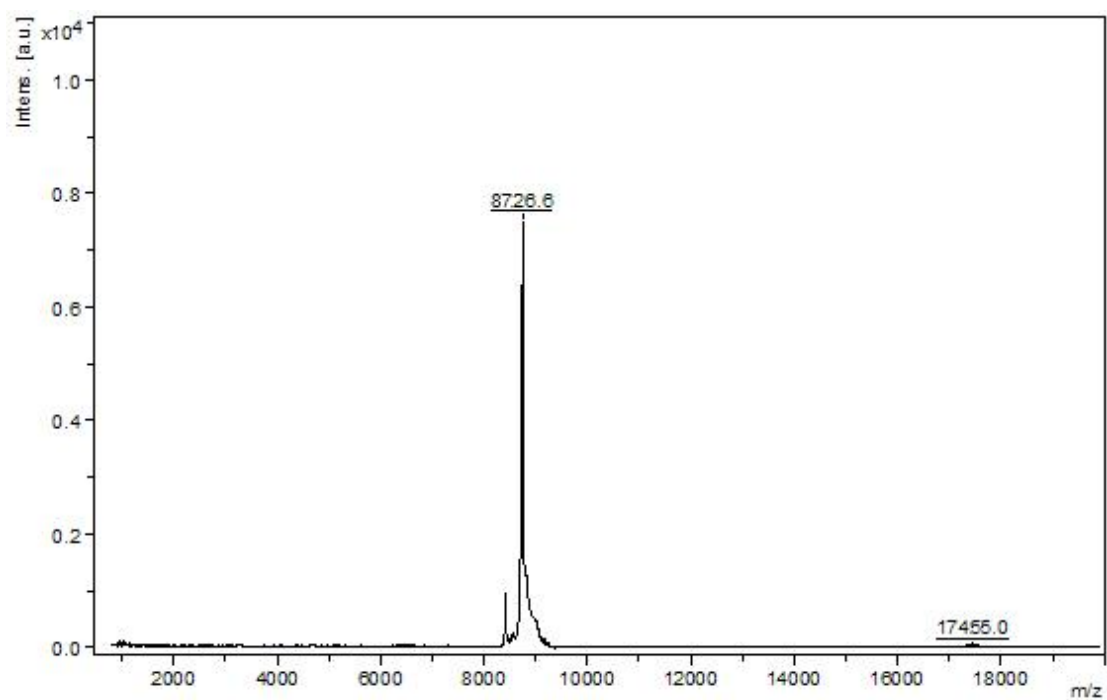


Figure S15. The mass spectrum of 38.801 min peak in Figure 1-D.

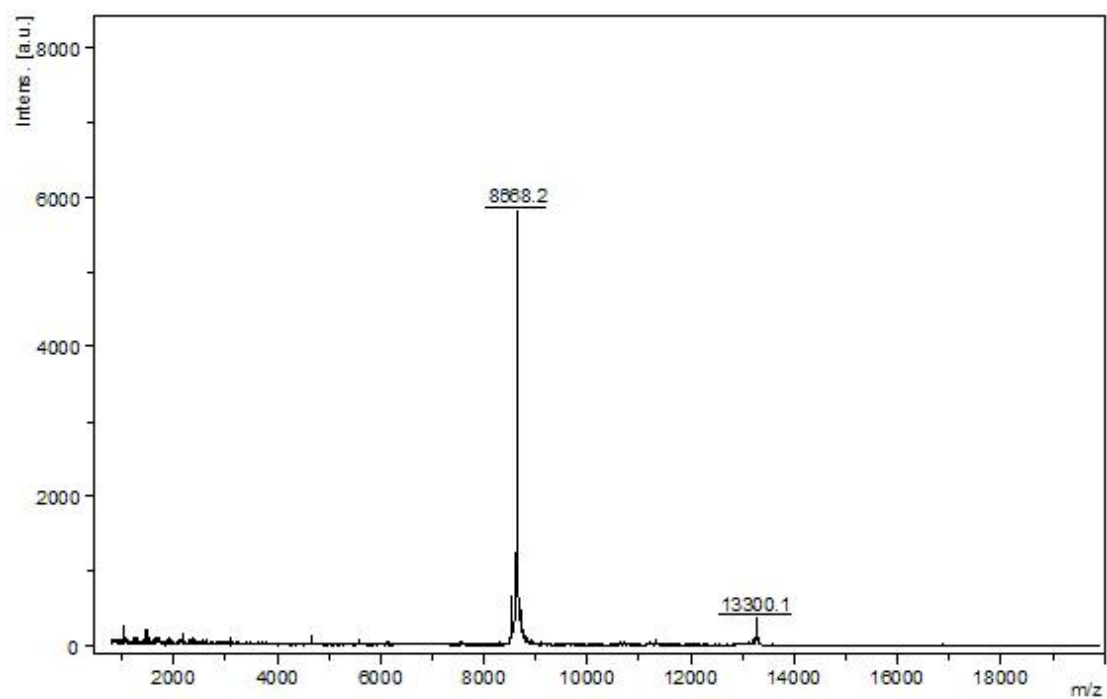


Figure S16. The mass spectrum of 39.954 min peak in Figure 1-D.

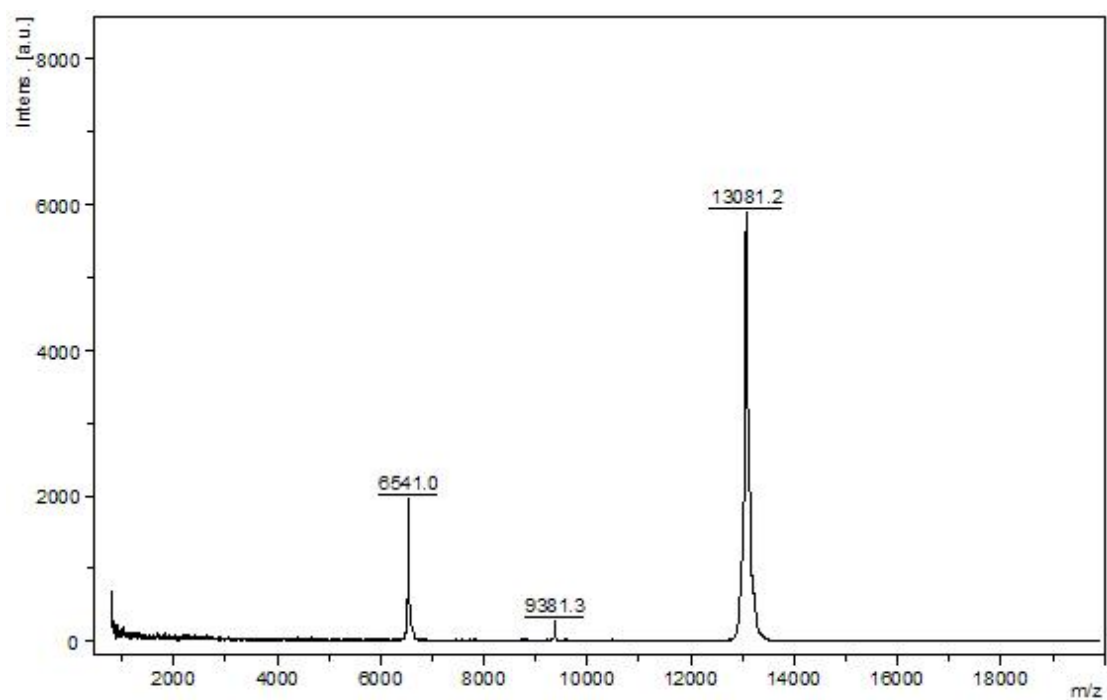


Figure S17. The mass spectrum of 40.133 min peak in Figure 1-D.

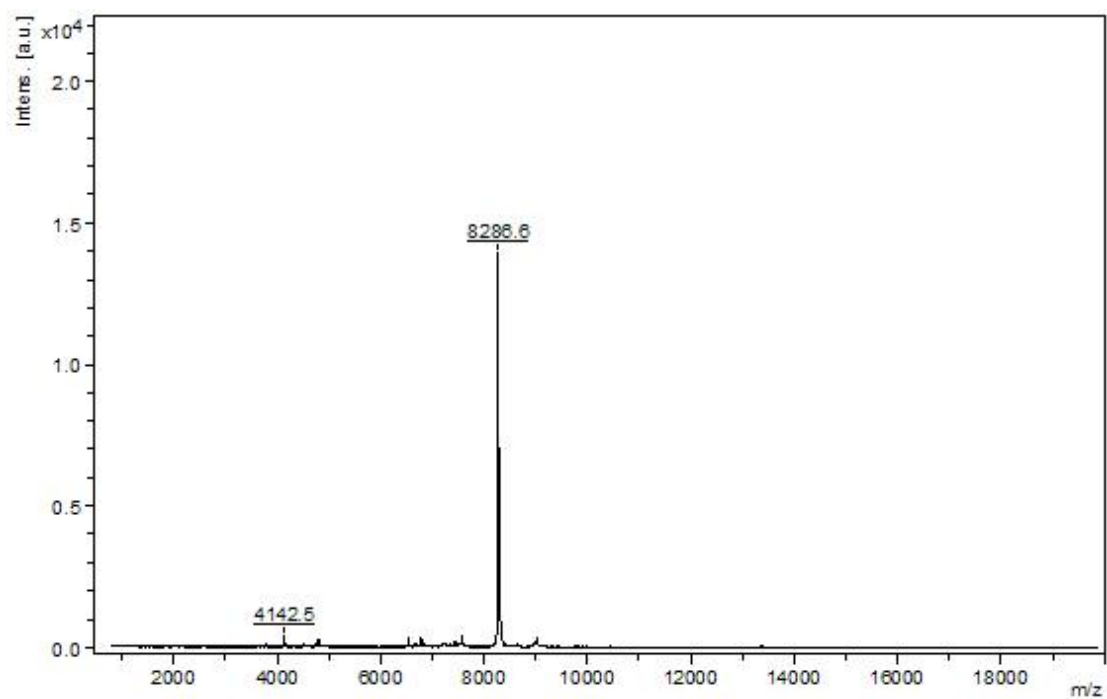


Figure S18. The mass spectrum of 40.691 min peak in Figure 1-D.

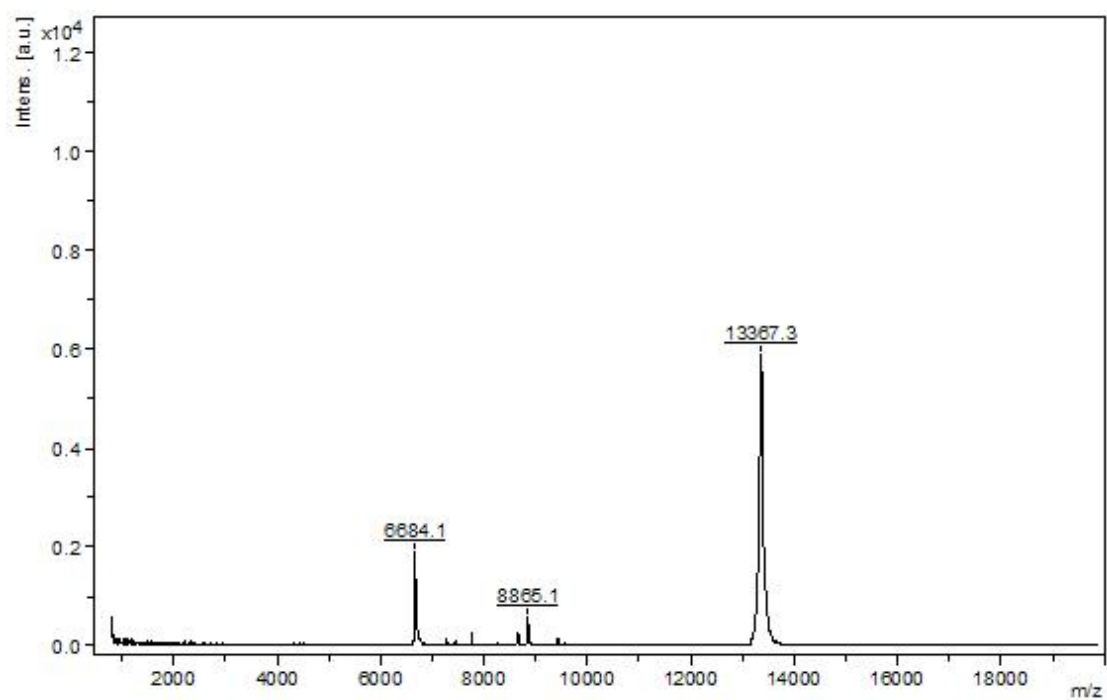


Figure S19. The mass spectrum of 41.546 min peak in Figure 1-D.

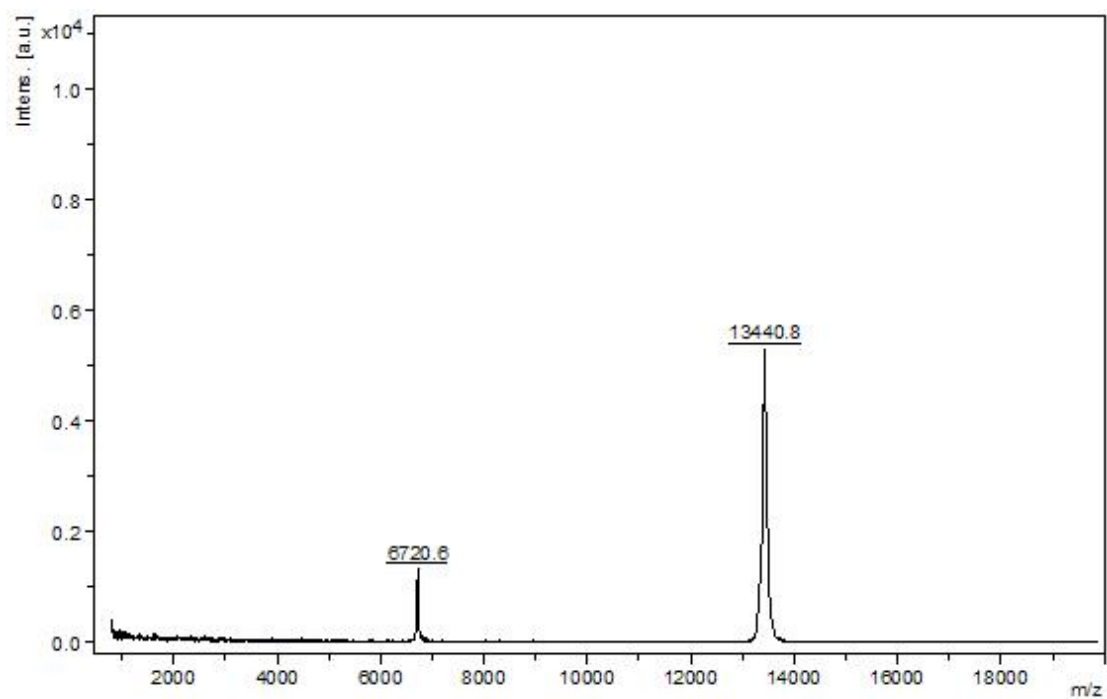


Figure S20. The mass spectrum of 42.176 min peak in Figure 1-D.

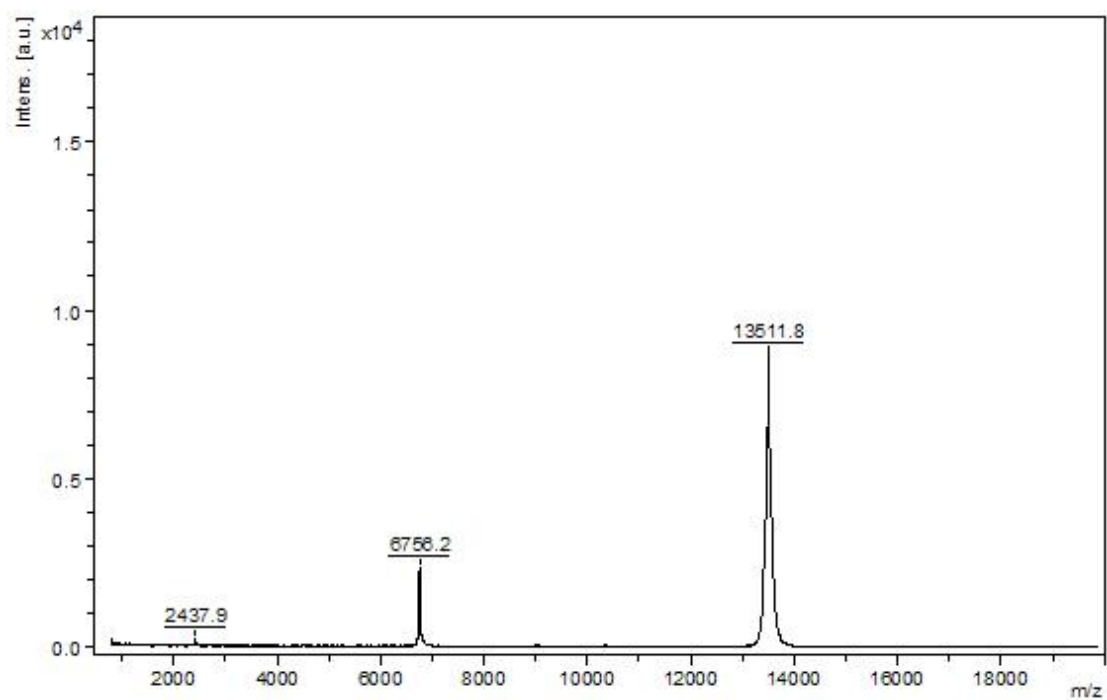


Figure S21. The mass spectrum of 43.689 min peak in Figure 1-D.

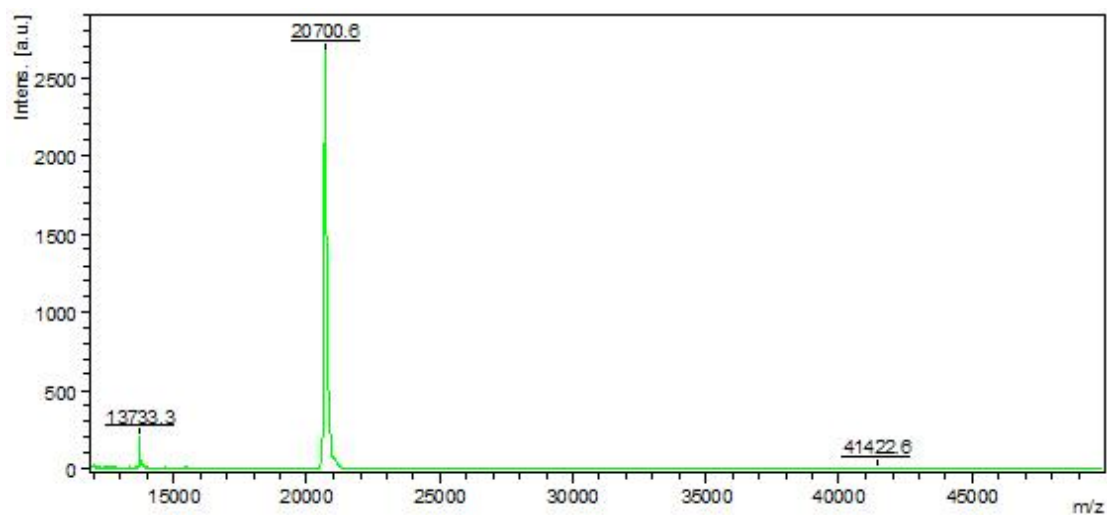


Figure S22. The mass spectrum of 44.441 min peak in Figure 1-D.

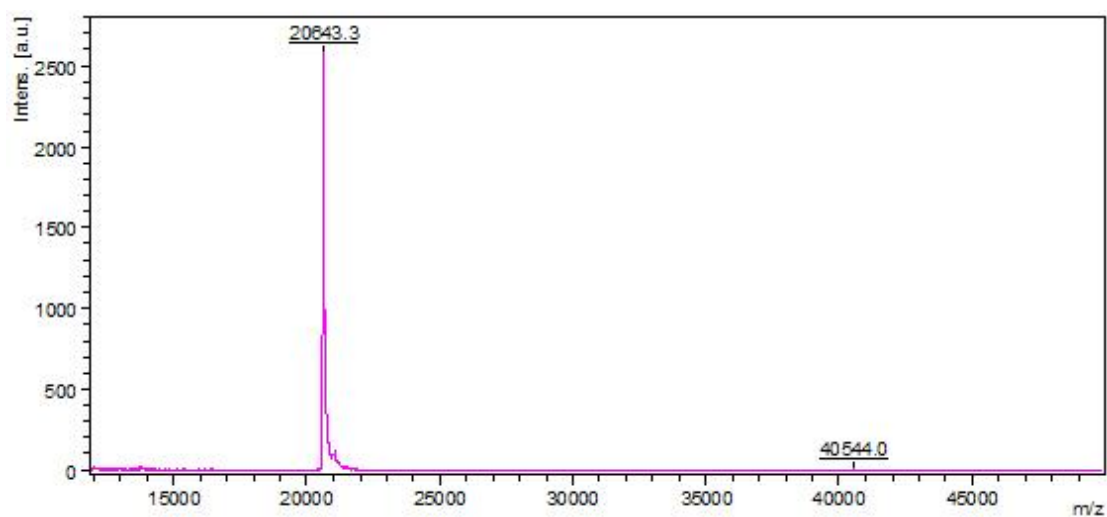


Figure S23. The mass spectrum of 45.725 min peak in Figure 1-D.

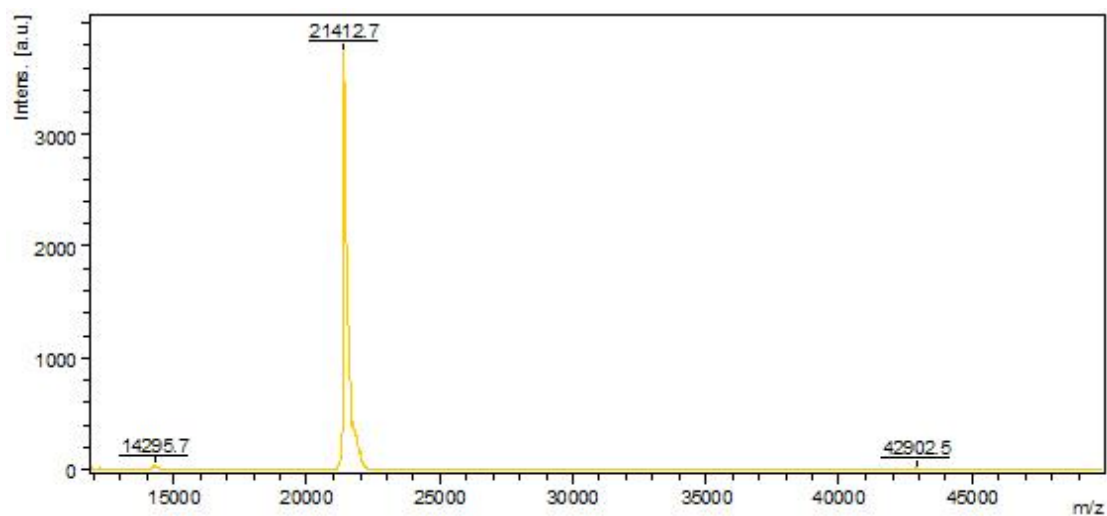


Figure S24. The mass spectrum of 52.197 min peak in Figure 1-D.

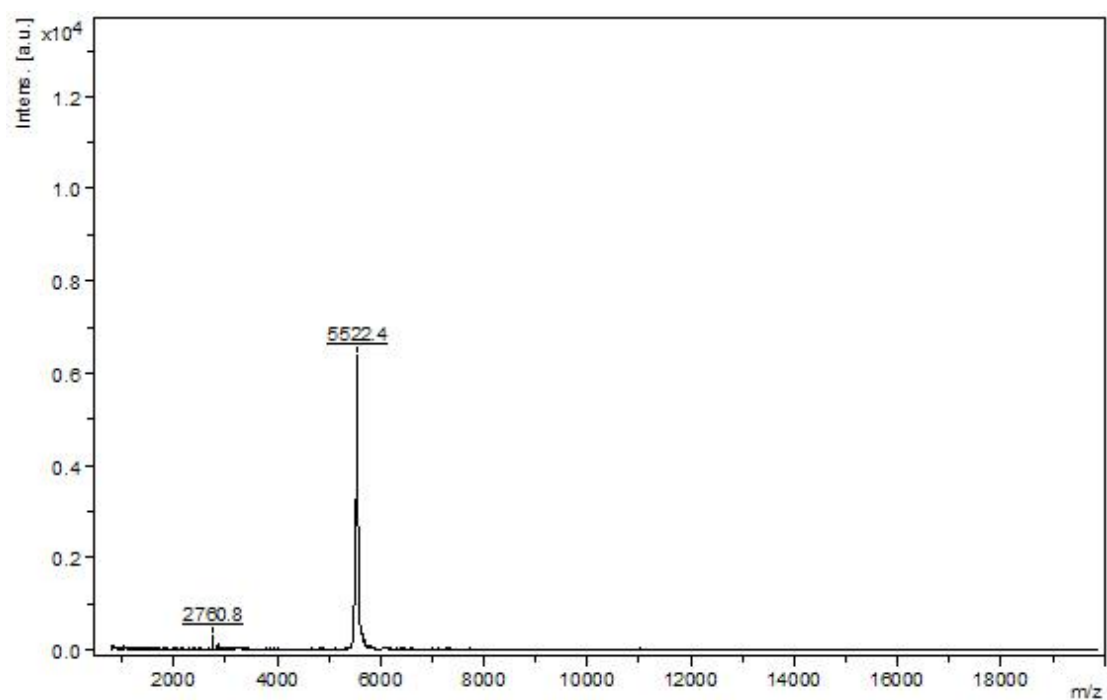


Figure S25. The mass spectrum of 27.956 min peak in Figure 1-F.

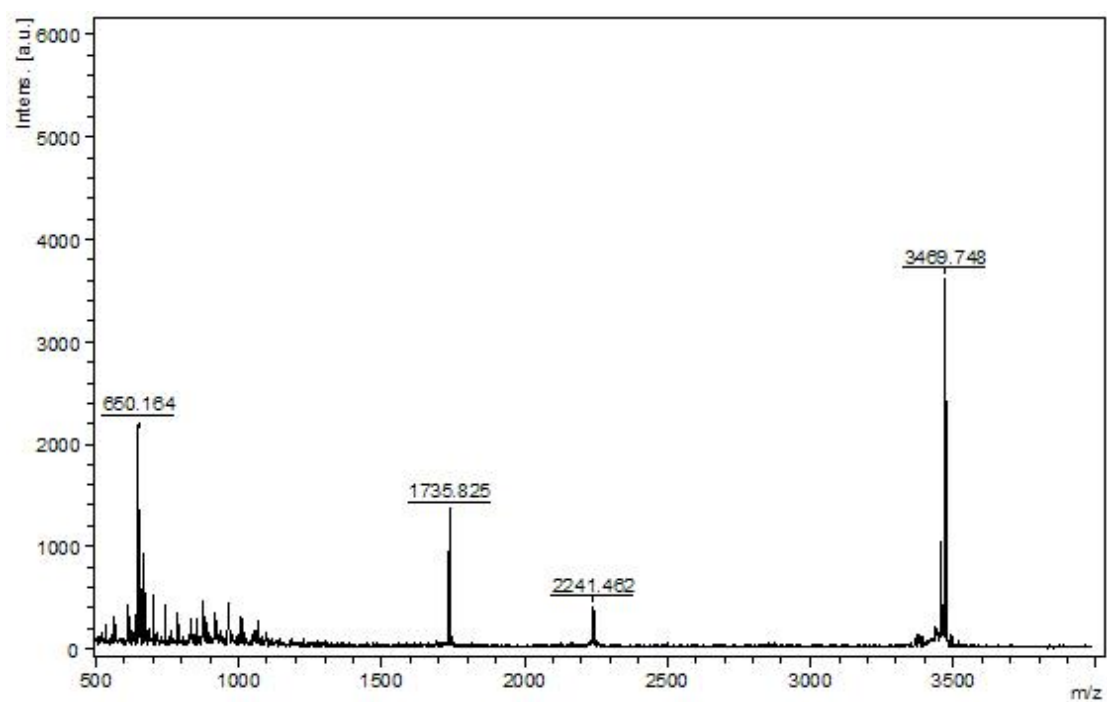


Figure S26. The mass spectrum of 29.749 min peak in Figure 1-H.

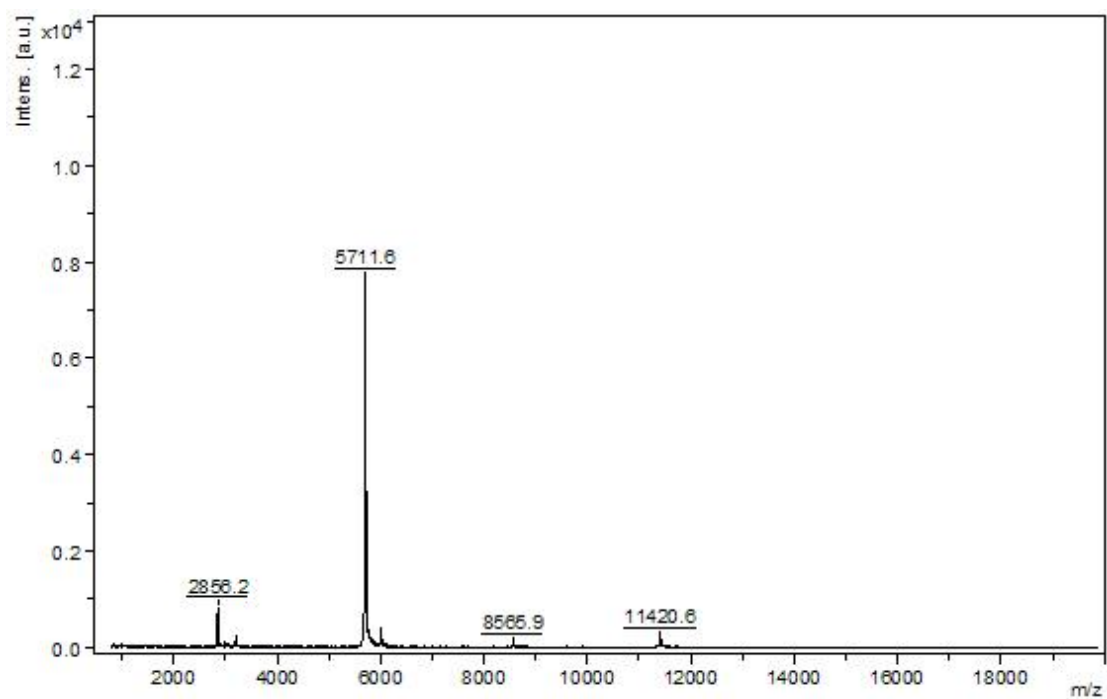


Figure S27. The mass spectrum of 35.652 min peak in Figure 1-H.

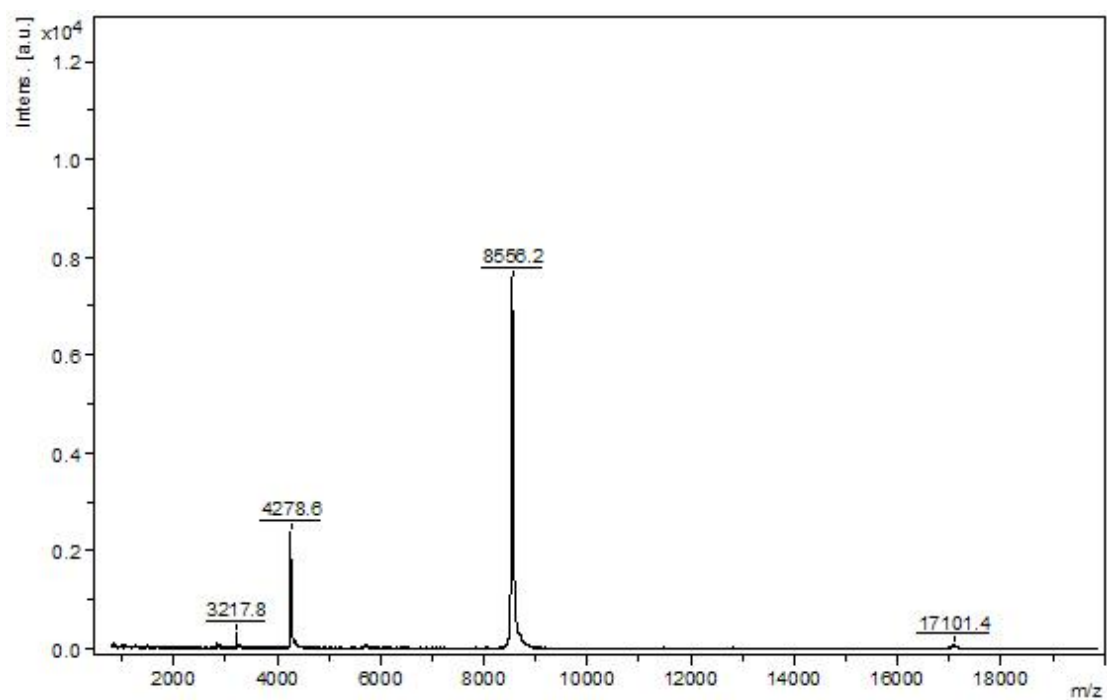


Figure S28. The mass spectrum of 36.169 min peak in Figure 1-H.

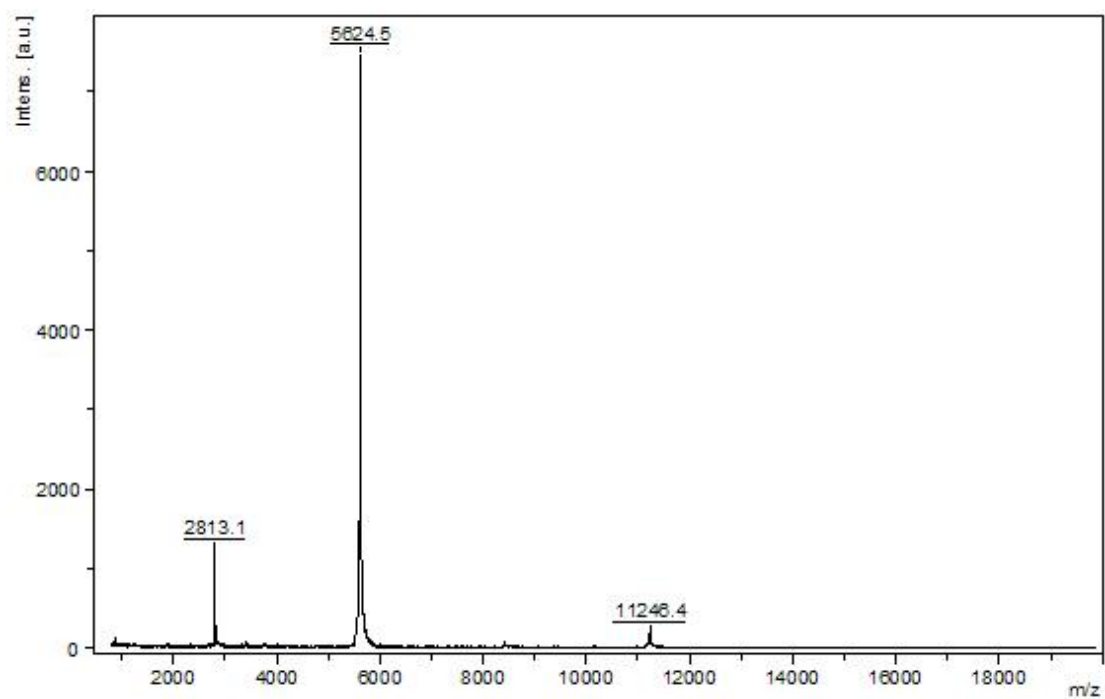


Figure S29. The mass spectrum of 28.725 min peak in Figure 1-I.

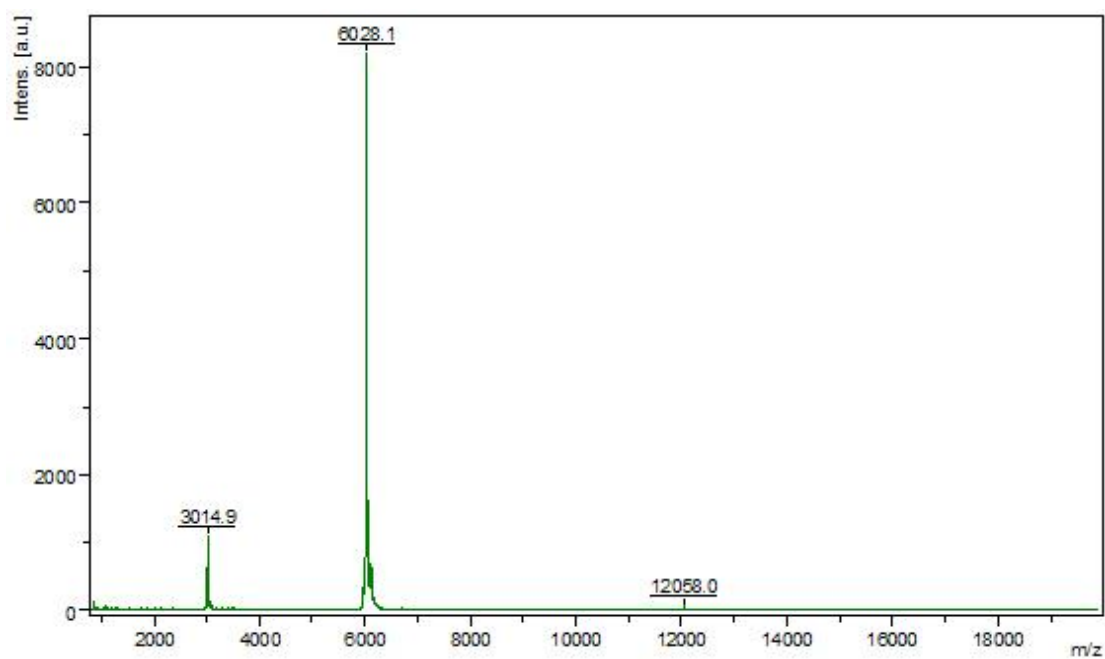


Figure S30. The mass spectrum of 30.966 min peak in Figure 1-I.

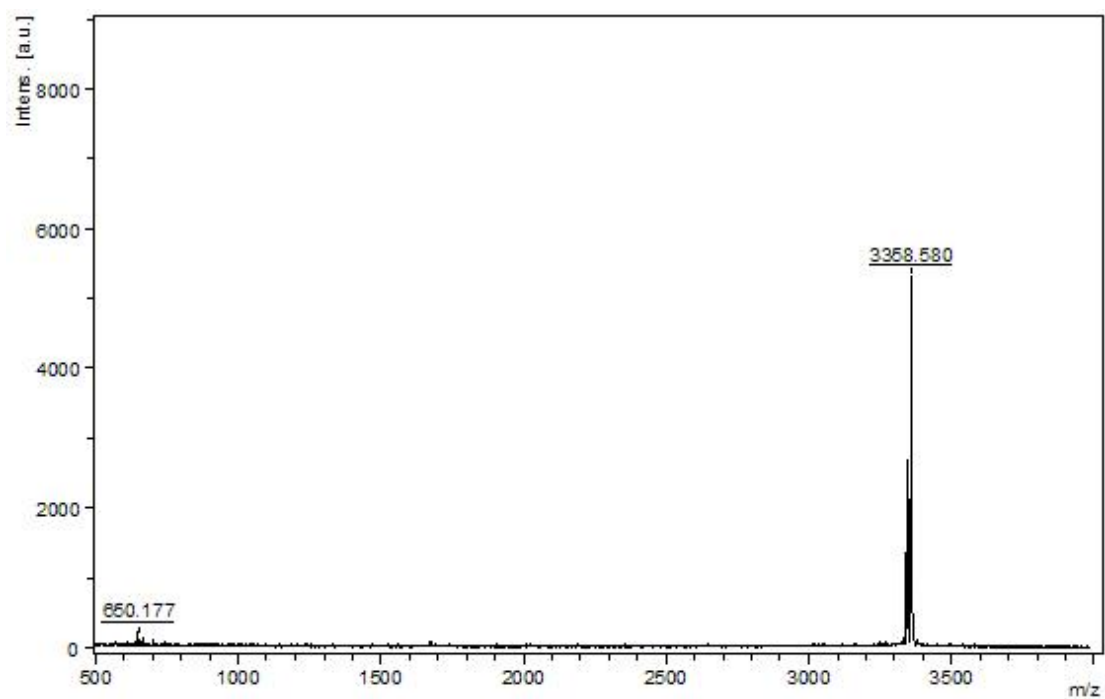


Figure S31. The mass spectrum of 28.544 min peak in Figure 1-J

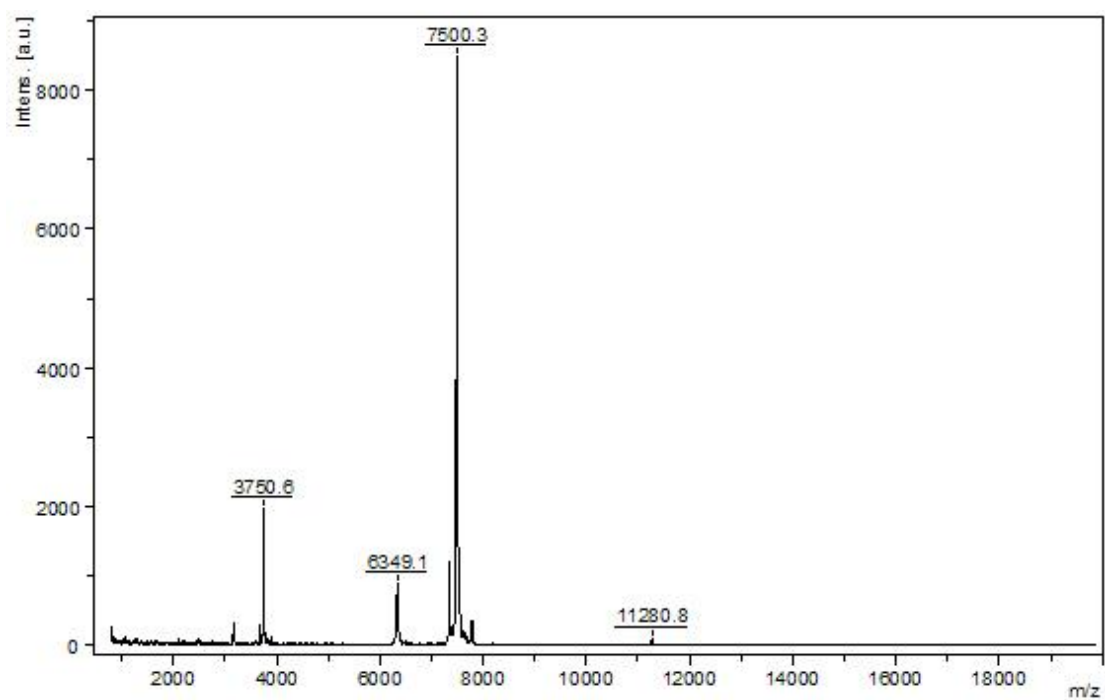


Figure S32. The mass spectrum of 34.392 min peak in Figure 1-J

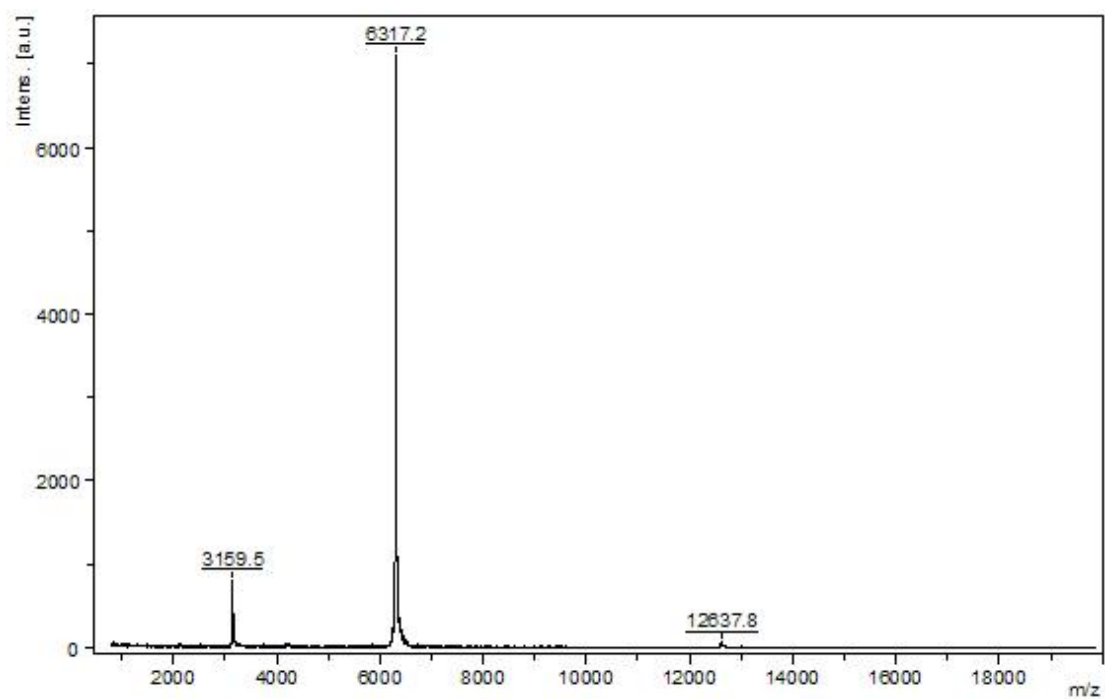


Figure S33. The mass spectrum of 35.094 min peak in Figure 1-J

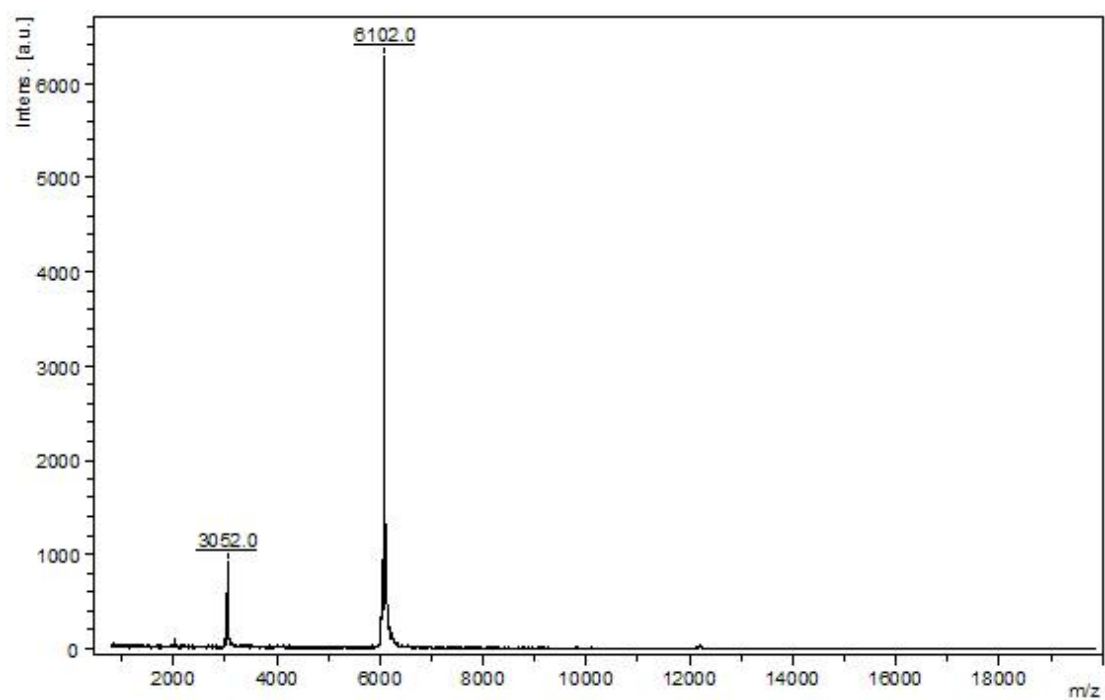


Figure S34. The mass spectrum of 29.208 min peak in Figure 1-K

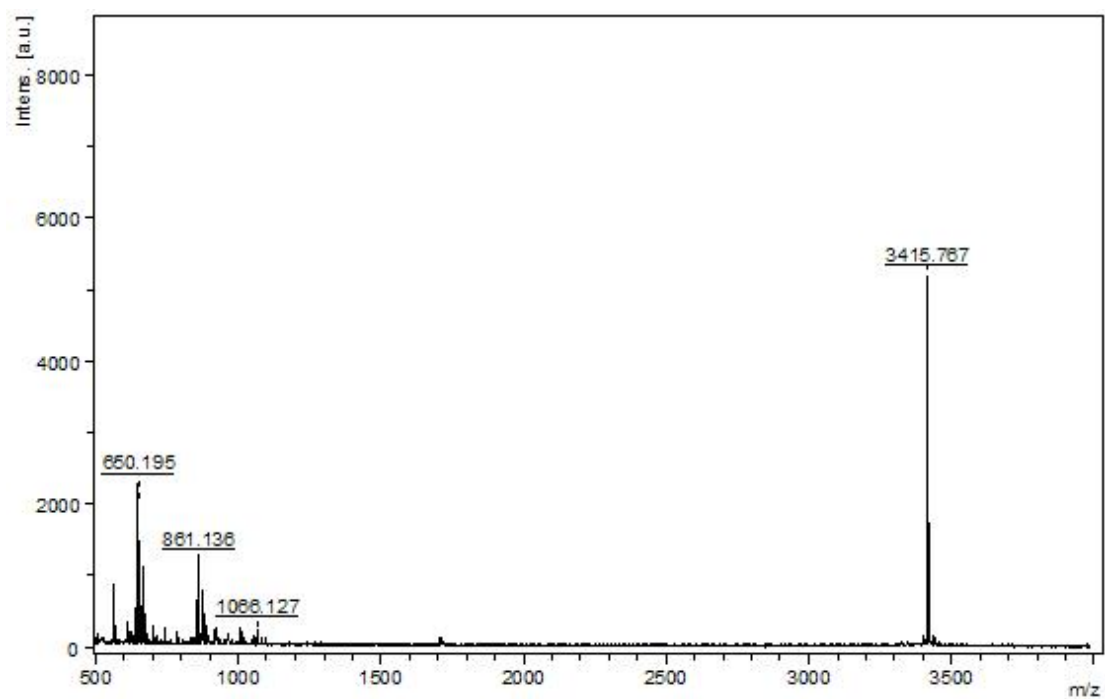


Figure S35. The mass spectrum of 32.976 min peak in Figure 1-K

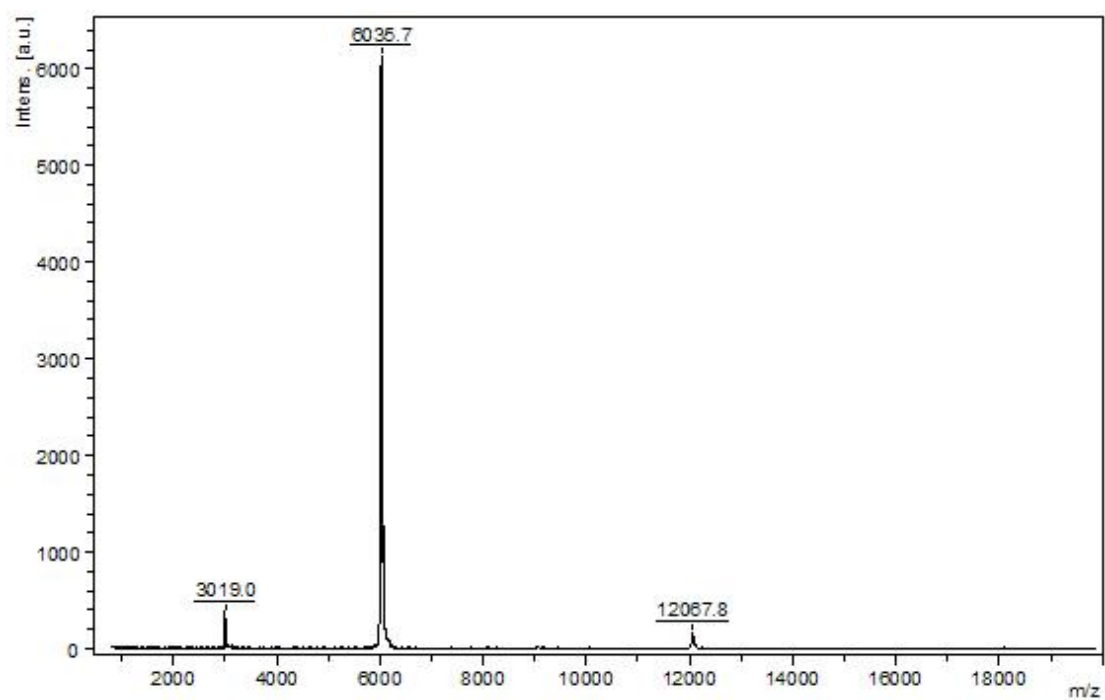


Figure S36. The mass spectrum of 27.672 min peak in Figure 1-L

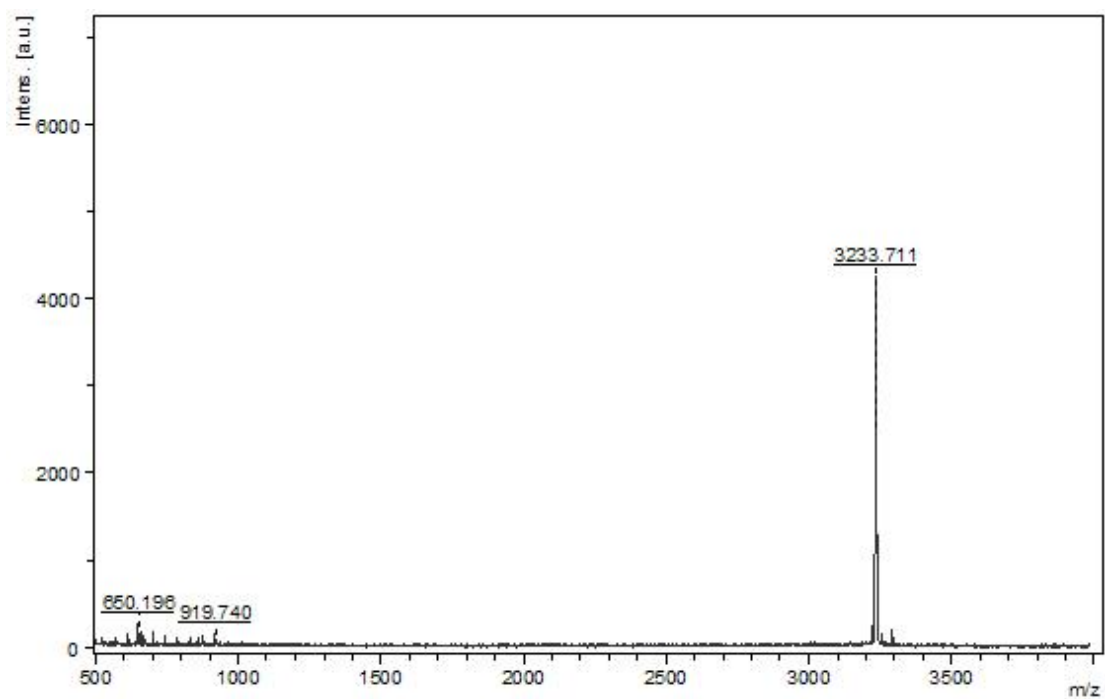


Figure S37. The mass spectrum of 28.944 min peak in Figure 1-L

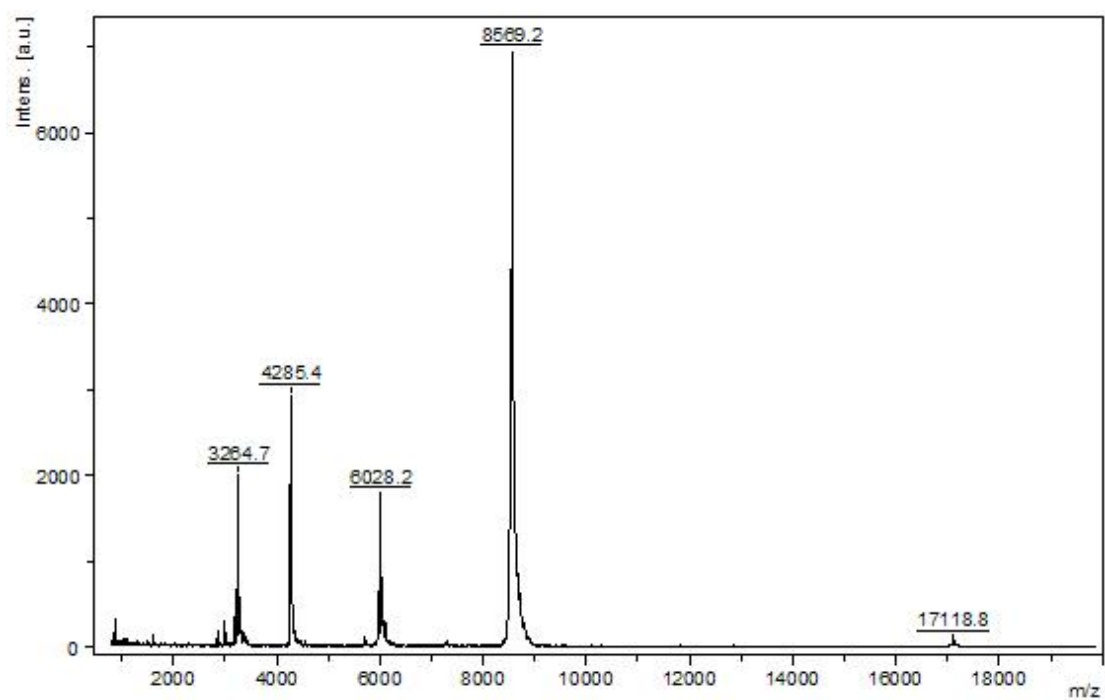


Figure S38. The mass spectrum of 29.458 min peak in Figure 1-L

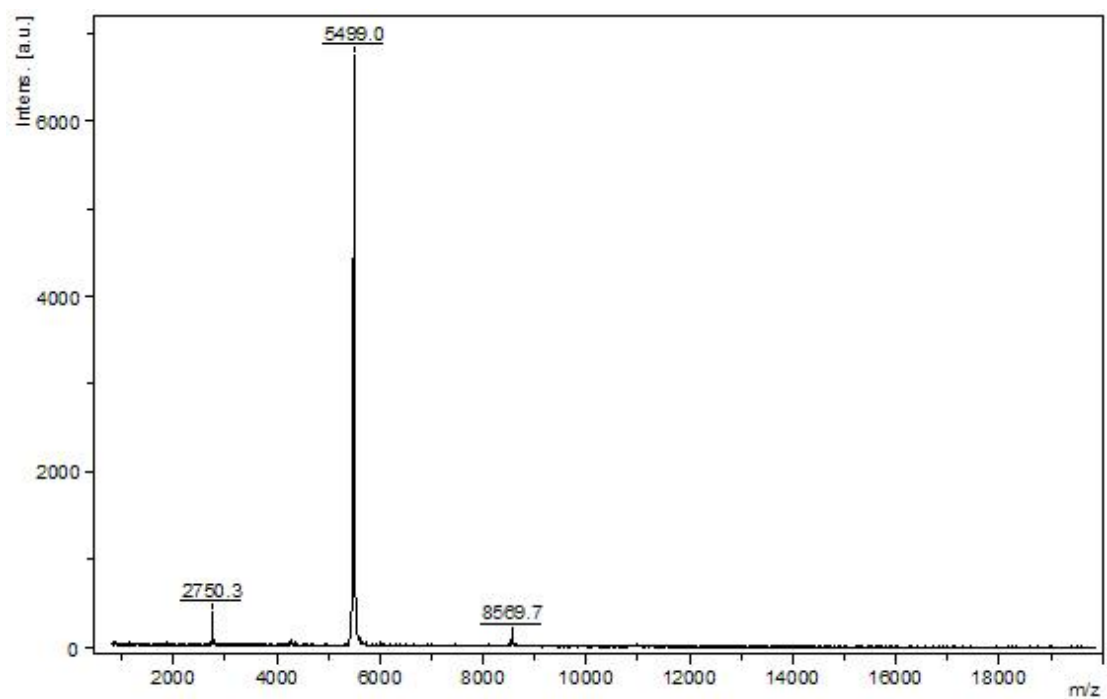
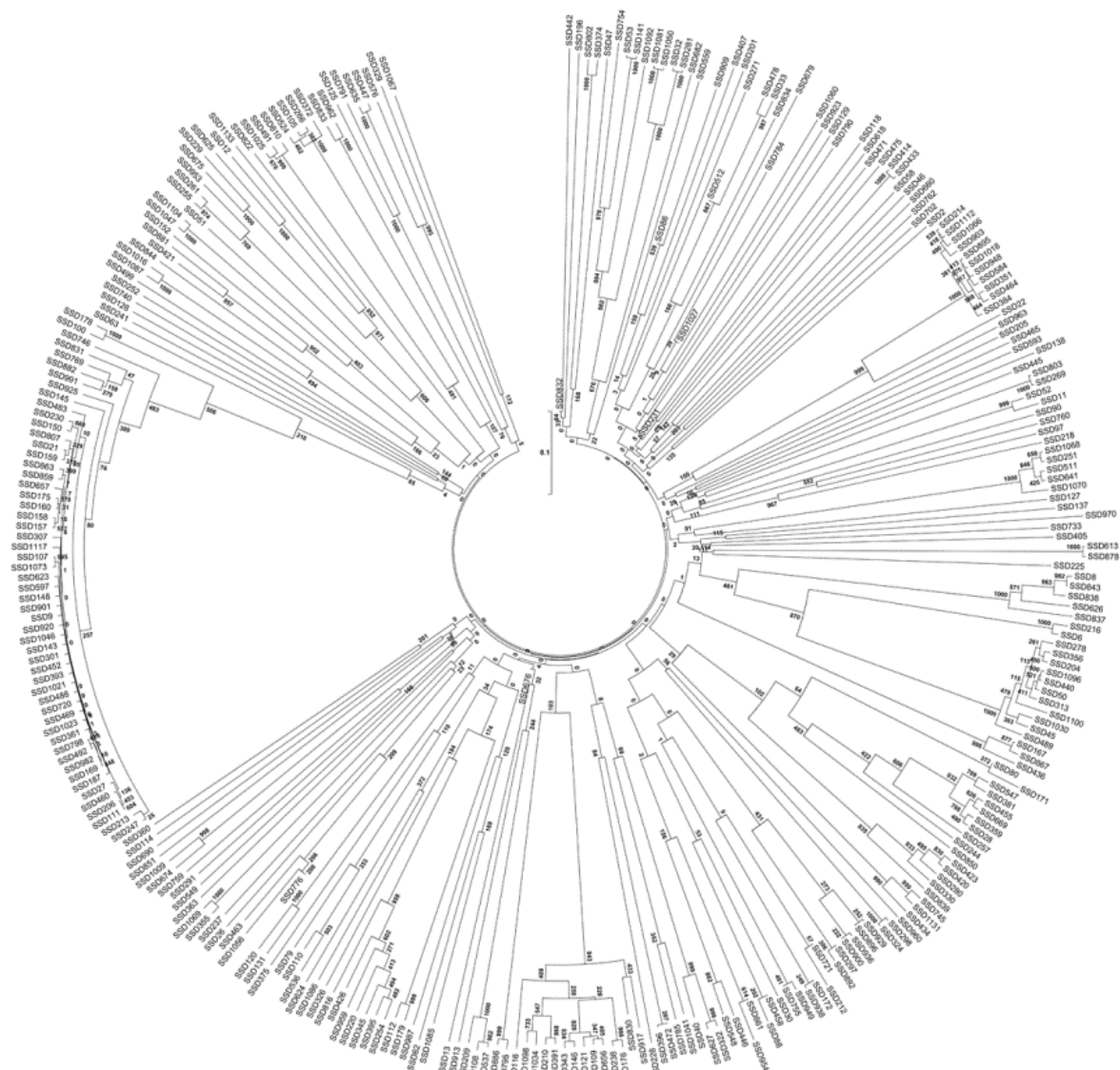


Figure S39. The mass spectrum of 30.681 min peak in Figure 1-L



Group1	Group2	Group3	Group4	Group5	Group6	Group7	Group8	Total
15	45	20	65	51	17	88	7	308

Figure S40. The Cluster of Proteins/peptides with no hits, 308 molecules.

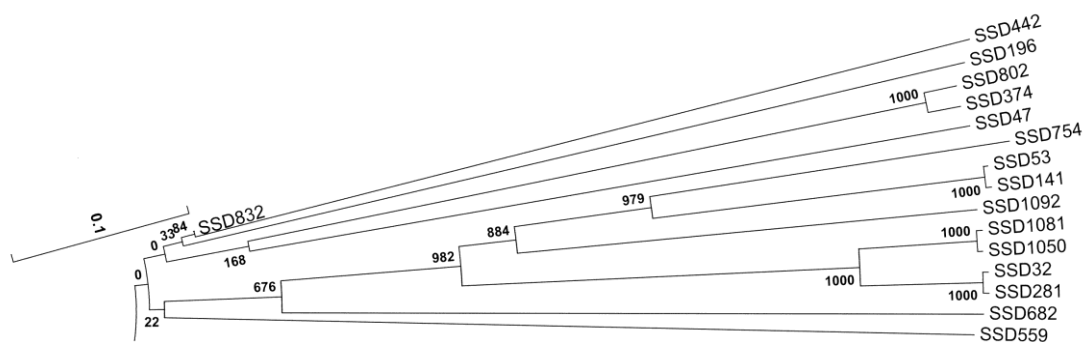


Figure S41. Enlargement of Group 1, 15 members.

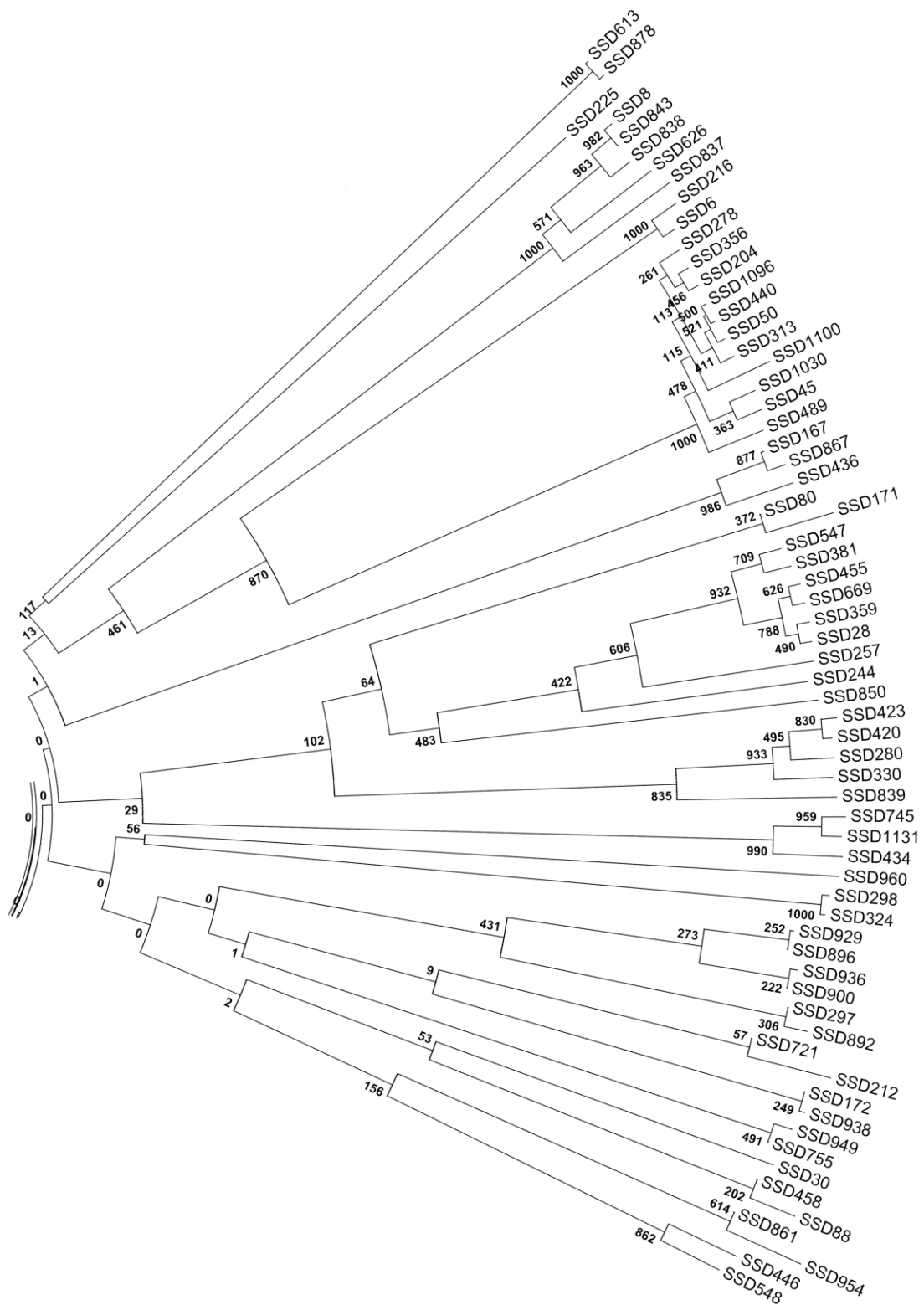


Figure S44. Enlargement of Group 4, 65 members.

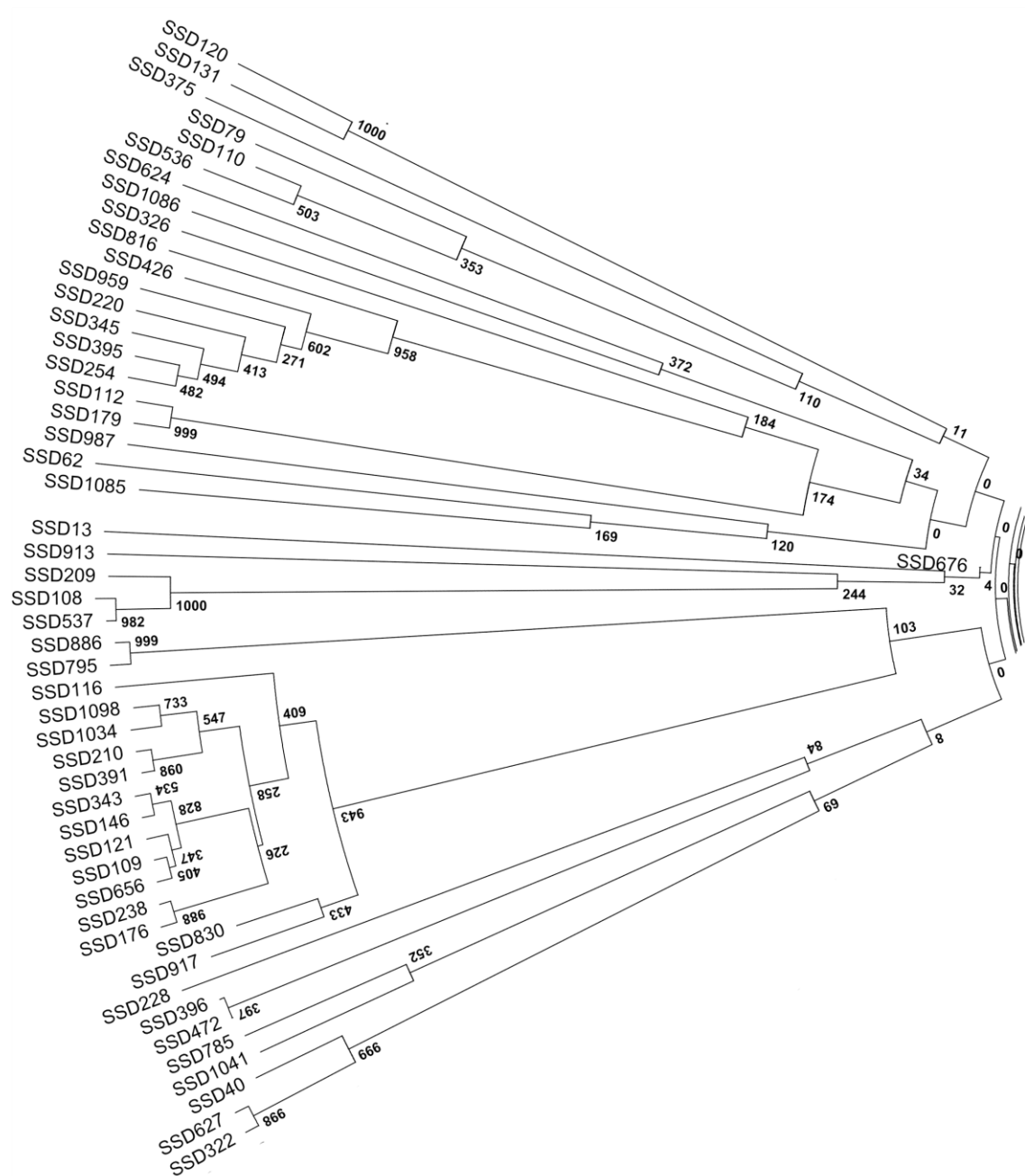


Figure S45. Enlargement of Group 5, 51 members.

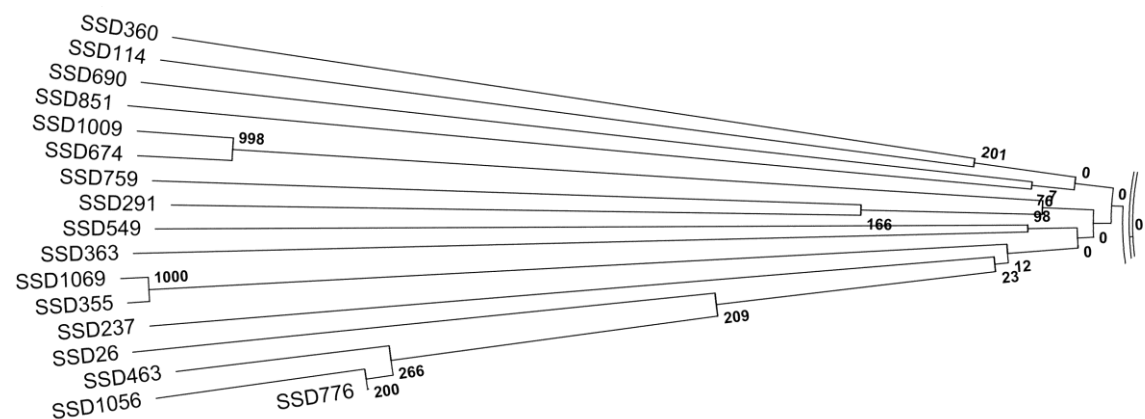


Figure S46. Enlargement of Group 6, 17 members.

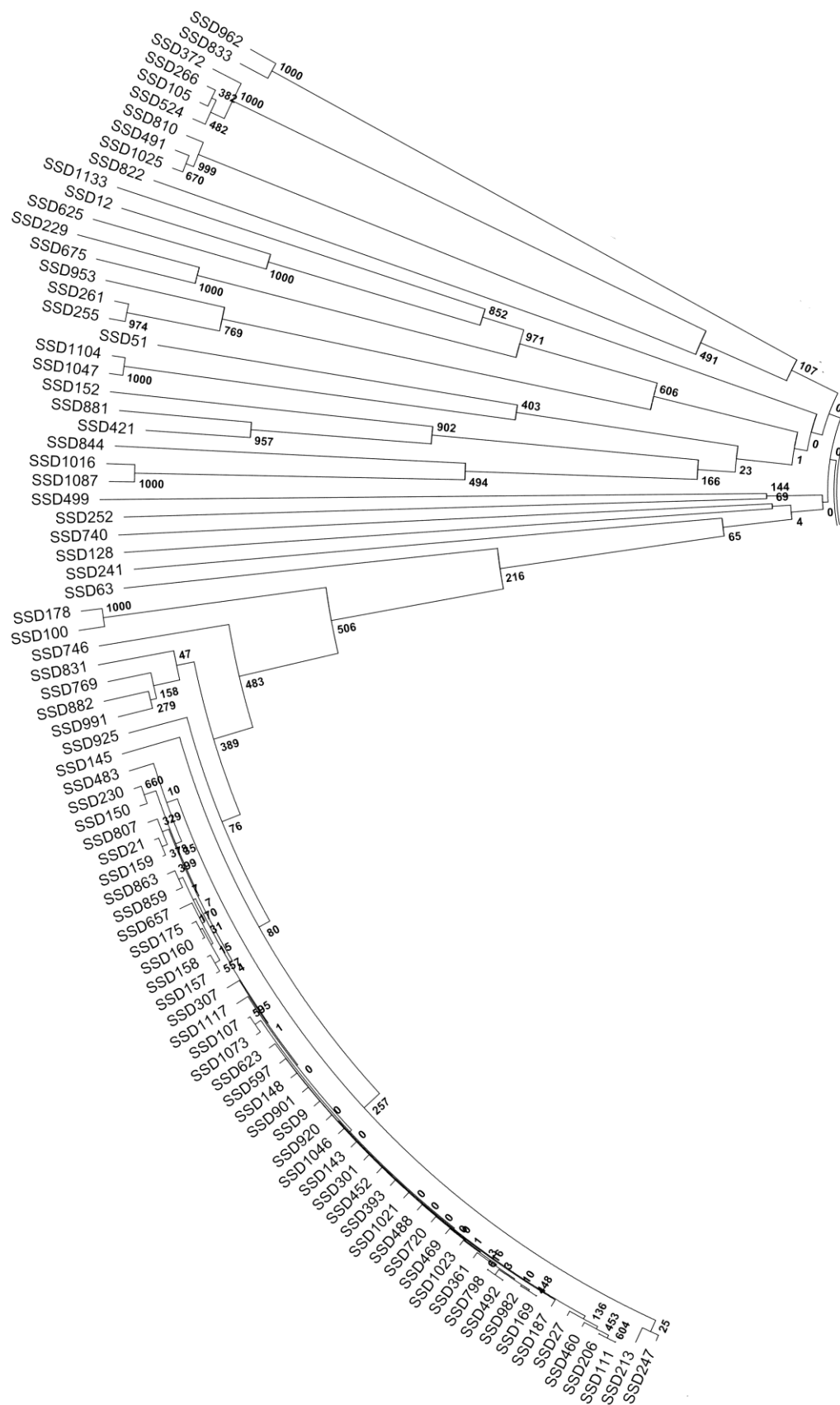


Figure S47. Enlargement of Group 7, 88 members.

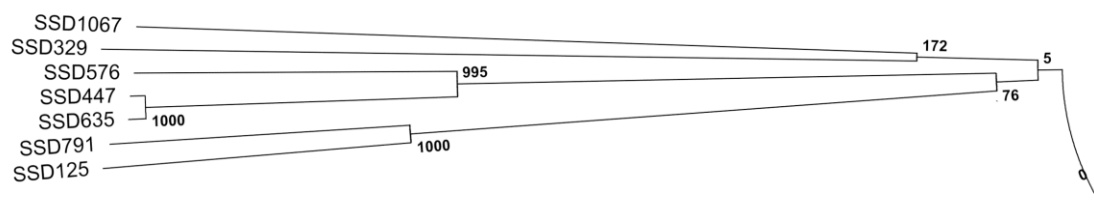


Figure S48. Enlargement of Group 8, 7 members.