Serine protease isoforms of *Deinagkistrodon acutus* venom: cloning, sequencing and phylogenetic analysis

Ying-Ming WANG*, Suei-Rong WANG† and Inn-Ho TSAI*†¹

*Institute of Biological Chemistry, Academia Sinica, P.O. Box 23-106, Taipei, Taiwan, and †Institute of Biochemical Science, National Taiwan University, P.O. Box 23-106, Taipei, Taiwan

The major coagulating fibrinogenase of *Deinagkistrodon acutus* venom, designated acutobin, was purified by anion-exchange chromatography, gel filtration and reverse-phase HPLC. Approximately 80% of its protein sequence was determined by sequencing the various fragments derived from CNBr cleavage and digestion with endoprotease. Extensive screening of the venom gland cDNA species after amplification by PCR resulted in the isolation of four distinct cDNA clones encoding acutobin and three other serine proteases, designated Dav-PA, Dav-KN and Dav-X. The complete amino acid sequences of these enzymes were deduced from the cDNA sequences. The amino-acid sequence of acutobin contains a single chain of 236

residues including four potential N-glycosylation sites. The purified acutobin (40 kDa) contains approx. 30% carbohydrate by weight, which could be partly removed by N-glycanase. The phylogenetic tree of the complete amino acid sequences of 40 serine proteases from 18 species of Crotalinae shows functional clusters reflecting parallel evolution of the three major venom enzyme subtypes: coagulating enzymes, kininogenases and plasminogen activators. The possible structural elements responsible for the functional specificity of each subtype are discussed.

Key words: cDNA cloning, complete sequence, thrombin-like enzyme.

INTRODUCTION

Viperid snake venom components are known to affect the haemostatic system by a variety of mechanisms. One group of therapeutically useful venom proteins is the serine proteases. They are widely distributed in the venoms of true vipers (Viperinae) and pit vipers (Crotalinae) and act on different steps of the blood coagulation cascade [1–4]. Some of them can hydrolyse fibrinogen specifically and release fibrinopeptide A or fibrinopeptide B or both [3], and some of them affect other substrates, e.g. kininogen [5,6], plasminogen [7] and protein C [8]. Apparently, variants of this family of venom proteases have undergone gene duplication and accelerated evolution to acquire special functions [9]. Some of the venom proteases have been used as anti-thrombotics to facilitate tissue oxygenation and prevent arterial embolism [10]. To improve their therapeutic use, the structure–activity relationships and the diversity of these venom proteases should be investigated.

Deinagkistrodon acutus (formerly *Agkistrodon acutus*, or hundred-pace snake) is a monotypic pit viper inhabiting Taiwan and South China. The major serine protease in the venom is a clotting enzyme that we now designate acutobin, although it has previously been named acutin [11], acuthrombin [12], acutase or defibrinum. Partly purified acutobin has been used clinically on many patients in China [13]. Another bradykinin-releasing serine protease from this venom species has been reported [14]. However, the structures of these enzymes have not been unequivocally solved. The present study aims at determining the primary structures of *D. acutus* venom serine proteases and investigating their structure–function relationships.

EXPERIMENTAL

Materials

Pooled venom powder of *D. acutus* (Taiwan) and the fresh venom gland of a live specimen of *D. acutus* (Fujian, China) were gifts from Dr M. Y. Liao (Institute of Disease Control, Taipei, Taiwan). Modification and restriction enzymes and the ‘pGEM-T easy’ vector were from Promega (Madison, WI, U.S.A.). Benzoyl-Pro-Phe-Arg-pNA (in which pNA stands for *p*-nitro-anilide), tosyl-Gly-Pro-Arg-pNA, benzoyl-Phe-Val-Arg-pNA and *p*-nitrophenyl *p*′-guanidinobenzoate were from Sigma Chemical Co. (St Louis, MO, U.S.A.). Acrylamide/bisacrylamide solution and other reagents for gel electrophoresis were from Bio-Rad Laboratories (Hercules, CA, U.S.A.). N-glycosidase F and endopeptidases Lys-C and Arg-C were from Boehringer Mannheim (Mannheim, Germany). Buffers and other chemicals were of reagent grade.

Purification of venom serine proteases

D. acutus venom powder was dissolved in 5 mM ammonium acetate and centrifuged. The supernatant was applied to a DEAE-Trisacryl ion-exchange column and eluted with a linear gradient of ammonium acetate from 5 to 320 mM. Each fraction was assayed for amidolytic activities towards the chromogenic substrate. The active fractions (designated acutobin) were pooled, dialysed, freeze-dried and further purified with a gel-filtration column (TSK G2000SW, 0.75 cm × 30 cm) equilibrated and eluted with 0.15 M ammonium acetate. Alternatively, acutobin was further purified by HPLC on a Vydac RP-C₄ column (4.6 mm × 250 mm).

Abbreviation used: pNA; *p*-nitroanilide.

¹ To whom correspondence should be addressed at P.O. Box 23-106, Taipei, Taiwan 10798 (e-mail bc201@gate.sinica.edu.tw).

² The nucleotide sequence data reported will appear in the DDBJ, EMBL and GenBank Nucleotide Sequence Databases under the accession numbers AF159057 (acutobin), AF159058 (Dav-PA), AF159059 (Dav-KN) and AF159060 (Dav-X).

Figure 1 DEAE-Trisacryl anion-exchange purification of crude venom

Crude venom (150 mg) of *D. acutus* in 5 mM ammonium acetate was loaded on a DEAE-Trisacryl column pre-equilibrated with the same buffer. The proteins were eluted with a gradient (— · — · —) of 0.005–0.32 M ammonium acetate, pH 7.2, at 40 ml/h. Fractions of 3 ml were collected. The column can separate up to 6 peaks of amidolytic activities from either the pooled crude venom (upper panel) or venom of a single specimen (lower panel).

Deglycosylation of the protease

Deglycosylation of acutobin with 1 unit of recombinant N-glycosidase F was performed in 50 μ l of 40 mM sodium phosphate (pH 7.2) containing 0.05% SDS, 1% (v/v) 2-mercaptoethanol and 2% (v/v) octylglucoside at 37 °C for 16 h. The molecular mass and purity of the venom protein with or without deglycosylation were analysed by SDS/PAGE [12.5% (w/v) gel].

Reductive alkylation, fragmentation and protein sequencing

Approximately 300 μ g (7.5 nmol) of acutobin was reduced for 2 h with 375 mM dithiothreitol in 6 M guanidinium chloride/0.25 M Tris/HCl (pH 8.5)/1 mM EDTA at 50 °C and subsequently S-carboxymethylated. The alkylated protein was desalted with a gel-filtration column (TSK G2000SW) and then cleaved for 18 h by CNBr in 70% (v/v) formic acid at 25 °C [15], or digested for 16 h with either endopeptidase Lys-C or endopeptidase Arg-C in 10 mM Tris/HCl (pH 8.0) at 37 °C and an enzyme-to-substrate ratio of 1:100 (w/w). The resultant peptides were purified by reverse-phase HPLC on a Vydac C_{18} column (4.6 mm $\times$ 150 mm) with a gradient of acetonitrile in 0.07% (v/v) trifluoroacetic acid. The N-terminal amino acid sequences of the enzyme and its fragments were determined by a sequencer (ABI 477A; Perkin-Elmer, Foster City, CA, U.S.A.) [16].

Enzyme specificity towards synthetic substrates and fibrinogen

The purified enzyme was titrated with *p*-nitrophenyl *p*'-guanidinobenzoate [17]. Tosyl-Gly-Pro-Arg-pNA, benzoyl-Pro-Phe-Arg-pNA and benzoyl-Phe-Val-Arg-pNA were dissolved in acetonitrile separately to make stock solutions. The reaction

mixture contained 690 μ l of 0.1 M Tris/HCl, pH 7.6, and 10 μ l of substrate stock; venom enzyme was added to start the reaction. The hydrolysis was followed at 25 °C by measuring the absorbance of *p*-nitroaniline at 405 nm [18]. The specificity constants, k_{cat}/K_m , were calculated from the pseudo-first-order rate constants divided by the active enzyme concentration. The clotting activity of the enzyme was assayed with human fibrinogen solution [13,14].

Isolation of cDNA clones and sequencing

The mRNA and the complement cDNA were prepared from the venom glands of a *D. acutus* specimen from Fujian Province, China, using mRNA extraction and the cDNA synthesis kits in accordance with the manufacturer's instructions (Stratagene, La Jolla, CA, U.S.A.). PCR was conducted to amplify the cDNA species of the venom serine proteases. Two primers were designed on the basis of the highly conserved cDNA sequences encoding TLf1 [9], calobin [19], batroxobin [20] and ancrod [21].

Primers 1 and 2 were 5'-ACGCAGAGTTGAAGCTATGGT-3' and 5'-GCAGTCCTATTTGAGTCTA-3' respectively. Primer 1 was in the sense orientation of the 5' end from the untranslated region to the first two amino acids of the signal peptide; primer 2 was in the anti-sense direction of a conserved region at the 3' end of the untranslated region. Total cDNA (as templates) and SuperTaq DNA polymerase (HT Biotech, Cambridge, U.K.) were used. As shown by 1% (w/v) agarose-gel electrophoresis, a 0.9 kb DNA fragment was specifically amplified. After being treated with polynucleotide kinase, it was inserted into the 'pGEM-T easy' vector and transformed into *Escherichia coli* strain JM109. White transformants were picked up and

sequenced on the DNA-Sequencing System (model 373A; PE Applied Biosystems, Foster City, CA, U.S.A.).

Phylogenetic tree of serine proteases based on protein sequences

Amino acid sequences of the serine proteases were aligned by using the PILEUP program implemented in the Genetics Computer Group (GCG). A cladistic analysis of sequence data was performed with PHYLIP [22]. The distance matrix for the aligned sequences was calculated with the two-parameter method of Kimura; this distance matrix was then used to reconstruct a phylogenetic tree by the neighbour-joining method [23]. Bootstrap estimates of the confidence levels were conducted in 1000 replicates with a program based on the complete-and-partial bootstrap technique [24].

RESULTS

Purification of acutobin

Because *D. acutus* venom contains both metalloproteases and serine proteases, but since only the latter hydrolyse pNA, we used tosyl-Gly-Pro-Arg-pNA as the substrate to follow the purification of the serine proteases from the crude venom. The venom enzymes that released pNA were fractionated with a

DEAE-Trisacryl column into six partly resolved peaks (Figure 1, upper panel), which were eluted at approx. 0.13–0.18 M ammonium acetate. Venom from a single specimen of *D. acutus* was also fractionated in parallel with the pooled venom and similar results were obtained (Figure 1, lower panel). The purified acutobin showed some microheterogeneities that affected its pI and column affinity. Each amidolytic peak was further purified with a C_4 column and was eluted at approx. 45% acetonitrile (Figure 2). They were analysed on a sequencer and found to have identical N-terminal sequences of VIGGVECDINEHRFL-VALYE/ (single-letter amino acid code). Part of the isolated acutobin glycoproteins was further purified with a gel-filtration column (TSK G2000SW) (Figure 3). The purification achieved was 18-fold in comparison with the specific activity of the crude venom of *D. acutus*; the yield was 1.6% (w/w). We were unable to isolate serine proteases with N-terminal sequences identical with those of Dav-KN and Dav-PA (see below).

Characterization of the purified enzyme

Purified acutobin coagulated human fibrinogen by hydrolysis of the A α chains but not the B β chains (results not shown), which is consistent with previous findings [25,26]. The molecular masses of native and N-glycosidase-treated acutobin were studied by SDS/PAGE (Figure 3, inset). Purified acutobin showed a single

Figure 2 Purification of the protease peaks by reverse-phase HPLC

The serine protease fractions (A–F) from Figure 1 were further purified by HPLC on a Vydac RP- C_4 column (4.6 mm $\times$ 250 mm). Elution buffers A and B were distilled water and 0.07% (v/v) trifluoroacetic acid in acetonitrile respectively. The elution was performed at 25% buffer B for 5 min and then 25–65% buffer B for 40 min at a flow rate of 1 ml/min. The asterisk denotes the serine protease peak.

Figure 3 Purification of acutobin by gel filtration and SDS/PAGE

The major protease fractions from DEAE-Trisacryl anion-exchange chromatography (upper panel, SP) were pooled, freeze-dried and further purified by gel filtration on a TSK G2000SW column in 0.15 M ammonium acetate, pH 6.5 (lower panel). The asterisk denotes the peak containing acutobin. Inset: SDS/PAGE of acutobin before (lane 1) and after (lane 2) treatment with N-glycosidase F. The positions of molecular mass markers are indicated at the left.

protein band at 41 kDa and the glycosidase-treated form was at 35 kDa. The kinetic specificity constants of acutobin towards the following synthetic chromogenic substrates were measured at pH 7.6 and 25 °C with $K_m > [S]$. The k_{cat}/K_m values for tosyl-Gly-Pro-Arg-pNA, benzoyl-Phe-Val-Arg-pNA and benzoyl-Pro-Phe-Arg-pNA were 1.80×10^5 , 1.46×10^3 , and $3.93 \times 10^3 \text{ M}^{-1} \cdot \text{s}^{-1}$ respectively.

Protein sequence of acutobin

The N-terminal amino acid sequence of the native protein was determined up to the 32nd residue. Fragments from the CNBr or

Figure 4 Fragmentation and sequencing of acutobin

Reduced and alkylated acutobin was subjected to digestions with CNBr and the endopeptidases Lys-C and Arg-C. The amino acid sequences of the HPLC-purified fragments were determined and are shown below the cDNA-deduced acutobin sequence. Fragments are abbreviated by their digesting agents: CNBr (CB), Lys-C (K) and Arg-C (R).

Figure 5 Alignment of the amino acid sequences of Crotaline venom serine proteases

The four *D. acutus* serine protease sequences were deduced from those of their venom-gland cDNA species. Sequences of batroxobin, ancrod, crotalase [20,21,28] and protein C activator from *A. contortrix* venoms (ACC-C) [29] are also listed for comparison. Chymotrypsinogen numbering is followed, and gaps (indicated by dashes) have been introduced to improve the alignment. Residues identical with those in the first row are shown by dots.

endopeptidase treatments of reduced and alkylated acutobin were purified by reverse-phase HPLC. We obtained a total of 16 purified polypeptides, including 4 from CNBr cleavage, 2 from Lys-C digestion and 10 from Arg-C digestion. The amino acid sequences of the overlapping fragments were determined; up to 185 amino acid residues, or 77% of acutobin, were sequenced (Figure 4).

Figure 6 Phylogenetic tree for snake venom serine proteases

The computer program PHYLIP and all the amino acid sequences of viperid venom serine proteases so far available were used to construct the tree. In addition to those shown in [11,37] and Figure 5, the enzyme abbreviations and references are as follows: Ghb-p1 and Ghb-p2 (Ghb-palladin 1 and 2) and Ah-Pallas (AF088996) from *Agkistrodon halys* Pallas [49]; Tm1–Tm5 from *Trimeresurus mucrosquamatus* [40]; TLF2 and TLF3 from *Trimeresurus flavoviridis*; TLG1, TLG2A, TLG2b, TLG2C and TLG3 from *Trimeresurus gramineus* [9]; Gh-1 (AF015727), Gh-2 (AF056033), Ghb-HALY2 and Haly-PA from *Gloydius halys breviceaudus* [39]; GU-CPI from *Gloydius ussuriensis* [50]; KN-BJ2 from *Bothrop jararaca* [37]; TSV-PA from *Trimeresurus stejnegeri* [7]; and halystase from *Agkistrodon halys bloomhoffii* [38]. RVV-Va, the factor V activator from Russell's viper, was used as an outgroup. Abbreviations for the subtypes: CL, coagulating enzyme; PA, plasminogen activator; KN, kininogenase. The numbers on the branches indicate the proportions of the bootstrap replicates, and only bootstrap values higher than 50% are shown. The scale at the bottom left denotes a branch length of 10^5 years.

Cloning and nucleotide sequencing of the protease cDNA species

After PCR amplification and cloning, the nucleotide sequences of up to 40 of the cDNA clones were determined; most of them encoded acutobin. We identified not only cDNA species for acutobin but also those encoding the other three distinct serine proteases, which were named Dav-KN, Dav-PA and Dav-X. Some truncated cDNA species with a deletion of 243 bp (corresponding to residues 68–147 of acutobin) were also cloned. All of the protease precursors contained a potential signal peptide of 24 residues followed by a protease domain of 235 ± 2 residues.

The complete amino acid sequences were predicted from the cDNA nucleotide sequences and are aligned with those of other venom serine proteases in Figure 5.

Phylogenetic study on 40 serine proteases

The relationships between 40 serine proteases from crotaline snakes were studied by constructing a phylogram with the program PHYLIP on the basis of their complete amino acid sequences. The sequence of factor V activator from Russell's viper venom, RVV-Va, was used as outgroup in the phylogenetic analysis. A single neighbour-joining tree was recovered (Figure 6). In this tree, three clusters of serine proteases were separated; they are designated CL (coagulating enzyme), KN (kininogenase) and PA (plasminogen activator). The nodes with bootstrap values of more than 0.70 on the tree are supported with more than 95% probability [27].

DISCUSSION

Venom serine proteases of *D. acutus* have been studied for 20 years [25,26]. The intrinsic heterogeneity associated with glycosylation of the enzymes and possible contamination with other venom proteins had caused difficulties in protein analyses in previous studies. The roles of the glycans in venom proteases remain to be investigated but in some cases the glycans seem to be needed for protein stabilization rather than for the catalytic function of the venom enzymes [6].

After the major serine protease (acutobin) had been purified, 77% of its amino acid sequence was solved by protein sequencing (Figure 4). The cDNA-deduced amino acid sequence of acutobin matched those determined for the purified acutobin fragments apart from three differences (positions 77, 174 and 229). The discrepancies probably resulted from glycosylation at specific Asn residues and geographical variation between the venom protein and the venom glands used. Four potential glycosylation sites were identified at Asn-77, Asn-81, Asn-100 and Asn-229 (Figure 4).

In addition to acutobin, the other three *D. acutus* serine proteases were also cloned. The predicted pI values and molecular masses of the four *D. acutus* proteases before glycosylation are as follows: acutobin, pI 5.7, 236 residues, 26 104 Da; Dav-X, pI 7.2, 236 residues, 26 500 Da; Dav-KN, pI 5.0, 236 residues, 25 623 Da; Dav-PA, pI 5.5, 234 residues, 25 321 Da. The sequences of Dav-X, Dav-PA and Dav-KN contain four, one and no potential N-glycosylation sites respectively. Only one of the N-glycosylation sites in Dav-X is identical with that in acutobin. All four *D. acutus* serine proteases share amino acid sequence identities of approx. 65% with one another and to the thrombin-like enzymes [20,21,28,29] from other venoms (Figure 5).

Compared with other members of the serine protease family, the presumed catalytic triad of the active site in serine proteases (His-57, Asp-102 and Ser-195; α -chymotrypsin numbering scheme) and their flanking sequences are highly conserved in the venom enzymes. It is noteworthy that all of them contain Asp-194, which is responsible for their selectivity for Arg at the P1 sites of the substrates [30], and 12 cysteine residues at conserved positions, except that Dav-KN contains a mutation of Cys-201 to Ser-201 (Figure 5). Their disulphide bridges are predicted to be Cys-22–Cys-157, Cys-42–Cys-58, Cys-91–Cys-250, Cys-136–Cys-201, Cys-168–Cys-182 and Cys-191–Cys-220 on the basis of the crystallographic structure of TSV-PA [31]. Assuming an essential role for Cys-136–Cys-201 in the stabilization of the enzyme structure, the Dav-KN mRNA probably represents a pseudogene.

Table 1 Amino acid sequences at the putative functional sites (residues 82–101 and 191–193) of representative venom serine proteases

Chymotrypsinogen numbering is used. Abbreviations of enzymes are as in Figure 6. Residues relatively conserved in all the venom proteases are boxed, and residues unique for the subtype are emboldened.

Type	Venom protease	82											90				96 97				100				191			
CL	Acutobin	G	R	E	P	I	E	K	Y	F	Y	N	C	-	S	-	N	N	L	T	T	R	D	K	C	N	R	
	TLg2b	G	R	H	P	K	E	K	Y	I	F	N	C	-	S	-	N	N	F	T	K	W	D	K	C	N	R	
	Bilineobin	R	R	Y	P	K	E	K	Y	L	F	R	C	-	S	-	K	N	F	T	K	W	D	K	C	K	R	
	Crotalase	R	R	F	P	K	E	K	Y	F	F	D	C	-	S	-	N	N	F	T	K	W	D	K	C	N	R	
	Lmm-TL	R	R	Y	P	K	K	K	Y	F	F	R	C	-	N	-	K	N	F	T	K	W	D	E	C	N	R	
	Ancrod	E	R	Y	P	K	K	R	Y	F	I	R	C	-	N	-	K	T	R	T	S	W	D	E	C	N	S	
	Dav-X	R	R	Y	A	I	E	E	H	Y	Y	R	C	-	D	-	E	K	L	T	K	W	E	K	C	H	Q	
KN	TLF2	T	R	V	P	K	E	K	-	F	F	-	C	L	S	S	K	T	H	T	R	W	N	K	C	K	A	
	Calobin	T	R	V	P	K	E	K	-	F	F	-	C	L	S	S	K	N	Y	T	L	W	D	K	C	R	G	
	Halystase	T	R	V	P	K	E	K	-	F	F	-	C	L	S	S	K	N	Y	T	L	W	D	K	C	R	G	
	KN-BJ2	T	R	V	A	K	E	K	-	F	F	-	C	L	S	S	K	N	Y	T	K	W	D	K	C	V	G	
	Tm1~Tm4	R	R	V	P	K	E	K	-	F	F	-	C	D	S	S	K	T	Y	T	K	W	N	K	C	V	G	
	Dav-KN	R	R	V	P	K	E	K	-	F	F	-	C	D	S	S	K	T	Y	T	K	W	N	K	C	G	F	
PA	PA-BJ	T	R	H	P	K	E	K	-	F	I	-	C	P	N	R	K	K	D	D	V	L	D	K	C	E	G	
	Ghb-Haly-PA	T	R	D	P	K	E	K	-	F	I	-	C	P	N	K	K	K	D	D	E	K	D	K	C	V	G	
	TSV-PA	T	R	D	P	K	E	K	-	F	F	-	C	P	N	R	K	K	D	D	E	V	D	K	C	H	F	
	Dav-PA	T	R	N	P	K	E	K	-	F	I	-	C	P	N	K	N	N	N	E	V	L	D	K	C	G	G	
	Flavoxobin	I	R	N	P	K	E	K	-	F	I	-	C	P	N	K	K	N	D	E	V	L	D	K	C	G	F	
	TLG1	I	R	N	P	K	E	K	-	F	I	-	C	P	N	K	K	N	N	E	V	L	D	K	C	G	G	

The sequence identity between the cDNA species encoding acutobin and Dav-PA is 75% for the mature protein-coding region but 98% for the signal peptide-coding region and 95% for the 3' untranslated region. Thus the latter two regions are highly conserved, which explains why a common pair of primers could be used in the PCR experiments successfully. This is in accord with previous findings of positive Darwinian evolution for other venom protein families [9,32]. It might be beneficial for the snakes to transcribe an existing venom gene at a low level even though the gene product is temporarily not translated at a high level. The existence of multiple mRNA species of low expression that encode venom protein variants (e.g. Dav-PA or Dav-X) might permit the efficient and fast adaptation of the snakes to changing environments.

The present study unequivocally solved the protein sequence of the major thrombin-like protease in the venom of *D. acutus*. The recently published sequence of 'acutin' (the thrombin-like enzyme from *D. acutus* venom; Jiangxi, China) was apparently a wrong match because the experimentally determined N-terminal sequence was not consistent with its cDNA-deduced sequence [11]. This sequence was only 65% identical with that of acutobin (Figure 4) but 97% identical with that of Dav-PA (Figure 5). The partial sequence of a kininogenase [14] and the complete sequences of agkisacutacin [33] and acutothrombin A-C [34,35] from *D. acutus* venom have been reported. It is noteworthy that these sequences are similar to those of the C-type lectin-like venom proteins. It is doubtful that the non-catalytic lectin family could have evolved into proteases, especially since the venoms had already been equipped with both serine proteases and metalloproteases. Without careful purification and quantification of the enzyme by active-site titration, these *D. acutus* proteins might have been heavily contaminated with the C-type lectin-like venom proteins.

The phylogenetic tree (Figure 6) reveals the independent evolution of three major enzyme subtypes, the coagulating enzymes (CL), the plasminogen activators (PA) and the kininogenases (KN). Deshimaru et al. [9] used the mature protein-coding DNA sequences of eight serine-protease cDNA sequences

from two Asian pit vipers to construct a tree that also revealed parallel evolution of several enzyme subtypes. In the venom of some pit vipers, e.g. *D. acutus*, *Trimeresurus flavoviridis* and *Trimeresurus gramineus*, all three major subtypes of serine proteases were found. The classification of the three functional subtypes might find support from the literature, as follows. (1) Most of the well-known venom-clotting enzymes that convert fibrinogen into a fibrin clot, e.g. batroxobin, ancrod, crotalase and acutobin, are linked together in the CL subtype and are characterized by being A α fibrinogens and highly glycosylated. (2) More than half of the enzymes of the KN subtype had been characterized as weakly clotting or non-clotting, but might cleave B β or both the A α and B β fibrinogen chains. Many of them have been characterized as being capable of cleaving kininogen to produce bradykinin [36–38], leading to a hypotensive effect. (3) At least two of the enzymes (TSV-PA and Haly-PA) of the PA subtype have been confirmed to have the role of a plasminogen activator and to facilitate the solubilization of fibrin clots, as their names imply [7,39]. Members of the KN and the PA subtypes usually are less or not glycosylated, so their molecular masses are lower than those of the CL subtype.

It should be pointed out that the complete sequences of serine proteases from Viperinae venoms are lacking, so the tree (Figure 6) contains only the enzymes from Crotalinae venoms. Further separations of the three subtypes might become possible if more sequence data could be used. For example, the KN subtypes might be further divided into two groups, one represented by halystase [38] and calobin [19], which cleave both A α and B β fibrinogen, and the other represented by *Trimeresurus mucrosquamatus* Tm1–5 [40], which cleave only B β fibrinogen. It was reported that both thrombin, a coagulating enzyme, can also cleave bradykinin [38]. It seems that both thrombin belongs to both the CL and KN subtypes. However, the functions of several enzymes of the PA subtype, such as flavoxobin [41], remain to be examined for an understanding of their evolution and their specificity towards plasminogen or other protein substrates.

Representative amino acid sequences of the venom proteases were aligned in accordance with the subtypes and compared

(Table 1). We noticed that three regions (residues 82–99, 192–193 and 217–219) bear unique substitutions that might be subtype-specific. These regions happen to be the known substrate-binding subsites of the enzyme as shown by crystallographic studies [31,42]. Moreover, site-directed mutagenesis studies have proved that Asp-97 and residue 193 of TSV-PA are involved in plasminogen recognition [43,44]. Table 1 shows that residue 97 of the enzymes in the PA subgroup is usually Asp or Glu (i.e. an acidic residue), whereas it is Thr-97 in both the CL and KN subgroups. It is noteworthy that positions 192 and 193 are basic residues in enzymes of the CL subtype, whereas they are usually Gly residues in the KN and PA subtypes. TSV-PA mutant F193G has a 9-fold higher plasminogenolytic activity than wild-type TSV-PA, whereas the activity of the mutant F193R was 6-fold lower [44]. Residue 192 in the coagulating serine proteases has been recognized as a key residue affecting the enzyme specificity for the P2 subsite [45,46], whereas residues 217 and 218 are close to the P3 subsite [42]. Residue 218 is usually Asp in the KN subtype and His in the PA subtype. Thus mutagenesis studies at these sites should enable us to pinpoint the residues involved in the specificity of the serine proteases for substrates in the blood.

The specificities of acutobin for the synthetic substrates seem to be consistent with the conclusion from related studies that the thrombin-like venom proteases have a higher specificity for Pro residues than for Gly and Phe at the P2 position of substrates [47,48].

The present paper has clarified the structure of the major serine proteases of *D. acutus*, which have been used clinically on many Chinese patients. Several other serine proteases, possibly present in small amounts in this venom, await further characterization. Phylogenetic analyses of venom proteins' sequences are proving to be a powerful tool for classifying the venom proteins by function or for predicting the functional subtype of a new venom's protein sequence. It will undoubtedly also facilitate meaningful sequence comparisons and subsequently rational mutagenesis in order to understand structural and functional relationships between the venom proteins.

This study was supported by research grants from Academia Sinica Taiwan, and National Science Council, Taiwan, ROC. We thank Dr Ming-Yi Liao for supplying the snake venom gland and venom.

REFERENCES

- Pirkle, H. (1998) Thrombin-like enzymes from snake venoms: an updated inventory. Scientific and Standardization Committee's Registry of Exogenous Hemostatic Factors. *Thromb. Haemostasis* **79**, 675–683
- Stocker, K. F. (1998) Research, diagnostic and medicinal uses of snake venom enzymes. In *Enzymes from Snake Venom* (Bailey, G. S., ed.), pp. 705–727. Alaken, Fort Collins, CO
- Markland, F. S. (1998) Snake venoms and the hemostatic system. *Toxicon* **36**, 1749–1800
- Ouyang, C., Teng, C. M. and Huang, T. F. (1992) Characterization of snake venom components acting on blood coagulation and platelet function. *Toxicon* **30**, 945–966
- Nikai, T. and Komori, Y. (1998) Kinin-releasing and kinin-degrading enzymes. In *Enzymes from Snake Venom* (Bailey, G. S., ed.), pp. 287–316. Alaken, Fort Collins, CO
- Komori, Y. and Nikai, T. (1998) Chemistry and biochemistry of kallikrein-like enzyme from snake venoms. *J. Toxicol. Toxin Rev.* **17**, 261–277
- Zhang, Y., Wisner, A., Xiong, Y. and Bon, C. (1995) A novel plasminogen activator from snake venom: purification, characterization, and molecular cloning. *J. Biol. Chem.* **270**, 10246–10255
- Stocker, K. (1994) Inventory of exogenous hemostatic factors affecting the prothrombin activating pathways. *Thromb. Haemostasis* **71**, 257–260
- Deshimaru, M., Ogawa, T., Nakashima, K., Nobuhisa, I., Chijiwa, T., Shimohigashi, Y., Fukumaki, Y., Niwa, M., Yamashina, I., Hattori, S. and Ohno, M. (1996) Accelerated evolution of Crotalinae snake venom gland serine proteases. *FEBS Lett.* **397**, 83–88
- Chang, M. C. and Huang, T. F. (1995) Characterization of a thrombin-like enzyme, grambin, from the venom of *Trimeresurus gramineus* and its in vivo antithrombotic effect. *Toxicon* **33**, 1087–1098
- Pan, H., Du, X. Y., Yang, G. Z., Zhou, Y. and Wu, X. (1999) cDNA cloning and expression of acutin. *Biochem. Biophys. Res. Commun.* **255**, 412–415
- Huang, Q. Q., Teng, M. K. and Liu, L. W. (1999) Purification and characterization of two fibrinogen-clotting enzymes from five-pace snake (*Agkistrodon acutus*) venom. *Toxicon* **37**, 999–1013
- Shen, J.-R., Zhu, J.-R., Wang, W.-X., Li, Z.-Z. and Shi, L.-T. (1998) The pure component thrombin-like enzyme from Chinese *Agkistrodon acutus* venom and the defibrinogen effect. *J. Toxicol. Toxin Rev.* **17**, 560–561
- Komori, Y., Nikai, T., Suzuki, Y. and Sugihara, H. (1996) Isolation and characterization of kinase from *Deinagkistrodon acutus* (hundred-pace snake) venom. *J. Nat. Toxins* **5**, 375–384
- Matsudaira, P. (1990) Limited N-terminal sequence analysis. *Methods Enzymol.* **182**, 610–611
- Hunkapiller, M. W. and Hood, L. E. (1983) Analysis of phenylthiohydantoins by ultrasensitive gradient high-performance liquid chromatography. *Methods Enzymol.* **91**, 486–493
- Chase, Jr, T. and Shaw, E. (1970) Titration of trypsin, plasmin and thrombin with *p*-nitrophenyl *p*'-guanidinobenzoate HCl. *Methods Enzymol.* **19**, 20–27
- Lottenberg, R., Christensen, I., Jackson, C. M. and Coleman, P. L. (1981) Assay of coagulation proteases using peptide chromogenic and fluorogenic substrates. *Methods Enzymol.* **80**, 341–361
- Hahn, B. S., Yang, K. Y., Park, E. M., Chang, I. M. and Kim, Y. S. (1996) Purification and molecular cloning of calobin, a thrombin-like enzyme from *Agkistrodon caliginosus* (Korean viper). *J. Biochem. (Tokyo)* **119**, 835–843
- Itoh, N., Tanaka, N., Mihashi, S. and Yamashina, I. (1987) Molecular cloning and sequence analysis of cDNA for batroxobin, a thrombin-like snake venom enzyme. *J. Biol. Chem.* **262**, 3132–3135
- Au, L. C., Lin, S. B., Chou, J. S., Teh, G. W., Chang, K. J. and Shih, C. M. (1993) Molecular cloning and sequence analysis of the cDNA for ancro, a thrombin-like enzyme from the venom of *Calloselasma rhodostoma*. *Biochem. J.* **294**, 387–390
- Felsenstein, J. (1992) PHYLIP: the PHYLogeny Inference Package, version 3.573, Computer program distributed by the University of Washington, Department of Genetics, Seattle
- Saitou, N. and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–425
- Felsenstein, J. (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783–791
- Teng, C. M. and Seegers, W. H. (1981) Production of factor X and factor Xa variants with thrombin, acutin and by autolysis. *Thromb. Res.* **22**, 213–220
- Komori, Y., Nikai, T., Sakairi, Y. and Sugihara, H. (1988) Characterization of clotting factors from *Agkistrodon acutus* venom. *Int. J. Biochem.* **20**, 387–392
- Hillis, D. M. and Bull, J. J. (1993) An empirical test of bootstrapping as a method assessing confidence in phylogenetic analysis. *Syst. Biol.* **42**, 182–192
- Henschen-Edman, A. H., Theoder, I., Edwards, B. F. P. and Pirkle, H. (1999) Crotalase, a fibrinogen-clotting snake venom enzyme: primary structure and evidence for a fibrinogen recognition exosite different from thrombin. *Thromb. Haemostasis* **81**, 81–86
- McMullen, B. A., Fujikawa, K. and Kisiel, W. (1989) Primary structure of a protein C activator from *Agkistrodon contortrix contortrix* venom. *Biochemistry* **28**, 674–679
- Perona, J. J. and Craik, C. S. (1997) Evolutionary divergence of substrate specificity within the chymotrypsin-like serine protease fold. *J. Biol. Chem.* **272**, 29987–29990
- Parry, M. A., Jacob, U., Huber, R., Wisner, A., Bon, C. and Bode, W. (1998) The crystal structure of the novel snake venom plasminogen activator TSV-PA: a prototype structure for snake venom serine proteinases. *Structure* **6**, 1195–1206
- Nakashima, K., Nobuhisa, I., Deshimaru, M., Nakai, M., Ogawa, T., Shimohigashi, Y., Fukumaki, Y., Hattori, M., Sakaki, Y., Hattori, S. and Ohno, M. (1995) Accelerated evolution in the protein-coding regions is universal in crotalinae snake venom gland phospholipase A2 isozyme genes. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 5605–5609
- Cheng, X., Qian, Y., Liu, Q., Li, B. X., Zhang, M. and Liu, J. (1999) Purification, characterization, and cDNA cloning of a new fibrinogenolytic venom protein, Agkiscutacin, from *Agkistrodon acutus* venom. *Biochem. Biophys. Res. Commun.* **265**, 530–535
- Liu, S., Huang, Q., Zhu, X., Teng, M. and Ziu, L. (1999) Purification, characterization, crystallization and preliminary X-ray diffraction of acuthrombin-B, a thrombin-like enzyme from *Agkistrodon acutus* venom. *Acta Crystallogr. D* **55**, 1193–1197
- Huang, Q. Q., Teng, M. K. and Niu, L. W. (1999) Purification and characterization of two fibrinogen-clotting enzymes from five-pace snake (*Agkistrodon acutus*) venom. *Toxicon* **37**, 999–1013
- Teng, C. M., Hsu, M. F. and Wang, J. P. (1992) Comparison of kinin-forming and amidolytic activities of four trimucases, oedema-producing and kinin-releasing enzymes, from *Trimeresurus mucrosquamatus* venom. *J. Pharm. Pharmacol.* **44**, 306–310

- 37 Serrano, S. M. T., Hagiwara, Y., Murayama, N., Higuchi, S., Mentele, R., Sampaio, C. A. M., Camargo, A. C. M. and Fink, E. (1998) Purification and characterization of a kinin-releasing and fibrinogen-clotting serine proteinase (KN-BJ) from the venom of *Bothrops jararaca*, and molecular cloning and sequence analysis of its cDNA. *Eur. J. Biochem.* **251**, 845–853
- 38 Matsui, T., Sakurai, Y., Fujimura, Y., Hayashi, I., Oh-ishi, S., Suzuki, M., Mamako, J., Yamamoto, Y., Yamazaki, J., Kinoshita, M. and Titani, K. (1998) Purification and amino acid sequence of halastase from snake venom of *Agkistrodon halys blomhoffii*, a serine protease that cleaves specifically fibrinogen and kininogen. *Eur. J. Biochem.* **252**, 569–575
- 39 Park, D., Kim, H., Chung, K., Kim, D. S. and Yun, Y. (1998) Expression and characterization of a novel plasminogen activator from *Agkistrodon halys* venom. *Toxicon* **36**, 1807–1819
- 40 Hung, C. C., Huang, K. F. and Chiou, S. H. (1994) Characterization of one novel venom protease with beta-fibrinogenase activity from the Taiwan habu (*Trimeresurus mucrosquamatus*): purification and cDNA sequence analysis. *Biochem. Biophys. Res. Commun.* **205**, 1707–1715
- 41 Shieh, T. C., Kawabata, S. I., Kihara, H., Ohno, M. and Iwanaga, S. (1988) Amino acid sequence of a coagulant enzyme, flavoxobin, from *Trimeresurus flavoviridis* venom. *J. Biochem. (Tokyo)* **103**, 596–605
- 42 Chang, Y. J., Hamaguchi, N., Chang, S. C., Ruf, W., Shen, M. C. and Lin, S. W. (1999) Engineered recombinant factor-VII Q217 variants with altered inhibitor specificities. *Biochemistry* **38**, 10940–10948
- 43 Zhang, Y., Wisner, A., Maroun, R. C., Choumet, V., Xiong, Y. and Bon, C. (1997) *Trimeresurus stejnegeri* snake venom plasminogen activator. Site-directed mutagenesis and molecular modeling. *J. Biol. Chem.* **272**, 20531–20537
- 44 Braud, S., Parry, M. A. A., Maroun, R. and Bon, C. (2000) The contribution of residues 192 and 193 to the specificity of snake venom serine proteinases. *J. Biol. Chem.* **275**, 1823–1828
- 45 Rezaie, A. R. and Esmon, C. T. (1996) Molecular basis of residue 192 participation in determination of coagulation protease specificity. *Eur. J. Biochem.* **242**, 477–484
- 46 Zhang, Y. L., Hervio, L., Strandberg, L. and Madison, E. L. (1999) Distinct contributions of residue 192 to the specificity of coagulation and fibrinolytic serine proteases. *J. Biol. Chem.* **274**, 7153–7156
- 47 Mackessy, S. P. (1993) Kallikrein-like and thrombin-like proteases from the venom of juvenile northern Pacific rattlesnakes (*Crotalus v. oreganus*). *J. Nat. Toxins* **2**, 223–239
- 48 Zhang, Y., Goa, R., Lee, W. H., Xiong, Y. L. and Wang, W. Y. (1998) Characterization of a fibrinogen-clotting enzyme from *Trimeresurus stejnegeri* venom, and comparative study with other venom proteases. *Toxicon* **36**, 131–142
- 49 Fan, C., Qian, Y., Gong, Y. and Yang, S. (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
- 50 Hahn, B. S., Back, K., Kim, W. S., Lee, C. S., Chang, I. L. and Kim, Y. S. (1998) Molecular cloning of capillary permeability-increasing enzyme-2 from *Agkistrodon caliginosus* (Korean viper). *Toxicon* **36**, 1887–1893

Received 16 August 2000/26 October 2000; accepted 28 November 2000