The Journal of Animal & Plant Sciences, 23(5): 2013, Page: 1398-1406
ISSN: 1018-7081

A NOVEL THROMBIN-LIKE GENE OBTAINED FROM CHINESE *GLOYDIUS SHEDAOENSIS SHEDAOENSIS*

Chunmei Guo, Shuqing Liu*, Tihua Zheng*, and Ming-Zhong Sun

Department of Biotechnology, *Department of Biochemistry, Dalian Medical University, Dalian 116044, China
Corresponding Author's (Ming-Zhong Sun) E-mail: mxs288@gmail.com

ABSTRACT

The snake of *Gloydius shedaoensis shedaoensis* (GSS) is the particular and sole snake species from Shedao island in Dalian, China. A novel gene named as GSSG-TLE was cloned from GSSG by PCR using GSSG-cDNA library plasmid as template. The GSSG-TLE cDNA was sequenced and encompasses an open reading frame of 786 bp in length and encoding a protein of 262 amino acid residues. The GSSG-TLE gene has a base composition of 14 Ala, 12 Cys, 14 Asp, 13 Glu, 10 Phe, 19 Gly, 7 His, 20 Ile, 15 Lys, 25 Leu, 4 Met, 12 Asn, 19 Pro, 7 Gln, 12 Arg, 20 Ser, 14 Thr, 13 Val, 5 Trp and 7 Tyr. GSSG-TLE includes a signal peptide of 18 amino acids, a proposed propeptide of 6 amino acids and a matured peptide of 238 amino acids. It also contains 12 Cys which form 6 disulfide bridges and has three conserved catalytically active sites of His^{67}, Asp^{112} and Ser^{202}. Sequence comparison revealed that GSSG-TLE amino acids sequence shared highly identity with other snake venom thrombin-like enzymes.

Key words *Gloydius shedaoensis shedaoensis,* Snake, Thrombin-like enzyme.

INTRODUCTION

There are two snake islands in the world, one is in Brasil, the other (Chinese Shedao Island, CSI) situates at Lvshunkou District, Dalian, China. Located 25-miles northwest of the Lvshunkou District in the Sea of Bohai, towering cliffs almost completely encompass CSI making it a true remote geographic wonder (Fig. 1A). CSI is about one square kilometers 217 meters above sea level. Interestingly, *Gloydius shedaoensis shedaoensis* (GSS) snake is the sole species segregated as a unique population in CSI. And about 20, 000 GSSs live in the island (Fig. 1B). The GSS are fed mainly by migratory birds and brown rats living on CSI. The snake venom of GSS (GSSV) is poorly understood (Guo *et al.*, 2013; Liu *et al.*, 2010; Liu *et al.*, 2011; Yang *et al.*, 2009; Yang *et al.*, 2003; Jiang *et al.*, 2009).

Snake venom proteases are mainly classified into serine proteinases and metalloproteinases (Giron *et al.*, 2013; Guo *et al.*, 2012; Lu *et al.*, 2000; Madrigal *et al.*, 2012; Menaldo *et al.*, 2012; Rokyta *et al.*, 2012; Yang *et al.*, 2002). Serine proteinases commonly functionalize in blood-clotting disorders and local tissue destruction by acting on the blood coagulation cascade (Giron *et al.*, 2013; Liu *et al.*, 2006; Menaldo *et al.*, 2012; Park *et al.*, 1998). Belonging to the serine proteinase family, snake venom thrombin-like enzymes (TLEs) have attracted great interests due to their potential therapeutic usage in myocardial infarction and thrombotic diseases. They act on fibrinogen leading to the defibrinogenation of blood with a consequent decrease in blood viscosity, consequently, the consistency and flow properties of blood are greatly improved (Farid *et al.*, 1989; Giron *et al.*, 2013; Madrigal *et al.*, 2012; Menaldo *et al.*, 2012; Rokyta *et al.*, 2012; Smolka *et al.*, 1998; Ouyang *et al.*, 1992; Yang *et al.*, 2002). A variety of TLEs from different snake venoms have been isolated and characterized.

Previously, we have successfully constructed a cDNA library of GSS venomic gland (GSSG) using switching mechanism at 5' end of RNA transcript (SMART) technique (Guo *et al.*, 2013; Liu *et al.*, 2010), and preliminarily characterized the proteome of GSSV by SDS-PAGE and 2D-PAGE combined to HPLC-nESI-MS/MS proteomic approach (Ma *et al.*, 2009; Liu *et al.*, 2011). In current work, we successfully cloned a new TLE gene from GSSG by a new approach. The primers for PCR of GSS- TLE were designed based on the identified peptide sequences derived by proteomics assay. A novel gene cDNA full length sequence of GSSG was obtained by using GSSG-cDNA library plasmid as template. Sequence comparison indicated that the novel gene is highly homogenous with other snake venom TLEs and designated as GSSG-TLE. The corresponding expression construct of GSSG-TLE was also obtained and the protein expression conditions for GSSS-TLE were also explored.

MATERIALS AND METHODS

Materials: GSSG-cDNA library was constructed and preserved by our group (Guo *et al.*, 2013; Liu *et al.*, 2010); Expression vectors pPIC9K and pGEX-6P-1, *E. coli* DH5α and Rosseta host strains were obtained commercially; pMD18-T vector, DNA Marker, Rnase A, isopropyl-β-D-thiogalactopyranoside (IPTG), TaKaRa Ex

Taq, T_4 DNA Ligase, 5-bromo-4-chloro-3-lndolyl-β-D-galactoside (X-gal), *EcoR* I, *Not* I and gel extraction purification kit were from TaKaRa (Dalian, Japan); Tryptone and yeast extract were from OXOID (England); Acrylamide (Acr), N,N'-methylenebisacrylamide (Bis), sodium dodecyl sulfate (SDS) and tetramethylethylenedia -mine (TEMED) were from Sigma (USA); All other chemicals were analytical grade from commercial sources; All primers were synthesized by Sangon Biotech (Shanghai, China).

Primers designing: We obtained 6 peptide sequences of GSSG-TLE by HPLC-nESI-MS/MS analysis (Table 1). The middle-reverse and middle-forward primers were designed according to MS/MS derived TLCAGTQQGG. The M13 forward and M13 reverse primers were designed based on GSSG-cDNA library was constructed using pDNR-LIB vector (Liu *et al*., 2010; Table 2). The 5'-sequence of GSSG-TLE was amplified by PCR using M13 forward primer and middle-reverse primer, the 3'-sequence of GSSG-TLE was obtained using M13 reverse primer and middle-forward primer. The full-length nucleotide sequence of GSSG-TLE was obtained by PCR using TLE forward and reverse primers employing GSSG-cDNA library plasmid as template.

PCR amplification: The PCR reagent with a total volume of 25 μL was composed by 2.5 μL10 × PCR buffer (Mg^{2+} plus), 2 μL dNTPs mixture (2.5 mM for each of the four dNTPs), 1 μL each primer, 1 μL GSSG-cDNA library plasmid template, 0.125 μL *Taq* DNA polymerase (5 U/μL) (Takara, Japan) and 18.5 μL ddH_2O. Reactions were performed at 95 ^oC for 2 min, followed by 30 cycles of 95 ^oC for 30 s, 55 ^oC or 60 ^oC for 30 s and 72 ^oC for 1 min, at last extended at 72 ^oC for 5 min. The amplified product was visualized by 1% agarose gel electrophoresis containing 0.01% ethidium bromide (EB).

Gene cloning: The PCR amplified product of GSSG-TLE was purified by miniBEST agarose gel DNA extraction kit (Takara, Japan). Then, the purified PCR product was ligated to pMD18-T vector (an T cloning vector) by T_4 DNA ligase at 16 ^oC overnight. 10 μL ligation mixture was transformed into 100 μL competent *E. coli* DH5α cell by heating at 42 ^oC for 90 s, the cells were incubated in 1 mL LB (Lysogeny broth) medium with shaking (225 rpm) at 37 ^oC for 1 h. Then, the transformed cells were spread onto LB agar plates containing 100 μg/mL ampicillin, 33 μL 24 mg/ml IPTG and 40 μL 20 mg/ml X- gal and incubated at 37 ^oC for 16 h. The white colonies were selected by blue-white screening and cultured in LB medium supplemented with 100 μg/mL AMP with shaking at 37 ^oC overnight. The positive clones were designated as recombinant plasmid pMD18-T-GSSG-TLE. The recombinant plasmid DNA was extracted by alkaline-lysis and confirmed by restriction enzyme digestion (*EcoR* I, *Not* I) and PCR assay. The correct recombinant pMD18-T-GSSG-TLE plasmids were sent to Dalian TaKaRa Biotech for sequencing.

Sequence analysis: The derived nucleotide sequence and deduced amino acid sequence was analyzed using BLAST network of NCBI. A homology search in Genbank was performed using NCBI BLAST and alignment of amino acid sequences was generated by CLUSTAL W program.

Site-directed mutagenesis for changing prokaryotes preferred codons: As we found constructed expression vector of GSSG-TLE could not be over-expressed in *E.coli*, we then mutated the prokaryotes preferred codons for benefiting protein expression of GSSG-TLE gene (an eukaryotic gene). According to *E.coli* preferred codons of Mr. Gene, the 3, 4, 6, 7, 9, 10, 12, 13, 170, 171 amino acid residues and stop codons of GSSG-TLE gene were mutated by overlap extension PCR for site-directed mutagenesis. Site-directed mutagenesis forward primer 1 and reverse primer 1, forward primer 2 and reverse primer 2 were designed targeting sequence to be mutated (Table 2). There were 27 overlap base sequences between site-directed mutagenesis reverse primer 1 and forward primer 2. The two fragments obtained using site-directed mutagenesis forward primer 1 and reverse primer 1, forward primer 2 and reverse primer 2, were further purified by gel extraction kit and combined together. Overlap extension PCR was then performed utilizing the two combined purified fragments as mutual primers and templates to amplify the mutated-GSSG-TLE gene.

Construction of GSSG-TLE expression vector: The correct recombinant pMD18-T-GSSG-TLE plasmid was digested with *EcoR* I and *Not* I and the insert was subcloned into pGEX-6P-1 expression vector. The ligation mixtures were transformed into competent *E. coli* DH5α by heating at 42 ^oC for 90 s, and incubated in 1 mL LB medium with shaking (225 rpm) at 37 ^oC for 1 h. The transformed cells were then spread onto LB agar plate containing 100 μg/mL AMP (ampicillin). The subclones were cultured in LB medium containing 100 μg/mL AMP overnight with shaking at 37 ^oC. The positive clones were designated as recombinant pGEX-6P-1-GSSG-TLE plasmid. The plasmid DNA was extracted by alkaline-lysis method and confirmed by restriction enzyme digestion (*EcoR* I, *Not* I) and PCR assay. The positive recombinant pGEX-6P-1-GSSG-TLE plasmids were transformed into competent *E. coli* Rosseta.

Expression of recombinant GSSG-TLE protein:. We optimized induction conditions for fusion GSSG-TLE protein expression in *E. coli* Rosseta. Once the absorbance at 600 nm of cell culture reached ~0.6, different concentrations of IPTG (0.1-1 mM) were added into the culture to induce protein expression at 25, 30 or

37 ^{o}C for 2-6 h. Cell pellets harvested were suspended and sonicated in 20 mL ice-cold lysis buffer (50 mM Tris -HCl, pH 8.8, 2 mM EDTA, 100 mM NaCl, 0.5% Triton X-100, 1 mg/mL lysozyme and 1 mM PMSF. The precipitate was collected by centrifugation at 12,000 rpm for 30 min. The pellet of inclusion bodies was resuspended in 2 × SDS loading buffer containing 1.0 mM Tris-HCl (pH 6.8), 20% glycerol, 10% SDS, 5% β-mercaptoethanol and 0.5% bromophenol blue, heated at 95 ^{o}C for 10 min and analyzed by SDS-PAGE (10%).

RESULTS AND DISCUSSION

GSSG-TLE gene was obtained successfully: 5'- and 3'-sequences of GSSG-TLE were obtained by PCR using M13 forward and middle-reverse primers, M13 reverse and middle-forward primers, respectively (Fig. 2A). Containing partial sequence of pDNR-LIB vector, the sizes of both GSSG- TLE PCR fragments were ~ 1000 bp (Fig. 2A). A PCR product band with the molecular size of approximate 780 bp was shown on 1% agarose gel electrophoresis (Fig. 2B), which is similar with those molecular sizes reported for other snake venom TLEs, indicating that we should already obtain the full-length cDNA sequence of GSSG-TLE.

pMD18-T-GSSG-TLE cloning vector was successfully constructed: The blue-white screening is a method for the detection of successful ligations in vector-based gene cloning. The PCR product of GSSG-TLE was inserted into pMD18-T plasmids and the positive recombinant clones were then selected by blue-white screening (Fig. 3A). The white positive recombinant pMD18-T-TLE were verified by restriction enzyme digestion analysis, PCR assay and sequencing analysis. A fragment of ~ 3500 bp was shown on agarose electrophoresis for recombinant pMD18-T-TLE digested with *EcoR* I and *Not* I (Fig. 3B). As pMD18-T is 2692 bp in length and PCR product of GSSG-TLE is ~780 bp, which strongly indicated the success of pMD18-T-TLE construction.

GSSG-TLE sequence and sequential characteristics: GSSG-TLE gene was composed by 786 nucleotides encoding 262 amino acids (Fig. 4). It contains 14 Ala, 12 Cys, 14 Asp, 13 Glu, 10 Phe, 19 Gly, 7 His, 20 Ile, 15 Lys, 25 Leu, 4 Met, 12 Asn, 19 Pro, 7 Gln, 12 Arg, 20 Ser, 14 Thr, 13 Val, 5 Trp and 7 Tyr. Conversed with *Trimeresurus albolabris* TLE protein (Lin *et al.,* 2009), GSSG-TLE consists of a 18-residues signal region (1-18 residue), a 6 residues propeptide region (19-24 residue) and a 238 residues mature peptide region. Most of known snake venom TLEs are single chain Cys-rich proteins containing 12 Cys residues by forming 6 disulfide bonds functionally required for their biological activities and structural stabilities. GSSG-TLE contains 12 Cys residues that may form 6 disulfide bonds at $Cys^{31\text{-}165}$, $Cys^{52\text{-}68}$, $Cys^{144\text{-}214}$, $Cys^{100\text{-}260}$, $Cys^{176\text{-}193}$ and $Cys^{204\text{-}229}$ based on the comparative analysis of conserved sequences (Castro *et al.,* 2004; Fan *et al.,* 1999; Lin *et al.,* 2009; Menaldo *et al.,* 2012; Rokyta *et al.,* 2012). Comparing to catalytic triad residues (His^{57}, Asp^{102} and Ser^{195}) owned conservatively by most SV-TLEs (Castro *et al.,* 2004), GSSG-TLE shows the catalytic residues of His^{67}, Asp^{112} and Ser^{202} (Fig. 4).Most snake venom TLEs own two glycosylation sites (Asn-X-Thr) (Au *et al.,* 1993; Magalhaes *et al.,* 2007; Menaldo *et al.,* 2012; Pan *et al.,*1999). GSSG-TLE exists two possible glycosylation sites at Asn^{105}-Tyr^{106}-Thr^{107} and Asn^{124}-Ser^{125}-Thr^{126} (Fig. 4). For *Trimeresurus elegans* elegaxobin II (a TLE), N-deglycosylation affects its interaction with macromolecules (fibrinogen and kininogen) instead of small molecules such as*p*-tosyl-L-arginine methylester (TAME) (Oyama *et al.,* 2003), suggesting its carbohydrate region in the enzyme recognition of these substrates. Nevertheless, for other SV-TLEs, glycans were reported to be important for protein structural stabilization rather than their catalytic functions (Komori and Nikai, 1998). Although widely present in many SV-TLEs, the specific effect and precise importance of carbohydrate moieties on the structure and biological activities of SV-TLEs need to be clarified (Castro *et al.,* 2004; Rokyta *et al.,* 2012).

GSSG-TLE shares high sequence homogeneity with serine proteinases and thrombin-like enzymes from other snakes. Nucleotide sequence of GSSG-TLE shows the highest homologies of 95.9% with calobin (*Agkistrodon ussuriensis*), 94.3% with TLE (*Agkistrodon halys pallas*), 90.7% with serine proteinase 6 (*Adamanteus*), 90.4% with serine proteinase (*Sistrurus catenatus edwardsi*), 89.8% with serine protease 1 (*Trimeresurus stejnegeri*) and serine protease (*Trimeresurus gramineus*), 89.7% with serine protease KN7 precursor (*Trimeresurus stejnegeri*) and 89.6% with serine protease (*Bothrops jararacussu*) and TLE (*Gloydius shedaoensis*). And the encoded amino acid sequence of GSSG-TLE shares the highest identities of 95% with calobin (*Gloydius ussuriensis*), 92% with pallabin-2 (*Gloydius halys*), 90.4% with TLE (*Gloydius shedaoensis*), 90% with pallabin (*Gloydius halys*), 87.8% with two TLEs from *Gloydius halys* and *Agkistrodon halys*.

Site-directed mutagenesis changed prokaryotes preferred codons for TLE expression: We tried hard for optimizing expression conditions of GST-tagged pGEX-6P-1-GSSG-TLE vector in *E. coli* Rosseta. However, despite how hard we tried, we only found protein over-expression of GST tag. We proposed that this phenomenon might contribute to codon preference of *E. coli* by blocking protein translation. Therefore, 11 codons were mutated using site-directed mutagenesis technology by overlap extension PCR. PA1 was non-site-directed mutagenesis sequence of GSSG-TLE, PA2 was

site-directed mutagenesis sequence of GSSG-TLE, blue represents mutated bases, the mutated bases did not changed the coded amino acid sequence (Fig. 5A), indicating that the site-directed mutagenesis was succeeded and further proved to benefit the protein expression of GSSG-TLE.

Construction of recombinant expression vector: Positive pMD18-T-GSSG-TLE plasmid cut with *EcoR* I and *Not* I was ligated to pGEX-6P-1 vector. The recombinant vector pGEX-6P-1-TLE was confirmed by restriction enzyme digestion and PCR and verified by 1% agarose gel electrophoresis (Fig. 5B). The PCR product amplified by TLE forward and TLE reverse primers was 786 bp and pGEX-6P-1 vector was 4900 bp. And the obtained fragments of pGEX-6P-1-GSSG-TLE plasmid digested with *EcoR* I and *Not* I with the size of ~5700 bp, indicating the expression plasmid of pGEX-6P-1-GSSG-TLE was constructed successfully (Fig. 5B).

Protein expression of recombinant GSSG-TLE: To obtain a highly expressed level of pGEX-6P- 1-TLE in *E. coli* Rosseta, we optimized expression temperature, IPTG induction concentration and time. pGEX-6P -1-TLE was optimally expressed in *E. coli* Rosseta induced by 0.2 mM IPTG at 37 ^{o}C for 3 h. However, the recombinant protein GSSG-TLE was expressed mainly in inclusion bodies. High level of GST-tag infused GSSG-TLE protein band with the size of ~ 56 kDa dominated on SDS-PAGE (Fig. 6.) SV-TLEs are known as single-chain enzymes with molecular masses ranged in 25 kDa and 35 kDa (Fan *et al.*, 1999; Lu *et al.*, 2000; Madrigal *et al.*, 2012; Menaldo *et al.*, 2012; Smolka *et al.*, 1998;). GST is ~26 kDa and GSSG-TLE is ~30 kDa, the above result indicated fusion GSSG-TLE protein was correctly expressed (Fig. 6). No doubt, further effort should be made to get more soluble GSSG-TLE, to characterize its molecular characteristics, biological functions and structural properties.

Table 1. Protein identification results for TLE from GSS by HPLC-ESI-MS/MS

MS/MS-derived Sequence	MH^+	z	Peptide Probability
			2.06E-07[&]
NFQM[#]LFGVHSK	1323.65	2	4.08E-04
DDEKDKDIM[#]LIR	1506.74	2	5.13E-02
AAYPVLLAGSSTLC*AGTQQGGK	2150.12	2	9.46E-13
DTC*VGDSGGPLIC*NGQIQGIVSWGAHPC*GQGSK	3412.63	2	1.82E-13
VFDHLDWIK	1186.57	3	2.69E-04
SIIAGNTAVTC*PP	1300.68	2	2.06E-06

[&] Protein probability; [#] Methionine oxidation; *Cysteine oxidation

Table 2. Primer sequences for PCR amplification

Name	Sequences (5' to 3')
M13 forward primer	GTAAAACGACGGCCAGT
middle-reverse primer	CTCCTTGCTGGGTACCTGCACACAATG
middle-forward primer	CATTGTGTGCAGGTACCCAGCAAGGAG
M13 reverse primer	AAACAGCTATGACCATGTTCA
TLE forward primer	*CGC*GAATTCATGGTGCTTATCAAAGTGCTAGC
TLE reverse primer	*TAT*GCGGCCGCCGGGGGGCAGGACGCATCTGTAT
site-directed mutagenesis forward primer 1	*CG*GAATTCATGGTGCTGATTAAAGTCCTGGCAAATCT
site-directed mutagenesis reverse primer 1	GCTGATCCTGCAGCTTTCTTATGCACAAAAATCTTC
site-directed mutagenesis forward primer 2	CATCTCATATTCCAGCAGGTTAATGTTAACACAATGAGGGACATCG
site-directed mutagenesis reverse primer 2	AACATTAACCTGCTGGAATATGAGATGTGTCGAGTACC
	*TA*GCGGCCGCTTACGGGGGGCAGGACGCATC

The underlined sequences are *EcoR* I and *Not* I restriction sites, respectively. The italic sequences are protective bases.

Fig. 1 The Shedao Island and the snake of *Gloydius shedaoensis shedaoensis* (GSS). (A) The photo of Shedao Island located 25-miles northwest of the Lvshunkou District in the Sea of Bohai. It is about 1 square kilometers rising 217 meters above sea level. About 20, 000 GSSs live in the island. Insert plot is a photo of a GSS snake. (B) GSS snakes are rest in rock cracks and trees.

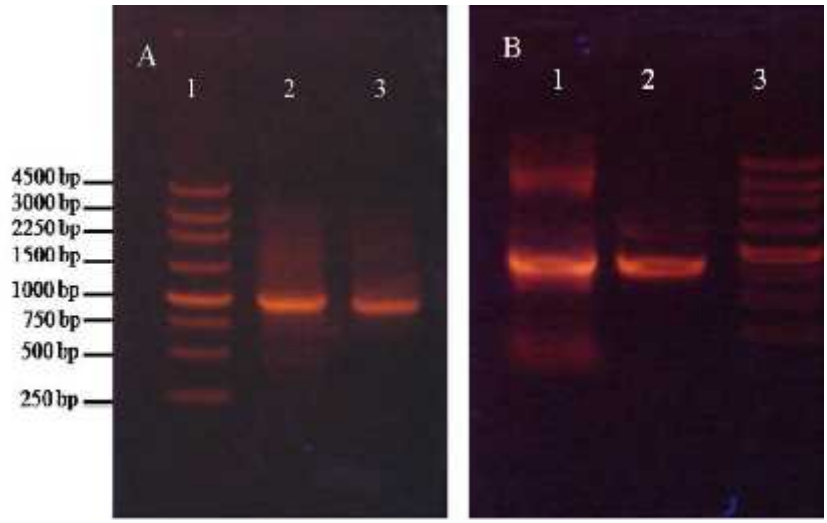

Fig. 2 Amplification of GSSG-TLE gene. (A) PCR amplification 5' sequences and 3' sequences of GSSG-TLE. Lane 1: 250 bp DNA marker; Lane 2: The 5'-sequence of GSSG-TLE was amplified by PCR using M13 forward primer and Middle-reverse primer; Lane 3: The 3'-sequence of GSSG-TLE was obtained using M13 reverse primer and Middle-forward primer. (B) PCR amplification full-length sequence of GSSG-TLE. Lane 1: The amplified product of GSSG-TLE using TLE forward primer and TLE reverse primer and GSSG-cDNA library plasmid as template; Lane 2: The purified GSSG-TLE PCR product by gel extraction kit; Lane 3: 250 bp DNA marker.

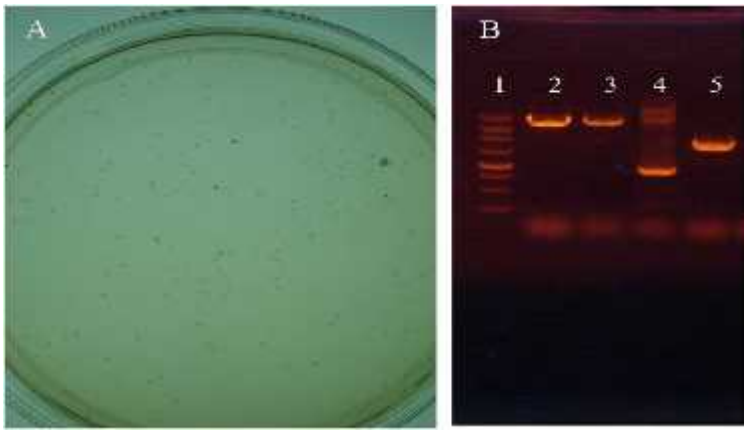

Fig. 3 Construction of recombinant pMD18-T-TLE cloning vector. (A) Blue-white screening of pMD18-T-GSSG-TLE positive recombinant plasmid. (B) Identification of pMD18-T-TLE recombinant plasmid by restriction enzyme digestion and PCR analysis. Lane 1: 250 bp DNA marker; Lane 2: The recombinant plasmid pMD18-T-GSSG-TLE were digested with *EcoR* I; Lane 3: The recombinant plasmid pMD18-T-GSSG-TLE were digested with *Not* I; Lane 4: The PCR product of recombinant plasmid pMD18-T-GSSG-TLE; Lane 5: The recombinant plasmid pMD18-T-GSSG-TLE.

```
1   ATGGTGCTTATCAAAGTGCTAGCAAACCTCCTGATACTACAGCTTTCTTATGCACAAAAA
1    M  V  L  I  K  V  L  A  N  L  L  I  L  Q  L  S  Y  A  Q  K

61  TCTTCTGAACTGATCATTGGAGGTGATGAATGTAACATAAATGAACATCGTTTCCTTGTA
21   S  K  K  V  P  N  E  D  E  Q  T  R  V  P  K  E  K  F  F  C

121 GCCTTATACACCTCTAGATTTAGGACTTTGTTCTGCGGTGGGACTTTGATCAACCAGGA
41   A  L  Y  T  S  R  F  R  T  L  F  C  G  G  T  L  I  N  Q  E

181 TGGGTGCTCACTGCTGCACACTGTGACAGGAAAAATTTCCGGATAAAGCTTGGTATTCT
61   W  V  L  T  A  A  H  C  D  R  K  N  F  R  I  K  L  G  I  H

241 AGCAAAAAGGTACCAAATGAGGATGAGCAGACAAGAGTCCCAAAGGAGAAGTTCTTGT
81   S  K  K  V  P  N  E  D  E  Q  T  R  V  P  K  E  K  F  F  C

301 CTCAGTAGCAAAAACTATACCCTTTGGGACAAGGACATCATGTTGATCAGGCTGGACAC
101  L  S  S  K  N  Y  T  L  W  D  K  D  I  M  L  I  R  L  D  S

361 CCTGTTAAGAACAGTACACACATTGCGCCTTTCAGCTTGCCCTCCAGCCCTCCCAGTGTG
121  P  V  K  N  S  T  H  I  A  P  F  S  L  P  S  S  P  P  S  V

421 GGCTCAGTTTGCCGTATTATGGGATGGGGCAGAATCTCACCCACTGAAGAGACTTATCCC
141  G  S  V  C  R  I  M  G  W  G  R  I  S  P  T  E  E  T  Y  P

481 GATGTCCCTCATTGTGTTAACATTAACCTACTCGAATATGAGATGTGTCGAGTACCTTAC
161  D  V  P  H  C  V  N  I  N  L  L  E  Y  E  M  C  R  V  P  Y

541 CCAGAATTTGGGTTGCCAGCGACAAGCAGAACATTGTGTGCAGGTATCCTGGAAGGAGGC
181  P  E  F  G  L  P  A  T  S  R  T  L  C  A  G  I  L  E  G  G

601 AAAGATACATGTCGGGGTGACTCTGGGGGACCCCTCATCTGTAATGGACAATTCCAGGC
201  K  D  T  C  R  G  D  S  G  G  P  L  I  C  N  G  Q  F  Q  G

661 ATTGCATCTTGGGGAGACGATCCTTGTGCCCAACCGCATAAGCCTGCCGCGTACACCAAG
241  I  A  S  W  G  D  D  P  C  A  Q  P  H  K  P  A  A  Y  T  K

721 GTCTTCGATCATCTTGACTGGATCAAGAGCATTATTGCAGGAAATACAGATGCGTCCTGC
261  V  F  D  H  L  D  W  I  K  S  I  I  A  G  N  T  D  A  S  C

781 CCCCCGTGA
281  P  P  *
```

Fig.4 The nucleotide sequence and amino acid sequence of GSSG-TLE. The cDNA length of GSSG-TLE is 286 nucleotides by sequencing and encoding 262 amino acids.

Fig. 5 (A) The site-directed mutagenesis of GSSG-TLE. PA1: non-site-directed mutagenesis sequence of GSSG-TLE; PA2: site-directed mutagenesis sequence of GSSG-TLE. Blue represents mutated bases. (B) Identification of pGEX-6P-1-GSSG-TLE recombinant plasmid by restriction enzyme digestion and PCR analysis. Lane 1: 250 bp DNA marker; Lane 2: The recombinant plasmid pGEX-6P-1-GSSG-TLE; Lane 3 and 4: The recombinant plasmid pGEX-6P-1-GSSG-TLE were digested with *EcoR* I and *Not* I; Lane 5: The PCR product of recombinant plasmid pGEX-6P-1-GSSG-TLE.

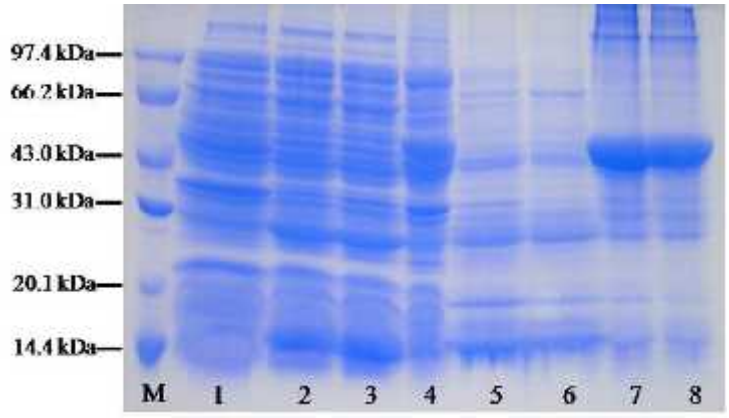

Fig. 6 The expression of GSSG-TLE protein in *E. coli* Rosseta induced by 0.2 mM TPTG for 3 h. Lane M: Protein molecular weight marker; Lane 1: The pGEX-6P-1 before induction; Lane 2: The pGEX-6P-1-GSSG-TLE before induction; Lane 3: The pGEX-6P-1 after induction; Lane 4: The pGEX-6P-1-GSSG-TLE after induction; Lane 5 and 6: The supernatant of pGEX-6P-1-GSSG-TLE after induction by ultrasonication; Lane 7 and 8: The precipitate of pGEX-6P-1-GSSG-TLE after induction by ultrasonication.

Acknowledgments: This work was supported by the Liaoning BaiQianWan Talent Project (2012921015), Distinguished Young Scholars of College and University Liaoning Province (LJQ2011094) and National Natural Science Foundation of China (81171957, 81100722, 81272186).

REFERENCES

Au, L. C., S. Lin, J. S. Chou, G. W. Teh, K. J. Chang and C. M. Shih (1993). Molecular cloning and sequence analysis of the cDNA for ancrod, a thrombin-like enzyme from the venom of *Calloselasma rhodostoma*. Biochem. J. 294:387-390.

Castro, H. C., R. B. Zingali, M. G. Albuquerque, M. Pujol-Luz and C. R. Rodrigues (2004). Snake venom thrombin-like enzymes: from reptilase to now. Cell Mol. Life Sci. 61:843-856.

Giron, M. E., A. Rodriguez-Acosta, A. M. Salazar, E. E. Sanchez, J. Galan, C. Ibarra and B. Guerrero (2013) Isolation and characterization of two new non-hemorrhagic metalloproteinases with fibrinogenolytic activity from the mapanare *(Bothrops colombiensis*) venom. Arch Toxicol 87(1):197-208.

Guo, C., S. Liu, Y. Yao, Q. Zhang and M-Z. Sun (2012). Past decade study of snake venom L-amino acid oxidase. Toxicon 60(3):302-311.

Guo, C., S. Liu and M-Z. Sun(2013). Sequence and bioinformatic characterization of ESTs originated from *Gloydius shedaoensis shedaoensis* venom gland. Anat. Record 296: 807-814.

Farid, T. M., A. T. Tu and M. F. el-Asmer (1989). Characterization of cerastobin, a thrombin-like enzyme from the venom of *Cerastes vipera* (Sahara sand viper). Biochemistry 28:337-371.

Fan, C., Y. Qian, S. Yang and Y. Gong (1999). Cloning, sequence analysis and expression in *E. coli* of the cDNA of the thrombin-like enzyme (pallabin) from the venom of *Agkistrodon halys pallas*. Biochem. Mol. Biol. Int. 47:217-225.

Jiang, X., J. Xu and Q. Yang (2009). Soluble expression, purification, and characterization of *Gloydius shedaoensis* venom gloshedobin in *Escherichia coli* by using fusion partners. Appl. Microbiol. Biotechnol. 85:635-642.

Komori, Y. and T. Nikai (1998). Chemistry and biochemistry of kallikrein-like enzyme from snake venoms. J. Toxicol. Toxin Rev. 17:261-277.

Laraba-Djebari, F., M. F. Martin-Eauclaire and P. Marchot (1992). A fibrinogen-clotting serine proteinase from *Cerastes cerastes* (horned viper) venom with arginine-esterase and amidase activities. Purification, characterization and kinetic parameter determination. Toxicon 30: 1399- 1410.

Lin, Y., X. Yu, Q. He, Li, H., D. Li, X. Song, Y. Wang, H. Wen, H. Deng and J. Deng (2009). Expression and functional characterization of chitribrisin, a thrombin-like enzyme, in the venom of the Chinese green pit viper (*Trimeresurus albolabris*). Protein Expr. Purif. 67:48-52.

Liu, S., C. Guo, Z. Hou, M-Z. Sun, L. Wang, X. Wang and L. Sun (2010). Construction of a cDNA library for Chinese *Gloydius shedaoensis shedaoensis* venom gland. Chin. J. Zoology 45:136-140.

Liu, S., M-Z. Sun and F. T. Greenaway (2006). A novel plasminogen activator from *Agkistrodon blomhoffii Ussurensis* venom (ABUSV-PA): purification and characterization. Biochem. Biophys. Res. Commun. 348:1279–1287.

Liu, S., Yang, F., Sun, M-Z., Zhang, Q., Gao, Y and S. Shao (2011). "Anatomical" view of the protein composition and protein characteristics for *Gloydius shedaoensis* snake venom via proteomics approach. Anat. Rec. (Hoboken) 294: 273-282.

Lu, Q., Jin, Y., Li, D., Wang, W and Y. Xiong (2000). Characterization of a thrombin-like enzyme from the venom of *Trimeresurus jerdonii*. Toxicon 38:1225-1236.

Ma, C., S. Liu., J. Zong., Z. Yuan., M-Z. Sun., Y. Wu., J. Li, L. Sun and Q. Zheng (2009). Proteomic analysis of *Gloydius shedaoensis shedaoensis* venom by two-dimensional electrophoresis combined to high performance liquid chromatography-electro-spray ionization tandem mass spectrometry: a preliminary study. Chin. J. Zoology 44: 70-77.

Madrigal, M., L. Sanz., M. Flores-Diaz., M. Sasa., V. Nunez., A. Alape-Gironand J. J. Calvete (2012). Snake venomics across genus Lachesis. Ontogenetic changes in the venom composition of *Lachesis stenophrys* and comparative proteomics of the venoms of adult *Lachesis melanocephala* and *Lachesis acrochorda*. J. Proteomics 77: 280-297.

Magalhaes, A., H. P. Magalhaes., M. Richardson., S. Gontijo., R. N. Ferreira., A. P. Almeida and E. F. Sanchez (2007). Purification and properties of a coagulant thrombin-like enzyme from the venom of *Bothrops leucurus*. Comp. Biochem. Physiol. A Mol. Integr. Physiol. 146:565- 575.

Menaldo, D. L., C. P. Bernardes., N. A. Santos-Filho., A. Moura Lde., A. L. Fuly., E. C. Arantes and S. V. Sampaio (2012). Biochemical characterization and comparative analysis of two distinct serine proteases from *Bothrops pirajai* snake venom. Biochimie 94(12):2545-2548.

Oyama, E and H. Takahashi (2003). Purification and characterization of a thrombin like enzyme, elegaxobin II, with lys-bradykinin releasing activity from the venom of *Trimeresurus elegans* (Sakishima-Habu). Toxicon 41:559-568.

Ouyang, C., C. M. Teng and T. F. Huang (1992). Characterization of snake venom components acting on blood coagulation and platelet function. Toxicon 30:945-966.

Park, D., H. Kim., K. Chung., D. S. Kim and Y. Yun (1998). Expression and characterization of a novel plasminogen activator from *Agkistrodon halys* venom. Toxicon 36:1807-1819.

Pan, H., X. Du., G. Yang., Y. Zhou and X Wu (1999). cDNA cloning and expression of acutin, a thrombin-like enzyme from *Agkistrodon actus.* Biochem. Biophys. Res. Commun. 255:412-415.

Rokyta, D. R., A. R. Lemmon., M. J. Margres and K. Aronow (2012). The venom-gland transcriptome of the eastern diamondback rattlesnake (*Crotalus adamanteus*). BMC Genomics 13:312

Smolka, M. B., S. Marangoni., B. Oliveira and J. C. Novello (1998). Purification and partial

characterization of a thrombin-like enzyme, balterobin, from the venom of *Bothrops alternatus*. Toxicon 36:1059-1063.

Yang, D., M. Peng., H. Yang., Q. Yang and J. Xu (2009). Expression, purification and characterization of *Gloydius shedaoensis* venom gloshedobin as Hsp70 fusion protein in *Pichia pastoris*. Protein Expr. Purif. 66:138-142.

Yang, Q., M. Li., J. Xu., Y. Bao., X. Lei and L. An (2003). Expression of gloshedobin, a thrombin-like enzyme from the venom of *Gloydius shedaoensis* in *Escherichia coli*. Biotechnol. Lett. 25:101-104.

Yang, Q., X. Xu., M. Li., X. Yuan., Z. Su., J. Janson and L. An (2002). Cloning and expression of defibrase cDNA from the venom of *Gloydius shedaoensis*. Biotechnol. Lett. 24:135-138.