动 物 学 研 究 2005, Aug. 26 (4): 337 - 343 CN 53 - 1040/Q ISSN 0254 - 5853
Zoological Research

Molecular Cloning and Sequence Analyses of cDNAs Encoding Seven C-type Lectin-like Protein Subunits from *Daboia russellii siamensis*

ZHONG Shu-rong[1, 2], JIN Yang[1], LI Dong-sheng[1], WANG Wan-yu[1], XIONG Yu-liang[1, *]

(1. *Department of Animal Toxinology, Kunming Institute of Zoology, the Chinese Academy of Sciences, Kunming 650223, China;*
2. *Graduate School of the Chinese Academy of Sciences, Beijing 100039, China*)

Abstract: Total mRNA was extracted from a venom gland of snake *Daboia russellii siamensis* following the manufacturer's protocol of the PolyATtract System 1000 kit purchased from Promega Biotech. cDNAs encoding C-type lectins were amplified by RT-PCR and subcloned into a pMD18-T vector. Fourteen positive clones were selected for nucleotide sequencing and seven cDNAs encoding various snake venom C-type lectin-like protein precursors, designated as DRS-L1, DRS-L2, DRS-L3, DRS-L4 DRS-L5, DRS-L6 and DRS-L7, were obtained. Amino acid sequences of these proteins were deduced and each contains a carbohydrate recognition domain. Of all the deduced protein sequences, only DRS-L1 seemed to represent a closer sequence similarity to α subunits of other known snake venom C-type lectin-like proteins using the BLAST program. Homology comparison combined with analysis of cysteine position indicate that DRS-L1 and DRS-L2 are probably the light chain LC2 and LC1 of factor Ⅹ activator from *Daboia russellii siamensis* venom, respectively. DRS-L3 and DRS-L4 might be the β subunits of higher molecular weight C-type lectin-like proteins while DRS-L5 and DRS-L6 might be β subunits of lower molecular weight C-type lectin-like proteins. DRS-L7 might be the β subunit of a platelet membrane glycoprotein Ib-binding protein similar to echicetin.

Key words: *Daboia russellii siamensis*; SV-CTTLs; SV-CTLPs; cDNA; Sequence analysis

圆斑蝰蛇毒七个 C - 型凝集素样蛋白亚基的 cDNA 克隆与序列分析

钟树荣[1,2], 金 扬[1], 李东升[1], 王婉瑜[1], 熊郁良[1,*]

(1. 中国科学院昆明动物研究所 动物毒素室, 云南 昆明 650223; 2. 中国科学院研究生院, 北京 100039)

摘要: 按照 Promega 公司的 mRNA 提取试剂盒操作手册, 从圆斑蝰蛇(*Daboia russellii siamensis*)的毒腺中提取 mRNA; 利用 RT-PCR 的方法进行体外扩增, 获得 C - 型凝集素蛋白的基因, 克隆到 pMD18-T 载体中。随机挑选 14 个阳性克隆进行核酸测序, 获得 7 个编码不同蛇毒 C - 型凝集素样蛋白亚基的 cDNA, 分别命名为 DRS-L1、DRS-L2、DRS-L3、DRS-L4、DRS-L5、DRS-L6 和 DRS-L7。由基因序列推导出的氨基酸序列表明, 克隆到的 7 个蛇毒 C - 型凝集素样蛋白的亚基中均有糖识别结构域存在。BLAST 分析显示, 仅有 DRS-L1 的蛋白序列与目前已知的蛇毒 C - 型凝集素样蛋白的 α 亚基相似。序列同源性比较并结合半胱氨酸位点分析, 推测 DRS-L1 和 DRS-L2 可能分别是圆斑蝰蛇毒Ⅹ因子激活剂的轻链 LC2 和 LC1。DRS-L3 和 DRS-L4 可能是高分子量的蛇毒 C - 型凝集素样蛋白的 β 亚基, 而 DRS-L5 和 DRS-L6 可能是低分子量的蛇毒 C - 型凝集素样蛋白的 β 亚基。DRS-L7 可能是类似于血小板膜糖蛋白 Ib 结合蛋白的 β 亚基。

关键词: 圆斑蝰蛇; 蛇毒 C - 型真凝集素; 蛇毒 C - 型凝集素样蛋白; cDNA; 序列分析
中图分类号: Q959.6; Q956; Q51 **文献标识码:** A **文章编号:** 0254 - 5853(2005)04 - 0337 - 07

Received date: 2005 - 04 - 28; Accepted date: 2005 - 06 - 11

* Corresponding author(通讯作者), E-mail: xiongyl@mail.kiz.ac.cn, Tel: +86 - 871 - 5192476, Fax: +86 - 871 - 5196747
收稿日期: 2005 - 04 - 28; 接受日期: 2005 - 06 - 11

Snake venoms contain a large number of proteins and polypeptides with diverse pharmacological activities. Some snake venom proteins are enzymes, such as L-amino acid oxidase, metalloproteinase, phoshpolipase A_2, serine proteinase, and hyaluronidase. Others are non-enzymatic proteins like neurotoxins, disintegrins and lectins (de Lima et al, 2005). Currently, the studies on lectins are hot topics in this field. Lectins are proteins or glycoproteins that specifically recognize and bind (or crosslink) to carbohydrates without catalyzing modifications of the carbohydrates except for enzymes and antibodies. They widely distribute in plants, animals and microorganisms (Dodd & Drickamer, 2001). Among animal lectins, Ca^{2+}-dependent-type lectins (C-type lectins) are the most diverse family. They are generally multidomain proteins which contain one or more copies of a highly conserved region composed of 115 – 130 amino acid residues, called the carbohydrate recognition domain (CRD). In the presence of Ca^{2+}, C-type lectins initiate a wide range of biological events, such as cell-cell adhesion, serum glycoprotein turnover and immune responses to potential pathogens (Weis et al, 1998; Drickamer, 1999).

Snake venom lectins belong to Group Ⅶ C-type lectin super-family and consist simply of CRDs (Weis et al, 1998; Drickamer, 1999). Based on their structural and functional characteristics, snake venom C-type lectins can be classified into two subgroups: snake venom C-type true lectins (SV-CTTLs) and snake venom C-type lectin-like proteins (SV-CTLPs) (Lee & Zhang, 2003). SV-CTTLs are composed of two covalently linked identical subunits, each consisting of 135 – 141 amino acid residues. They have saccharide-binding activities and can induce erythrocyte aggregation. SV-CTLPs are similar in structure to SV-CTTLs, but they are heterodimers lacking carbohydrate binding activity and have different pharmacological actions on thrombosis and hemostasis as exogenous modulators (Ogawa et al, 2005). At present, many snake venom C-type lectins have been purified and characterized (Ogawa et al, 2005). Recently, some SV-CTTLs have been cloned from various snake venom glands (Hirabayashi et al, 1991; Xu et al, 1999; Li et al, 2003; Zha et al, 2001) and large numbers of SV-CTLPs also have been cloned (Shin et al, 2000; Wei et al, 2002; Lee et al, 2003; Jasti et al, 2004). Recently, several factor Ⅹ activators, such as RVV-Ⅹ from *D. russellii siamensis* venom and VLFXA from *V. lebetina* venom have been purified and characterized (Takeya et al, 1992; Siigur et al, 2004).

Daboia russellii (alias *D. russelli*, *Vipera russellii*, *V. russelli* or Russell's viper) has been classified into five subspecies according to its geographic origin. *D. russellii siamensis* (alias *D. russelli siamensis*, *V. russellii siamensis* or *V. russelli Siamensis*) is a subspecies which mainly distributes in southern China, Burma, Thailand and Cambodia (Wuster et al, 1997; Prasad et al, 1999). Some components such as phospholipase A_2 (Li et al, 1985) and coagulation factor Ⅹ activator (Takeya et al, 1992), have been purified from the venom of this subspecies. Up to now, there are no reports on purification and cloning of snake venom C-type lectins from this subspecies venom. It is of great interest to know if there are C-type lectins expressed in *D. russellii siamensis* venom gland. Here, we reported seven cDNAs (DRS-L1 to DRS-L7) encoding various SV-CTLP subunits from *D. russellii siamensis* venom gland.

1 Materials and Methods

1.1 Materials

Living *D. russellii siamensis* snake was collected from Baise City in Guangxi Province, China. Ex *Taq*™ DNA polymerase, dNTP, DNA marker, pMD18-T vector and *Escherichia coli* strain JM109 were purchased from TaKaRa Biotechnology Co., Ltd (Dalian). The PolyATtract System 1000 kit and Reverse Transcription System kit were from Promega Biotech.

1.2 Design of oligonucleotide primers

Two oligonucleotide primers were designed according to the highly conserved non-coding regions of cDNA sequence of TSL, a snake venom C-type true lectin isolated from *Trimeresurus stejnegeri* venom and other related cDNA sequences of factor Ⅸ or Ⅹ binding proteins and platelet membrane glycoprotein binding proteins. The two primers were named as Lec 5 (5′-AGGGAAGGAAGGAAGACCATGG-3′, in the sense direction) and Lec 3 (5′-GGGGGCTTCCTTGCTTCTC-CAG-3′, in the anti-sense direction), respectively (Lee et al, 2003). Both two primers were synthesized by GeneCore Biotechnologies (Shanghai) with an ABI 3948 DNA synthesizer.

1.3 cDNA cloning and sequencing

Isolation of mRNA and reverse transcription were conducted using PolyATtract System 1000 kit and Reverse Transcription System kit, respectively, following the manufacturer's protocols. DNAs were amplified by PCR using total RT-PCR products as template and the two oligonucleotide primers. The amplification reaction

was catalyzed by *Taq* DNA polymerase and began with an initial 5 min denaturation at 94 ℃, followed by 35 cycles of 30 s denaturation at 94 ℃, 40 s annealing at 52 ℃ and 1 min extension at 72 ℃, a 10 min final extension at 72 ℃ on a GeneAmp PCR System 2400 (Perkin Elmer). Then PCR product was subcloned into pMD18-T vector and transformed into *Escherichia coli* strain JM109. The white single colonies grew in the LB agar media containing ampicillin, IPTG and X-gal were screened by PCR. The positive clones were sequenced by GeneCore Biotechnologies (Shanghai) with an ABI 377 DNA automatic sequencer.

1.4 Sequence analysis

Amino acid sequences were deduced from nucleotide sequences using Vector NTI Suite 8.0 software. The sequences of cDNAs and the deduced proteins were compared with those in the GenBank database using the BLAST program (Altschul et al, 1997). The sequences of snake venom C-type lectins were searched with the non-redundant protein sequence database using FASTA program (Pearson & Lipman, 1988). Alignment was made by CLUSTAL W program (Thompson et al, 1994).

2 Results and Analyses

PCR amplification of total cDNAs with the designed primers resulted in a DNA fragment estimated about 500 bp (data not shown). Then PCR product was subcloned into pMD18-T vector. Fourteen positive clones which contained inserts were selected for nucleotide sequencing. The results indicated that each of these clones contained a cDNA fragment which encoded a C-type lectin. Seven cDNAs encoding various C-type lectin precursors, designated as DRS-Ls (i.e. DRS-L1 to DRS-L7) were obtained from *D. russellii siamensis* venom gland. Among the fourteen cDNA sequences, each of DRS-L4, DRS-L5, DRS-L6 and DRS-L7 was coded by a single clone. Meanwhile, two identical clones for each of DRS-L1 and DRS-L2 were found. Six identical clones were found for DRS-L3.

The complete nucleotide sequences of DRS-Ls were submitted to GenBank and accession numbers are as follows: DRS-L1 (DQ060414), DRS-L2 (DQ060415), DRS-L3 (DQ060416), DRS-L4 (DQ060417), DRS-L5 (DQ060418), DRS-L6 (DQ060419) and DRS-L7 (DQ060420), respectively. The deduced amino acid sequences of DRS-Ls are listed in Fig. 1. Comparison of the deduced amino acid sequences of DRS-Ls with those of proteins in the BLAST database showed that DRS-Ls have strong similarities to other known SV-CTLP subunits. These deduced protein sequences of DRS-Ls have highly conserved regions composed of 23 amino acid residues (82.6% identity) which are putative signal peptide regions when compared with other SV-CTLP subunits reported previously.

3 Discussion

Drickamer (1992) indicated that the saccharide-binding specificity of CRDsis related to the amino acid sequence found in positions 96 - 98 (RSL numbering system). In C-type galactose-binding protein, this tripeptide sequence is Gln-Pro-Asp (QPD). In C-type mannose-binding protein, the sequence of this position is Glu-Pro-Asn (EPN). On the other hand, three residues (Glu-127, Asn-142 and Asp-143) form a Ca^{2+}-binding site in this type of proteins according to Drickamer (1992). Neither saccharide-binding site nor Ca^{2+}-binding site is found in DRS-Ls. So it is presumed that none of DRS-Ls is SV-CTTL.

Generally, SV-CTLPs consist of simply CRDs and have no extra domains. However, SV-CTLPs are also found in multidomain proteins like factor X activators from the venoms of Viperinae. Each of them comprises one heavy chain which belongs to metalloproteinase and two light chains which belong to C-type lectins. Fig. 2 compared the deduced protein sequences of DRS-L1 and DRS-L2 with those of reported SV-CTLPs. The results showed that DRS-L2 has the most similarity with RVV X LC1 (82.9%). RVV X LC1 is one of RVV-X light chains isolated from *D. russellii siamensis* venom and it was sequenced by Edman degradation (Takeya et al, 1992). If the putative signal peptide is removed from DRS-L2, 98.4% identity can be observed between DRS-L2 and RVV X LC1. Only two amino acid residues are different between the two molecules. It showed a K for E substitution at position 51 and a V for E substitution at position 107 in mature DRS-L2 when compared with RVV X LC1. So DRS-L2 may probably the light chain LC1 of factor X activator the same as RVV X LC1 expressed in *D. russellii siamensis* venom gland. Like RVV X LC1, DRS-L2 includes seven cysteine residues and resembles β subunits of lower molecular weight SV-CTLPs. The deduced protein sequence of DRS-L1 has the highest identity with VLF X ALC2 (82.3%), the light chain LC2 of VLF X A isolated from *V. lebetina* venom (Siigur et al, 2004), since the primary structure of RVV X LC2 is undetermined. So DRS-L1 is probably the light chain LC2 of factor X activator expressed in *D. russellii siamensis* venom

```
DRS-L1    (1) MGRFISVSFGLLVVFLSLSGTGAGLDCPPDSSLYRYFCYRVFKEHKTWEA
DRS-L2    (1) MGRFISVSFGWLVVFLSLSGTEAVLDCPSGWLSYEQHCYKGFNDLKNWTD
DRS-L3    (1) MGRFISVSFGLLVVFLSLSGTEAAFCCPSGWSAYDQNCYKVFTEEMNWAD
DRS-L4    (1) MGRFISISFGLLVVFLSLSGTEAAFCCPSGWSAYDQNCYKVFTEEMNWAD
DRS-L5    (1) MGRFISISFGLLVMFLSLSGTGAKQDCLSDWSFYEGYCYKVFNEKKTWED
DRS-L6    (1) MGRFISVSFGLLVVFLSLSGTGA--DCPSEWSSHEGHCYKVFKLLKTWED
DRS-L7    (1) MGRFISISFGLLVVFLSLSGTGAKQDCLSDWSFYEGYCYKVFNEKKTWED

DRS-L1   (51) AERFCMEHPNNGHLVSIESMEEAEFVAKLLSNTTGKFITHFWIGLMIKDK
DRS-L2   (51) AEKFCTEQKKGSHLVSLHSREEEKFVVNLIS----ENLEYPATWIGLGNM
DRS-L3   (51) AEKFCTEQHKGSHLLSLHNIAEADFVLKKTL----AMLKDGVIWMGLNDV
DRS-L4   (51) AEKFCTEQKKGSHLVSLHSREEEKFVVNLIS----ENLEYPATWIGLGNM
DRS-L5   (51) AEKFCTEQHKGSHLLSLHNIAEADFVLKKTL----AMLKDGVIWMGLNDV
DRS-L6   (49) AEKFCTEQKKGSHLVSLHSREEEKFVVNLIS----ENLEYPATWIGLGNM
DRS-L7   (51) AEKFCNEQVNGGYLVSFRSSEEMDFVIRMTF----PIFRFDFFWIGLRDF

DRS-L1  (101) EQECSSEWSDGSSVSYDKLGKQEFRKCFVLEKESGYRMWFNRNCEERYLF
DRS-L2   (97) WKDCRMEWSDRGNVKYKALAEE--SYCLIMITHE--KVWKSMTCNFIAPV
DRS-L3   (97) WNECNWGWTDGAKLDYKAWNEG--TNCFVFKIAK--NHWSHMDCSSTHNF
DRS-L4   (97) WKDCRMEWSDRGNVKYKALAEE--SYCLIMITHE--KVWKSMTCNFIAPV
DRS-L5   (97) WNECNWGWTDGAKLDYKAWNEG--TNCFVFKIAK--NHWSHMDCSSTHNF
DRS-L6   (95) WKDCRMEWSDRGNVKYKALAEE--SYCLIMITHE--KVWKSMTCNFIAPV
DRS-L7   (97) WRDCYWRWSDGVNLDYKAWSRE--PNCFVSKTTD--NQWLRWNCNDPRYF

DRS-L1  (151) VCKVPPEC
DRS-L2  (143) VCKF----
DRS-L3  (143) VCKFRV--
DRS-L4  (143) VCKF----
DRS-L5  (143) VCKFRV--
DRS-L6  (141) VCKF----
DRS-L7  (143) VCKSRVSC
```

Fig. 1 Deduced amino acid sequences of *D. russellii siamensis* lectins (DRS-Ls)
The putative signal peptide is in bold. Conserved cysteine residue is shaded. The extra cysteine residue is underlined.

gland. Like VLFXALC2, DRS-L1 includes eight cysteine residues and resembles α subunits of higher molecular weight SV-CTLPs such as flavocetin-A which has an extra cysteine in C-terminus. Meanwhile, a potential N-glycosylation site, Asn-X-Thr, is found in both DRS-L1 and DRS-L2, respectively (Fig. 2).

In SV-CTLPs, the positions of cysteine residues are highly conserved. Commonly, lower molecular weight SV-CTLPs contain seven cysteine residues in each subunit with the exception of echicetin β subunit (Mizuno et al, 1997; Peng et al, 1994; Jasti et al, 2004). While in higher molecular weight SV-CTLPs, there are eight cysteine residues in each subunit. Compared with protein sequences of lower molecular weight SV-CTLPs, the additional cysteine residues are found in the C-terminal regions of α subunits (Cys135) and N-terminal regions of β subunits (Cys3) in higher molecular weight SV-CTLPs (Fukuda et al, 2000). DRS-L3 and DRS-L4 include eight cysteines and resemble β subunits of higher molecular weight SV-CTLPs like flavocetin-A, stejnulxin and TMVA which have extra cysteines in their N-terminus (Fig. 3). So DRS-L3 and DRS-L4 may probably β subunits of higher molecular weight SV-CTLPs expressed in *D. russellii siamensis* venom gland. DRS-L5 and DRS-L6 include seven cysteine residues and resemble β subunits of lower molecular weight SV-CTLPs like agkisacutacin (Fig. 3). This indicates that DRS-L5 and DRS-L6 may probably β subunits of lower molecular weight SV-CTLPs expressed in *D. russellii siamensis* venom gland.

Many SV-CTLPs were characterized as platelet agonists or inhibitors, through targeting platelet membrane glycoprotein receptors. Among which, platelet membrane glycoprotein Ib-binding proteins (GPIb-BPs) represent the predominant ones. Based on the determined crystal structure of flavocetin-A, one well-studied GPIb-BP from *T. flavoviridis* venom, the platelet GPIb-binding sites of flavocetin-A are mainly attributed to two hydrophilic patches in β subunit but not in α subunit. Hydrophilic patch I is formed by Ser-46, Ser-47, Glu-48 and Asp-51 and hydrophilic

(A) α subunit

```
DRS-L1          (1) MGRFISVSFGLLVVFLSLSGTGAGLDCPPDSSLYRYFCYRVFKEHKTWEA
VLFXALC2        (1) MGRSISVSFGLLAVFLSLSGTGAGLDCPPDSSPYRYFCYRVFKEQKNWAD
TMVA            (1) MGRFTFVSFGLLVVFLSLSGTGADFDCIPGWSAYDRYCYQAFSEPKNWED
stejnulxin      (1) MGRFISVSFGLLVVFLSLSGTGADFDCPSGWSAYDWYCYKPFNEPQTWDD
flavocetin-A    (1) MERLIFVSFGLLVVILSLSGTGADFDCIPGWSAYDRYCYQAFSKPKNWED

DRS-L1         (51) AERFCMEHPNNGHLVSIESMEEAEFVAKLLSNTTGKFITHFWIGLMIKDK
VLFXALC2       (51) AERFCAERPNNGHLVSIESMEEAEFVAQLLSKITGKFITHFWIGLRIEDK
TMVA           (51) AESFCEEGVKTSHLVSIESSGEGDFVAQLVAEKIKTSFQYVWIGLRIQNK
stejnulxin     (51) AERFCTEQAKGGHLVSIESSGEADFVGQLVSENIQRPEIYVWIGLRDRRK
flavocetin-A   (51) AESFCEEGVKTSHLVSIESSGEGDFVAQLVAEKIKTSFQYVWIGLRIQNK

DRS-L1        (101) EQECSSEWSDGSSVSYDKLGKQEFRKCFVLEKESGYRMWFNRNCEERYLF
VLFXALC2      (101) KQQCRSEWSDGSSVSYDNLLKREFRKCFGLEKGTGYRSWFNLNCEEPYPF
TMVA          (101) EQQCRSEWSDASSVNYENLFKQSSKKCYALKKGTELRTWFNVYCGRENPF
stejnulxin    (101) EQQCSSEWSDGTSIIYVNWNKGESQMCQGLSKWTNFLKWDNTDCQAKNPF
flavocetin-A  (101) EQQCRSEWSDASSVNYENLVKQFSKKCYALKKGTELRTWFNVYCGTENPF

                                  Identity
DRS-L1        (151) VCKVPPEC      100.0 %
VLFXALC2      (151) VCKVPPNC       82.3 %
TMVA          (151) VCKYTPEC       55.1 %
stejnulxin    (151) VCKFPPQC       54.4 %
flavocetin-A  (151) VCKYTPEC       53.2 %
```

(B) β subunit

```
DRS-L2          (1) MGRFISVSFGWLVVFLSLSGTEAVLDCPSGWLSYEQHCYKGFNDI
RVVXLC1         (1) ----------------------VLDCPSGWLSYEQHCYKGFNDL
VLFXALC1        (1) MGRFISVSFGLLVVFLSLSGTGADFDCPSDWVSYDQHCYKAFNDL
TSV-FIX-BP      (1) MGRFIFVSFGLLVVFLSLSGTA--ADCLSGWSSYEGHCYKPFNEL
echicetin       (1) MGRFISVSFGLLVLLLSLSGTG--ANCLPDWSVYEGYCYKVFKER

DRS-L2         (46) KNNWTDAEKFCTEQKKGSHLVSLHSREEEKFVVNLISENLEYPATW
RVVXLC1        (23) KNNWTDAEKFCTEQKKGSHLVSLHSREEEEFVVNLISENLEYPATW
VLFXALC1       (46) KNNWTDAEKFCTEQNKGSHLVSLHSSEEEDFVVNLASQSLQYPVAW
TSV-FIX-BP     (44) KNWADAENFCTQQQAGGHLVSFQSSEEADFVVKLAFQTFDHSIFW
echicetin      (44) MNWADAEKFCTKQHKDGHLVSFRNSKEVDFVISLAFPMLKNDLVW

DRS-L2         (91) IGLGNMWKDCRMEWSDRGNVKYKALAEESYCLIMITHEKVWKSMT
RVVXLC1        (68) IGLGNMWKDCRMEWSDRGNVKYKALAEESYCLIMITHEKEWKSMT
VLFXALC1       (91) IGLGNMWKECRSEWSDGGNVKYKALAEESYCLLINTHKKGWRSMT
TSV-FIX-BP     (89) MGLSNVWNQCNWQWSNAAMLRYKAWAEESYCVYFKSTNNKWRSRS
echicetin      (89) IGLTDYWRDCNWEWSDGAQLDYKAWDNERHCFIYKNTDNQWTRRD

                                  Identity
DRS-L2        (136) CNFIAPVVCKF--  100.0 %
RVVXLC1       (113) CNFIAPVVCKF--   82.9 %
VLFXALC1      (136) CNNMAHVICKF--   79.5 %
TSV-FIX-BP    (134) CRMMANFVCEF--   54.1 %
echicetin     (134) CTWTFSFVCKCPA   45.9 %
```

Fig. 2　Comparison of the deduced amino acid sequences of DRS-L1 and DRS-L2 with those of other known snake venom C-type lectin-like proteins

(A)　The deduced amino acid sequence of DRS-L1 was aligned with the following sequences: VLF Ⅹ ALC2 (AAT91068) from *V. lebetina* venom, TMVA α subunit (AAM43808) from *T. mucrosquamatus* venom, stejnulxin α subunit (AAQ15166) from *T. stejnegri* venom and flavocetin-A α subunit (AAN72438) from *T. flavoviridis* venom.

(B)　The deduced amino acid sequence of DRS-L2 was aligned with the sequences as follows: RVV Ⅹ LC1 (AAB22478) from *D. russellii siamensis* venom, VLF Ⅹ ALC1 (AAQ17468) from *V. lebetina* venom, TSV-F Ⅸ-BP β subunit (AAQ15155) from *T. stejnegri* venom and echicetin β subunit (AAP41219) from *Echis carinutus* venom. Gaps have been inserted to maximize homology. Conserved cysteine residues are shaded. The extra cysteine residue is underlined. Identity percentage is shown on the right. The putative N-glycosylation site is boxed.

```
          DRS-L3    (1) MGRFISVSFGLLVVFLSLSGTEAAFC PSGWSAYDQN YKVFTEE
          DRS-L4    (1) MGRFISISFGLLVVFLSLSGTEAAFC PSGWSAYDQN YKVFTEE
          DRS-L5    (1) MGRFISISFGLLVMFLSLSGTGAKQD LSDWSFYEGY YKVFNEK
          DRS-L6    (1) MGRFISVSFGLLVVFLSLSGTGAD-- PSEWSSHEGH YKVFKLL
          DRS-L7    (1) MGRFISISFGLLVVFLSLSGTGAKQD LSDWSFYEGY YKVFNEK
    flavocetin-A    (1) MGQFIFVSFGFLVVATSLSGTEAGFC PLGWSSYDEH YQVFQQK
      stejnulxin    (1) MGQFIFVSFGLLVVLLSLSGAGAGFC PLGWSSYDLY YKVFKQQ
            TMVA    (1) MGRFIFVSFGLLVVFISLSGTEAGFC PLGWSSYDEH YQVFQQK
    agkisacutacin   (1) MGRFIFVSFGLLVVFLSLSGTAAD-- PSDWSSYEGH YKPFDEP
      TSV-FIX-BP    (1) MGRFIFVSFGLLVVFLSLSGTAAD-- LSGWSSYEGH YKPFNEL
       echicetin    (1) MGRFISVSFGLLVLLLSLSGTGAN-- LPDWSVYEGY YKVFKER
           EMS16    (1) MGRLISVRFSLLVVFLSLSGIGAGLC PLGWSSFDQH YKVFEPV
      botrocetin    (1) ------------------------D PPDWSSYEGH YRFFKEW

          DRS-L3   (46) MNWADAEKF TEQHKGSHLLSLHNIAEADFVLKKTLAMLKDGVIW
          DRS-L4   (46) MNWADAEKF TEQKKGSHLVSLHSREEEKFVVNLISENLEYPATW
          DRS-L5   (46) KTWEDAEKF TEQHKGSHLLSLHNIAEADFVLKKTLAMLKDGVIW
          DRS-L6   (44) KTWEDAEKF TEQKKGSHLVSLHSREEEKFVVNLISENLEYPATW
          DRS-L7   (46) KTWEDAEKF NEQVNGGYLVSFR**SSE**EM**D**FVIRMTFPIFRFDFFW
    flavocetin-A   (46) MNWEDAEKF TQQHKGSHLVSFH**SSE**EV**D**FVTSKTFPILKYDFVW
      stejnulxin   (46) MNWTDAEKF TEQHTGSHLVSFHSSEEADFVVNMTYPILKLDFVW
            TMVA   (46) MNWEDAEKF TQQHTGSHLVSYE**SSE**EV**D**FVVSKTLPILKASFVW
    agkisacutacin  (44) KTWADAEKF TQQHKGSHLASFHSSEEADFVVTLTTPSLKTDLVW
      TSV-FIX-BP   (44) KNWADAENF TQQQAGGHLVSFQSSEEADFVVKLAFQTFDHSIFW
       echicetin   (44) MNWADAEKF TKQHKDGHLVSFRN**S**KEV**D**FVISLAFPMLKNDLVW
           EMS16   (46) KNWTEAEEI MQQHKGSRLASIHGSEEEAFVSKLASKALKFTSMW
      botrocetin   (21) MHWDDAEEF TEQQTGAHLVSFQSKEEADFVRSLTSEMLKGDVVW

          DRS-L3   (91) MGLNDVWNE NWGWTDGAKLDYKAWN--EGTN FVFKIAKN--HW
          DRS-L4   (91) IGLGNMWKD RMEWSDRGNVKYKALA--EESY LIMITHEK--VW
          DRS-L5   (91) MGLNDVWNE NWGWTDGAKLDYKAWN--EGTN FVFKIAKN--HW
          DRS-L6   (89) IGLGNMWKD RMEWSDRGNVKYKALA--EESY LIMITHEK--VW
          DRS-L7   (91) IGLRDFWRD YWRWSDGVNLDYKAWS--REPN FVS**KTTDN**--QW
    flavocetin-A   (91) IGLSNVWNE TKEWSDGTKLDYKAWS--GGSD IVS**KTTDN**--QW
      stejnulxin   (91) IGLSNVWNQ NSEWSDGTKLDYKDWS--GESE IASKTVEN--QW
            TMVA   (91) IGLSNVWNA RLQWSDGTELMYNAWT--AESE IAS**KTTDN**--QW
    agkisacutacin  (89) IGLKNIWNG YWKWSDGTKLDYKDWR--EQFE LVSRTVNN--EW
      TSV-FIX-BP   (89) MGLSNVWNQ NWQWSNAAMLRYKAWA--EESY VYFKSTNN--KW
       echicetin   (89) IGLTDYWRD NWEWSDGAQLDYKAWD--NERH FIY**KNTDN**--QW
           EMS16   (91) IGLNNPWKD KWEWSDNARFDYKAWK--RRPY TVMVVKPDRIFW
      botrocetin   (66) IGLSDVWNK RFEWTDGMEFDYDDYYLIAEYE VASKPTNN--KW

          DRS-L3  (132) SHMD SSTHNFV KFRV-----
          DRS-L4  (132) KSMT NFIAPVV KF-------
          DRS-L5  (132) SHMD SSTHNFV KFRV-----
          DRS-L6  (130) KSMT NFIAPVV KF-------
          DRS-L7  (132) LRWN NDPRYFV KSRVSC---
    flavocetin-A  (132) LSMD SSKRYVV KFQA-----
      stejnulxin  (132) WTKS SRTHYVV KFQA-----
            TMVA  (132) WSMD SSKRYVV KF-------
    agkisacutacin (130) LSMD GTTCSFV KFQA-----
      TSV-FIX-BP  (130) RSRS RMMANFV EF-------
       echicetin  (130) TRRD TWTFSFV KCPA-----
           EMS16  (134) FTRG EKSVSFV KFLTDPAV-
      botrocetin  (109) WIIP TRFKNFV EFQA-----
```

Fig. 3 Comparison of the deduced amino acid sequences of DRS-L3, DRS-L4, DRS-L5, DRS-L6 and DRS-L7 with those of other known snake venom C-type lectin-like protein β subunits

The sequences are: flavocetin-A β subunit (AAN72437) from *T. flavoviridis* venom, stejnulxin β subunit (AAQ15168) from *T. stejnegri* venom, TMVA β subunit (AAM43809) from *T. mucrosquamatus* venom, agkisacutacin β subunit (AAF26287) from *Deinagkistrodon acutus* venom, TSV-FⅨ-BP β subunit (AAQ15155) from *T. stejnegri* venom,

echicetin β subunit (AAP41219) from *E. carinutus* venom, EMS16 β subunit (BAC77707) from *E. multisquamatus* venom and botrocetin β subunit (P22030) from *Bothrops jararaca* venom. Gaps have been inserted to maximize homology. Conserved cysteine residues are shaded. The extra cysteine residue is underlined. The two hydrophilic patch Ⅰ and Ⅱ in GPIb-BPs are indicated in bold.

patch Ⅱ by Lys-102, Thr-104, Asp-105, Asn-106 and Gln-107 in the β subunit (flavocetin-A numbering system) (Fukuda et al, 2000). When comparing the deduced protein sequence of DRS-L7 with those of other known GPIb-BP β subunits, it can be found that the hydrophilic patch Ⅰ and Ⅱ are well conserved among DRS-L7 and other known snake venom GPIb-BPs like flavocetin and TMVA (Fig. 3). Moreover, DRS-L7 has the highest identity with echicetin (62.0%). So it is presumed that DRS-L7 is probably the β subunit of a platelet membrane GPIb-BP expressed in *D. russellii siamensis* venom gland.

In this work, seven cDNAs encoding various SV-CTLP subunits from *D. russellii siamensis* venom gland were obtained. The biological effects of these proteins in snake venom need to be clarified further. Our results provide the information for studying the structure-function relationships of SV-CTLPs.

References:

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs[J]. *Nucleic Acids Res.*, **25** (17): 3389－3402.

de Lima DC, Abreu PA, de Freitas CC, Santos DO, Borges RO, dos Santos TC, Cabral LM, Rodrigues CR, Castro HC. 2005. Snake venom: Any clue for antibiotics and CAM[J]. *eCAM*, **2**(1): 39－47.

Dodd RB, Drickamer K. 2001. Lectin-like proteins in model organisms: Implications for evolution of carbohydrate-binding activity[J]. *Glycobiology*, **11**(5): 71R－79R.

Drickamer K. 1992. Engineering galactose-binding activity into a C-type mannose-binding protein[J]. *Nature*, **360**(6400): 183－186.

Drickamer K. 1999. C-type lectin-like domains[J]. *Curr. Opin. Struct. Biol.*, **9**(5):585－590.

Fukuda K, Mizuno H, Atoda H, Morita T. 2000. Crystal structure of flavocetin-A, a platelet glycoprotein Ib-binding protein, reveals a novel cyclic tetramer of C-type lectin-like heterodimers[J]. *Biochemistry*, **39**(8): 1915－1923.

Hirabayashi J, Kusunoki T, Kasai K. 1991. Complete primary structure of a galactose-specific lectin from the venom of the rattlesnake *Crotalus atrox*: Homologies with Ca^{2+}-dependent-type lectins[J]. *J. Biol. Chem.*, **266**(4): 2320－2326.

Jasti J, Paramasivam M, Srinivasan A, Singh TP. 2004. Crystal structure of echicetin from *Echis carinatus* (Indian saw-scaled viper) at 2.4 A° resolution[J]. *J. Mol. Biol.*, **335**(1): 167－176.

Lee WH, Zhang Y. 2003. Research development of snake venom C-type lectins[J]. *Zool. Res.*, **24**(2): 151－160.[李文辉,张 云. 2003. 蛇毒C－型凝集素研究进展. 动物学研究,**24**(2):151－160.]

Lee WH, Du XY, Lu QM, Clemetson KJ, Zhang Y. 2003. Stejnulxin, a novel snake C-type lectin-like protein from *Trimeresurus stejnegeri* venom is a potent platelet agonist acting specifically via GPVI[J]. *Thromb. Haemost.*, **90**(4): 662－671.

Li R, Wei Q, Wang WY, Xiong YL. 2003. Cloning of cDNA encoding a C-type lectin from snake *Trimeresurus jerdonii*[J]. *Zool. Res.*, **24**(4): 293－296.[李 瑞,魏 勤,王婉瑜,熊郁良. 2003. 菜花烙铁头蛇毒C－型凝集素基因的克隆与序列分析. 动物学研究,**24**(4):293－296.]

Li YS, Liu KF, Wang QC, Ran YL, Tu GC. 1985. A platelet function inhibitor purified from *Vipera russelli siamensis* (Smith) snake venom [J]. *Toxicon*, **23**(6): 895－903.

Mizuno H, Fujimoto Z, Koizumi M, Kano H, Atoda H, Morita T. 1997. Structure of coagulation factors Ⅸ/Ⅹ-binding protein, a heterodimer of C-type lectin domains[J]. *Nat. Struct. Biol.*, **4**(6): 438－441.

Ogawa T, Chijiwa T, Oda-Ueda N, Ohno M. 2005. Molecular diversity and accelerated evolution of C-type lectin-like proteins from snake venom[J]. *Toxicon*, **45**(1): 1－14.

Pearson WR, Lipman DJ. 1988. Improved tools for biological sequence comparison[J]. *Proc. Natl. Acad. Sci. USA*, **85**(8): 2444－2448.

Peng M, Holt JC, Niewiarowski S. 1994. Isolation, characterization and amino acid sequence of echicetin beta subunit, a specific inhibitor of von Willebrand factor and thrombin interaction with glycoprotein Ib[J]. *Biochem. Biophys. Res. Commun.*, **205**(1): 68－72.

Prasad NB, Uma B, Bhatt SKG, Gowda VT. 1999. Comparative characterization of Russell's viper (Daboia/Vipera russelli) venoms from different regions of the Indian peninsula[J]. *Biochim. Biophys. Acta*, **1428**: 121－136.

Shin Y, Okuyama I, Hasegawa J, Morita T. 2000. Molecular cloning of glycoprotein Ib-binding protein, flavocetin-A, which inhibits platelet aggregation[J]. *Thromb. Res.*, **99**(3): 239－247.

Siigur E, Aaspollu A, Trummal K, Tonismagi K, Tammiste I, Kalkkinen N, Siigur J. 2004. Factor Ⅹ activator from Vipera lebetina venom is synthesized from different genes[J]. *Biochim. Biophy. Acta*, **1702** (1): 41－51.

Takeya H, Nishida S, Miyata T, Kawada S, Saisaka Y, Morita T, Iwanaga S. 1992. Coagulation factor Ⅹ activating enzyme from Russell's Viper venom (RVV-Ⅹ). A novel metalloptroteinase with disintegrin (platelet aggregation inhibitor)-like and C-type lectin-like domains[J]. *J. Biol. Chem.*, **267**(20): 14109－14117.

Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice[J]. *Nucleic Acids Res.*, **22**(22): 4673－4680.

Wei Q, Lu QM, Jin Y, Li R, Wei JF, Wang, WY, Xiong YL. 2002. Purification and cloning of a novel C-type lectin-like protein with platelet aggregation activity from *Trimeresurus mucrosquamatus* venom[J]. *Toxicon*, **40**(9): 1331－1338.

Weis WI, Taylor ME, Drickamer K. 1998. The C-type lectin superfamily in the immune system[J]. *Immunol. Rev.*, **163**: 19－34.

Wuster W, Golay P, Warrell DA. 1997. Synopsis of recent developments in venomous snake systematics[J]. *Toxicon*, **35**(3): 319－340.

Xu Q, Wu XF, Xia QC, Wang KY. 1999. Cloning of a galactose-binding lectin from the venom of *Trimeresurus stejnegeri*[J]. *Biochem. J.*, **341**(Pt3): 733－737.

Zha HG, Lee WH, Zhang Y. 2001. Cloning of cDNAs encoding C-type lectins from Elapidae snakes *Bungarus fasciatus* and *Bungarus multicinctus*[J]. *Toxicon*, **39**(12): 1887－1892.