Accelerated evolution of crotalinae snake venom gland serine proteases

Masanobu Deshimaru[a], Tomohisa Ogawa[a], Kin-ichi Nakashima[a], Ikuo Nobuhisa[a], Takahito Chijiwa[a], Yasuyuki Shimohigashi[a], Yasuyuki Fukumaki[b], Mineo Niwa[c], Ikuo Yamashina[d], Shosaku Hattori[e], Motonori Ohno[a,*]

[a]*Department of Chemistry, Faculty of Science, Kyushu University, Higashi-ku, Fukuoka 812, Japan*
[b]*Institute of Genetic Information, Kyushu University, Higashi-ku, Fukuoka 812, Japan*
[c]*Faculty of Pharmaceutical Sciences, Tokushima University, Tokushima, Tokushima 770, Japan*
[d]*Faculty of Engineering, Kyoto Industrial University, Kita-ku, Kyoto 603, Japan*
[e]*Institute for Medical Science, University of Tokyo, Oshima-gun, Kagoshima 894-15, Japan*

Received 29 July 1996; revised version received 1 October 1996

Abstract Eight cDNAs encoding serine proteases isolated from *Trimeresurus flavoviridis* (habu snake) and *T. gramineus* (green habu snake) venom gland cDNA libraries showed that nonsynonymous nucleotide substitutions have accumulated in the mature protein-coding regions to cause amino acid changes. Southern blot analysis of *T. flavoviridis* genomic DNAs using two proper probes indicated that venom gland serine protease genes form a multigene family in the genome. These observations suggest that venom gland serine proteases have diversified their amino acid sequences in an accelerating manner. Since a similar feature has been previously discovered in crotalinae snake venom gland phospholipase A_2 (PLA_2) isozyme genes, accelerated evolution appears to be universal in plural isozyme families of crotalinae snake venom gland.

Key words: Snake venom; Serine protease family; cDNA cloning; Evolution

1. Introduction

Snake venom gland contains a great variety of enzymes as toxic components. Among them, diverse serine proteases have been isolated from the venoms of viperidae snakes [1,2]. Three fibrinogenolytic serine proteases were isolated from the venoms of two *Trimeresurus* species and were characterized to be different in fibrinopeptide specificity. Flavoxobin isolated from *T. flavoviridis* (habu snake, crotalinae) venom excises fibrinopeptide A selectively [3], whereas okinaxobins I [4] and II [5] isolated from *T. okinavensis* venom release fibrinopeptide B and fibrinopeptides A and B, respectively. Besides, various types of snake venom serine proteases which display other physiological activities than fibrinogenolytic activity have been reported: for example, platelet-aggregating enzyme from *Bothrops jararaca* venom [6], protein C activator from *Agkistrodon contortrix contortrix* venom [7], factor V activators from *Vipera russelli* venom [8], kallikrein-like enzyme (KLE) from *Crotalus atrox* venom [9], capillary permeability increasing (CPI) enzymes from *A. caliginosus* venom [10,11] and plasminogen activator from *Trimeresurus stejnegeri* venom [12]. Some multifunctional serine proteases have also been reported from snake and lizard venoms: crotalase from *C. adamanteus* venom possesses fibrinogenolytic and kallikrein-like activities [13] and gilatoxin from *Heloderma horridum horridum* venom shows kallikrein-like, angiotensin-degrading and fibrinogenolytic activities [14]. In spite of such a variety of physiological properties, they are composed of approximately 235 amino acids and are highly homologous in sequence. Thus, these serine proteases seem to have diverged from a common ancestral molecule and constitute a gene family. Ito et al. reported that the exon/intron organization of the gene for batroxobin, a fibrinogenolytic enzyme from *Bothrops atrox* (fer-de-lance, crotalinae) venom gland, is different from that of thrombin but similar to that of kallikrein [15]. This suggests that snake venom gland serine proteases including batroxobin have diverged from glandular kallikrein and that their functions have altered in the process of evolution.

On the other hand, *T. flavoviridis* venom contains phospholipase A_2 (PLA_2) isozymes with a variety of physiological activities [16]. Comparison of their cDNAs [17] and genes [18] showed that the protein-coding regions are much more diversified than the noncoding regions including introns and that nonsynonymous nucleotide substitutions are greater than or close to synonymous nucleotide substitutions in the protein-coding regions. These unusual features indicated that *T. flavoviridis* venom gland PLA_2 isozymes have evolved via accelerated evolution to gain diverse physiological functions [19].

In the work reported here, we investigated the molecular evolution of crotalinae snake venom gland serine proteases. A series of cDNA clones encoding *T. flavoviridis* and *T. gramineus* (green habu snake, crotalinae) venom gland serine proteases were isolated from their cDNA libraries and sequenced. Their evolutionary features are discussed based on computational analysis of their nucleotide sequences.

2. Materials and methods

Specimens of *T. flavoviridis* and *T. gramineus* were collected in Amami-Oshima Island, Japan, and Taiwan, respectively. Venom gland total RNAs were extracted from a single specimen of each species by the guanidinium isothiocyanate/CsCl method [20]. Poly-$(A)^+$ RNAs were purified using mRNA purification kit (Pharmacia). *T. flavoviridis* genomic DNAs were prepared from the liver of a single specimen by proteinase K/SDS treatment (0.05% (w/v) and 5% (w/v), respectively), extracted twice with phenol and three times with phenol/chloroform, and precipitated by 2-propanol.

Venom gland cDNA libraries were constructed with cDNA synthesis kit and λgt10 cloning system (Amersham). cDNA first strand was synthesized with venom gland poly$(A)^+$ RNA (2 μg) and

*Corresponding author. Fax: (81) (92) 632-2734.

Abbreviations: KLE, kallikrein-like enzyme; CPI, capillary permeability increasing; PLA_2, phospholipase A_2; TBP, TATA box binding protein; PCR, polymerase chain reaction; RT, reverse transcription; UTR, untranslated region

PII S0014-5793(96)01144-1

oligo(dT)$_{12-18}$ primer. Completed libraries contained 2.3×10^7 (*T. flavoviridis*) and 3.4×10^5 (*T. gramineus*) primary recombinants.

In order to obtain a partial fragment of cDNA encoding flavoxobin, reverse transcription-polymerase chain reaction (RT-PCR) was carried out for *T. flavoviridis* venom gland mRNAs. cDNA first strand was synthesized with oligo(dT)$_{12-18}$ primer and PCR was performed with flavoxobin-degenerated sense primer FX-N (5′-GAT/CGAA/GTGT/CAAT/CATA/T/CAAT/CGAG/ACA-3′, corresponding to the amino acid sequence, ^{5}Asp-Glu-Cys-Asn-Ile-Asn-Glu-^{12}His) and antisense primer FX-C (5′-TGG/A/TATCCAG/A/T/CGCA/GTTA/GTAA/GTCA/GAA-3′, corresponding to the sequence, ^{216}Phe-Asp-Tyr-Asn-Ala-Trp-Ile-^{223}Gln) [21].

For screening of serine protease cDNAs, approximately 5×10^5 plaques from *T. flavoviridis* or *T. gramineus* venom gland cDNA libraries were screened by the plaque hybridization method [22] with the PCR-derived probe and batroxobin cDNA [23]. Hybridization was carried out at 42°C for 12 h in a mixture of 5×standard saline citrate (1×SSC: 1.0 M sodium chloride, 0.1 M sodium citrate), 50 mM sodium phosphate pH 6.5, 0.5% SDS, 2×Denhardt's solution (100×Denhardt's solution: 0.1% Ficoll 400, 0.1% polyvinyl-pyrrolidone, 0.1% bovine serum albumin), 50% formamide and 100 mg/ml sonicated salmon sperm DNA. The membranes were washed finally in 0.2×SSC and 0.1% SDS at 65°C (PCR-derived DNA) or in 2×SSC and 0.1% SDS at 42°C (batroxobin cDNA) twice for 30 min, respectively. Five to ten positive clones were isolated and their DNAs were extracted.

For Southern blot analysis, *T. flavoviridis* genomic DNAs (15 μg) were digested with *Eco*RI, *Hin*dIII and *Pst*I. DNA fragments were electrophoresed on a 0.7% agarose gel and transferred to Hybond-N$^+$ nylon membrane (Amersham). The 5′ probe (5′-untranslated region (UTR) plus signal peptide-coding region, corresponding to nucleotides 1–245, 245 bp) and the coding probe (mature protein-coding region, corresponding to nucleotides 246–956, 711 bp) for TLf1, a serine protease cDNA, were prepared by PCR, internal-labelled with α-[^{32}P]dCTP and employed for hybridization experiments [22]. The membranes were washed finally in 2×SSC and 0.1% SDS at 45°C and analyzed by BAS 2000 Bio imaging analyzer (Fuji Photo Film Co. Ltd., Japan).

The DNASIS package developed by Hitachi Software Engineering was employed for analysis and alignment of DNA sequences. The number of nucleotide substitutions per site (K_N) in the 5′- and 3′-UTRs and the numbers of nucleotide substitutions per synonymous site (K_S) and per nonsynonymous site (K_A) in the protein-coding region were computed for all the pairs of cDNAs according to the method of Nei and Gojobori [24] using ODEN package developed by Y. Ina (National Institute for Genetics, Mishima, Japan). A synonymous site is a site of codon at which base substitution causes no amino acid change. A nonsynonymous site is a site of codon at which base substitution causes amino acid change. The phylogenetic trees for serine proteases were constructed using the various components of the cDNA sequences by the one-parameter method [25] and the neighbor-joining method [26].

The nucleotide sequence data reported in this paper are available from DDBJ, EMBL and GenBank databases with accession numbers of D67078–D67085.

3. Results

RT-PCR of *T. flavoviridis* mRNAs with flavoxobin-degenerated primers, FX-N and FX-C, gave an amplified 656-bp DNA fragment. Sequence analysis revealed that this DNA encoded the amino acid sequence identical to flavoxobin [21] except for substitutions of five amino acids. *T. flavoviridis* venom gland cDNA library was screened with this DNA fragment as a probe. Five positive clones were obtained from 5×10^5 plaques by washing in high stringency (0.2×SSC, 0.1% SDS, 65°C). These cDNAs encoded the same protein which is 97.9% identical to flavoxobin, although the 5′- and 3′-UTRs were different in length. Such cDNAs were designated TLf1. Similarly, five cDNAs selected from *T. gramineus* venom gland cDNA library encoded the same protein which is 84.3% identical to flavoxobin and were designated TLg1.

In order to obtain cDNAs encoding serine proteases less homologous to flavoxobin, the protein-coding region of batroxobin cDNA [23] was employed as a probe and the membranes were washed in low stringency (2×SSC, 0.1% SDS, 42°C). In this experiment, first screening of *T. flavoviridis* and *T. gramineus* cDNA libraries (approximately 5×10^5 plaques each) gave more than 200 positive plaques and 15 cDNA clones were selected based on signal strength. TLf2 and TLf3 were isolated from *T. flavoviridis* cDNA library and TLg2a, TLg2b, TLg2c and TLg3 were from *T. gramineus* cDNA library. Their deduced amino acid sequences were 62.7%, 63.1%, 66.5%, 67.4%, 63.1% and 62.3% identical to flavoxobin, respectively.

Fig. 1 shows the nucleotide sequence of TLf1, one of the isolated cDNAs. All cDNAs were similar in total length (approximately 1.6 kb) and encoded nearly 260 amino acids including a potential signal peptide of 24 amino acids. The amino acid sequences are shown in Fig. 2 together with those of flavoxobin [21] and batroxobin [23]. Catalytic triads common to serine proteases (His-43, Asp-88 and Ser-182) are conserved except that proteins for TLf2 and TLg2a contain arginine instead of His-43. The positions of 12 half-cystines are identical among all the sequences, suggesting that the proteins encoded will take similar tertiary structures.

The nucleotide sequences of *T. flavoviridis* and *T. gramineus* venom gland serine protease cDNAs and batroxobin cDNA were compared to one another by homology plot analysis of the DNASIS program. Since all the combinations gave the similar patterns in terms of nucleotide substitutions, 11 combinations are shown in Fig. 3. The identical regions are indicated by filled boxes and the blank spaces correspond to the substituted nucleotides. It should be noted that substitution of one nucleotide is to be expressed as a 10-nucleotide-sized blank space. For all the combinations, the nucleotide substitutions occur predominantly in the protein-coding region except for the signal peptide domain, indicating that the mature protein-coding region is much more variable than the 5′- and 3′-UTRs. The sectional identities between TLf1 and TLf2, for

Table 1
The K_N, K_S, K_A, K_N/K_S and K_A/K_S values for 11 combinations within species and between species of venom gland serine protease cDNAs

Combination	K_N	K_S	K_A	K_N/K_S	K_A/K_S
Within species					
T. flavoviridis					
TLf1–TLf2	0.0592	0.179	0.183	0.330	1.021
TLf1–TLf3	0.0637	0.142	0.188	0.449	1.322
TLf2–TLf3	0.0607	0.097	0.159	0.625	1.636
T. gramineus					
TLg1–TLg2a	0.0480	0.153	0.167	0.314	1.093
GLg1–TLg3	0.0699	0.147	0.186	0.476	1.268
TLg2a–TLg3	0.0488	0.113	0.153	0.431	1.350
Between species					
T. flavoviridis vs. T. gramineus					
TLf1–TLg1	0.0439	0.066	0.071	0.661	1.065
TLf2–TLg2a	0.0353	0.038	0.034	0.927	0.883
TLf3–TLg3	0.0726	0.108	0.072	0.675	0.674
T. flavoviridis vs. B. atrox					
TLf1–batroxobin	0.0763	0.177	0.177	0.431	1.001
T. gramineus vs. B. atrox					
TLg1–batroxobin	0.0754	0.148	0.184	0.510	1.248

```
cttgtcagcgttccagattgttggccacccagctgcttaatttgatcaaa   50
taaagtgctgcttgatcaagaagtctccgcttggcttatctgattagatt  100
aatacggtatctcaagtttaagtttggaactgggatcttacaggcaaaca  150
gctttccacgcagagttgaagctATGGTTCTGATCAGAGTGCTAGCAAAC  200
CTTCTGATACTACAGCTTTCTTACGCACAAAAGTCTTCTGAATTGGTCAT  250
TGGAGGTGATGAATGTAACATAAATGAACATCCTTTCCTTGTAGCCTTGT  300
ATGATGCTTGGTCTGGAAGATTTCTCTGTGGTGGGACTTTGATCAACCCG  350
GAATGGGTGCTCACTGCTGCACACTGCGACAGTAAAAATTTCAAGATGAA  400
GCTTGGTGCGCATAGCAAAAAGGTACTAAATGAGGATGAACAGATAAGAA  450
ACCCAAAGGAGAAGTTCATTTGTCCCAATAAGAAAAACGATGAAGTACTG  500
GACAAAGATATCATGTTGATCAAGCTGGACAGTCCTGTTAGCTACAGTGA  550
ACACATCGCGCCTCTCAGCTTGCCTTCCAGCCCTCCCAGTGTGGGCTCAG  600
TTTGCCGTATTATGGGATGGGGATCAATCACACCTGTTGAAGAGACTTTT  650
CCCGATGTCCCCCATTGTGCTAACATTAACCTACTCGATGATGTGGAGTG  700
TAAACCAGGTTATCCAGAGTTGCTGCCAGAATACAGAACATTGTGTGCAG  750
GTGTCCTGCAAGGAGGCATAGATACATGTGGGTTTGACTCTGGGACACCC  800
CTCATCTGTAATGGACAATTCCAGGGTATTGTATCTTATGGGGGGCATCC  850
TTGTGGCCAAAGTCGTAAGCCTGGTATCTACACCAAGGTCTTTGATTATA  900
ATGCCTGGATTCAGAGCATTATTGCAGGAAATACAGCTGCAACTTGCCTC  950
CCGTGAaaactgtaaatgtaacatattagtacatgtcttctatatcccta 1000
accatatccgactgcattggaatatattcccaggcagtaagctttttta 1050
gactcaaataggactgcctttggagtaagaaatgctcaaaatagtgctgc 1100
agggatcatgtcccatttaatttcagtataaaacaatctcagtaaaatag 1150
aggcctgttttagggtcagtgcgaaatttttctgactctaaaatggacca 1200
ttccaaatattttaagctctgaatatctttccatttctgaccacttctgg 1250
gacagtggggtccttgatgctctctgagcttgtcttcttgcagacatttc 1300
attacccagctaggtaacatcatcagtgctagaatattctcttctattgg 1350
tacttctgtggcatttacaatacgctcatatggagtcatgcagtcacccc 1400
acaaacatatccatatacctaggtcccactgttgcctaaaaaggatccca 1450
gattaacctccacttcccaatcaacaaatagagtcttttgagaatcatgt 1500
tttcatgtaaattctcaagtaactacaggaataaaattgtataaatcgtt 1550
aaaaaaaaaaaaaaaaaaaa 1570
```

Fig. 1. The nucleotide sequence of TLf1, a cDNA encoding a *T. flavoviridis* venom gland serine protease. The protein-coding region is shown in upper-case letters and the 5′- and 3′-UTRs are in lower-case letters. The initiation codon (ATG) and polyadenylation signal (aataaa) are underlined. The mature protein-coding region which correspond to the amino acid sequence of flavoxobin is boxed.

example, were 91.7% for the 5′-UTR, 95.8% for the signal peptide-coding domain, 80.5% for the mature protein-coding region and 93.3% for the 3′-UTR. This structural feature of serine protease cDNAs is in accord with those previously seen in crotalinae snake venom gland PLA_2 isozyme cDNAs [17] and genes [18] but is the opposite to that of general isoprotein genes in which the protein-coding regions are much more conserved than the UTRs [27,28].

Mathematical analysis was made for the nucleotide sequences of cDNAs encoding *T. flavoviridis* and *T. gramineus* venom gland serine proteases and batroxobin to obtain further information about their evolutionary behavior. The K_N values for the 5′- and 3′-UTRs and the K_S and K_A values for the protein-coding region were computed for all the pairs of cDNAs [24]. Table 1 shows the data of 11 out of 36 combinations. The others also showed similar values. Two remarkable features are noted. First, K_N/K_S values are smaller than 1, indicating that nucleotide substitutions have more frequently occurred in the protein-coding region than in the UTRs. Second, K_A/K_S values are greater than or close to 1. This indicates that nucleotide substitutions in the protein-coding region tend to cause amino acid change. These two features are in sharp contrast to those observed in general isozyme genes in which the noncoding region is more variable than the protein-coding region and the K_A/K_S values were calculated to be about 0.2 [27,28].

Genomic DNAs extracted from *T. flavoviridis* liver were digested with *Eco*RI, *Hin*dIII and *Pst*I and employed for Southern blot analysis with two types of probes. When the coding probe which is the mature protein-coding region of TLf1 was used, nearly 10 bands were detected as shown in Fig. 4. These bands may represent DNAs with high homology to the protein-coding region of TLf1. On the other hand, when the 5′ probe which corresponds to the 5′-UTR plus signal peptide domain of TLf1 was employed, numerous bands lying between 3.0 and 10 kb were hybridized almost without break (Fig. 4). Since the 5′-UTR is conserved in serine protease cDNAs, all the hybridized DNA fragments are presumed to reflect the number of venom gland serine protease genes in the genome. These results indicate that there are a number of venom gland serine protease genes which form a multigene family.

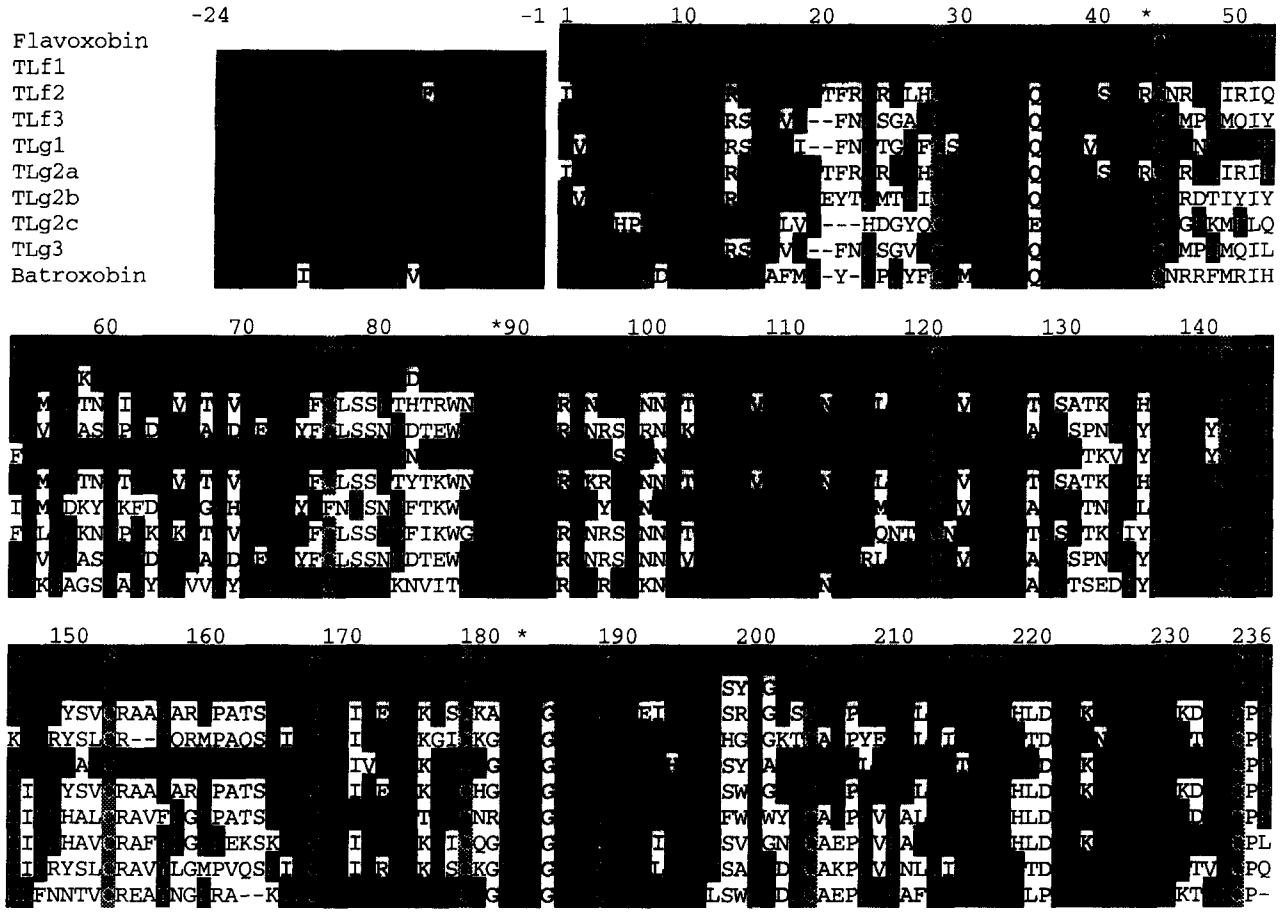

Fig. 2. Amino acid sequences of *T. flavoviridis* and *T. gramineus* venom gland serine proteases deduced from their cDNAs and those of flavoxobin and batroxobin. Residues −24 to −1 correspond to a potential signal peptide. Flavoxobin is taken as a standard for the mature protein sequences and TLf1 for the signal peptide sequences. The identical amino acids are in reverse-shaded letters. Dashes represent the deleted residues. Half-cystine residues are shaded. The catalytic triad residues (His-43, Asp-88 and Ser-182) of serine proteases are indicated by an asterisk.

4. Discussion

Eight cDNAs which encode proteins homologous to flavoxobin were isolated from *T. flavoviridis* and *T. gramineus* venom gland cDNA libraries. It is noted that proteins encoded by this series of cDNAs contain a number of amino acid substitutions throughout the mature protein sequences (Fig. 2). There are some tendencies in their amino acid sequences

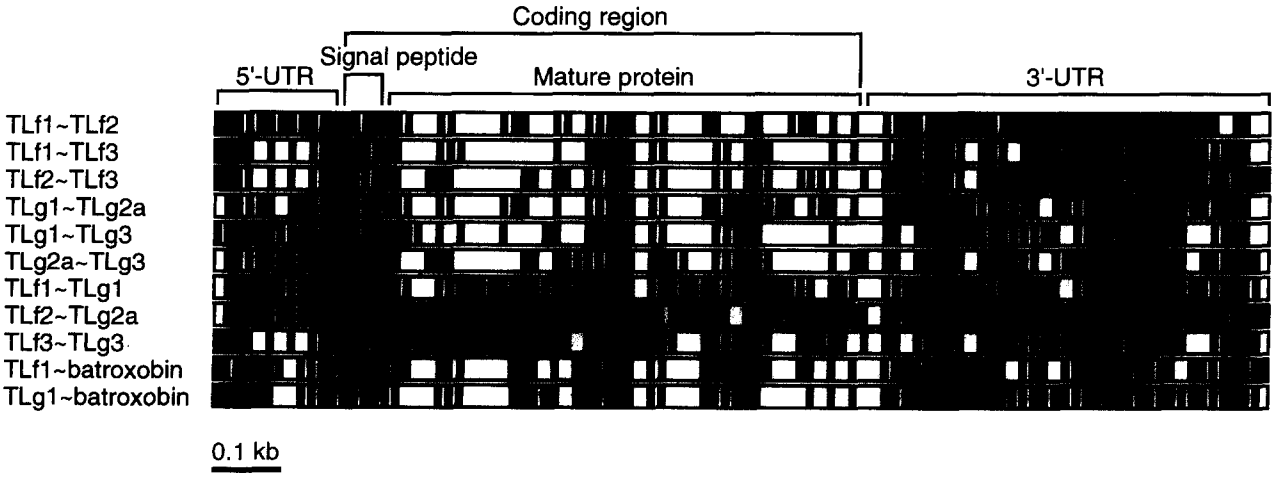

Fig. 3. Schematic representation of homology plot analysis for 11 combinations of cDNAs encoding *T. flavoviridis* and *T. gramineus* venom gland serine proteases and batroxobin. The regions in which a set of sequential 10 nucleotides is completely identical between paired cDNAs are represented by filled boxes. The blank spaces correspond to the positions of substituted nucleotides. Shaded boxes indicate the regions for which the nucleotide sequences could not be compared due to deletion of more than two consecutive nucleotides in one of paired cDNAs.

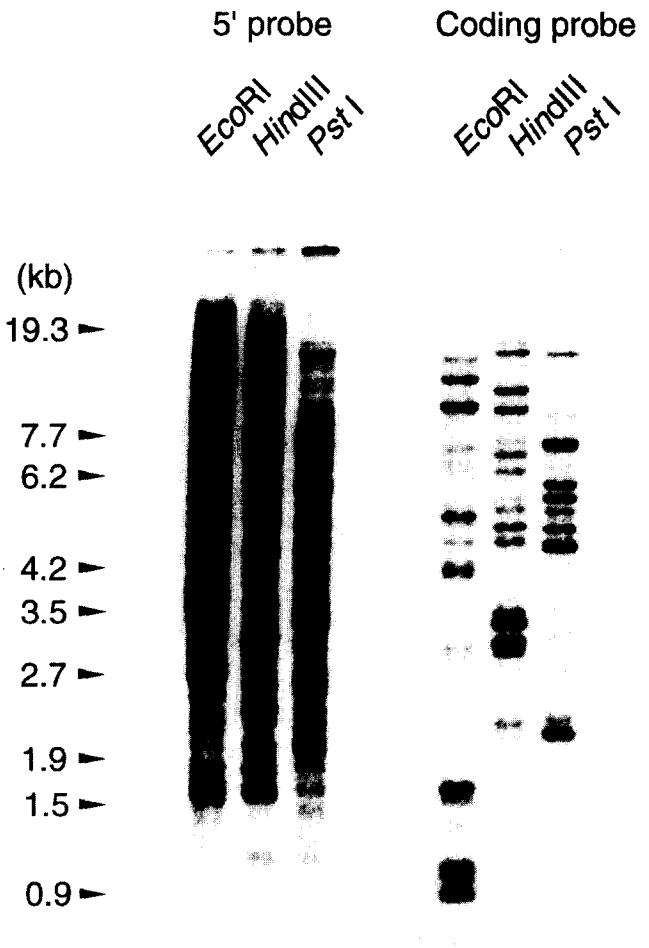

Fig. 4. Southern blot analysis of *T. flavoviridis* genomic DNAs. The DNAs (5 μg) were digested with *Eco*RI, *Hin*dIII and *Pst*I, electrophoresed on a 0.7% agarose gel, blotted to nylon membranes and hybridized with the coding probe (left) or the 5′ probe (right). The 5′ probe is the 5′-UTR plus signal peptide-coding region of TLf1 and the coding probe is the mature protein-coding region of TLf1. The hybridized bands were analyzed by BAS 2000 Bio imaging analyzer (Fuji Photo Film Co. Ltd., Japan).

which enable us to surmise their own physiological activities. Since TLf1 and TLg1 encode proteins highly homologous to flavoxobin [21] with only five and 37 amino acid substitutions, respectively, they might be expected to exhibit thrombin-like activity. In particular, TLf1 appears to encode flavoxobin itself because errors in amino acid sequence determination of flavoxobin could not be excluded completely [21]. Gln-58, Thr-84, Tyr-199, Ile-200 and Ser-202 of flavoxobin were replaced by Lys-58, Asp-84, Ser-199, Tyr-200 and Gly-202, respectively, in the deduced sequence of TLf1 (Fig. 2). When RT-PCR experiments for *T. flavoviridis* venom gland RNAs were carried out with two different types of oligonucleotides specific to residues 198–203 of flavoxobin and TLf1, DNA was amplified only by TLf1 specific oligonucleotide (data not shown). RNA samples from 10 individual specimens gave the same result, strongly suggesting that mRNA corresponding to the amino acid sequence of flavoxobin does not occur.

The proteins encoded by TLf2, TLg2a and TLg2c contain many more arginine residues than the others and their isoelectric points (pI) predicted by DNASIS program were 9.22, 9.01 and 8.14, respectively. The proteins encoded by other cDNAs were predicted to be acidic (pIs 5.09–5.92). Actually, most of known venom gland serine proteases have been reported to be acidic: for example, pI values for flavoxobin [3], *C. atrox* KLE (EI) [9] and *A. caliginosus* CPI enzyme-2 [11] are 4.8, 4.7 and 3.5, respectively. Even though it appears to be inappropriate to conjecture their activities from pIs, basic isozymes might be expected to have their own specific physiological activities considerably different from known acidic serine proteases. The N-terminal regions of proteins encoded by TLf3 and TLg3 are highly homologous to those of *A. caliginosus* CPI enzymes 1 and 2 [10,11], so that it is conceivable that they might exhibit similar activity as CPI enzymes. On the other hand, the activities of proteins encoded by TLf2 and TLg2a which contain arginine instead of His-43, a residue constituting the catalytic triad, are unknown. Phylogenetic tree constructed for these serine protease cDNAs using the nucleotide sequences of mature protein-coding region is shown in Fig. 5. Eight cDNAs of *Trimeresurus* species appear to be separated into three groups. Each group includes a pair consisting of one each from *T. flavoviridis* and of *T. gramineus* cDNAs and the paired cDNAs might encode the proteins with similar properties as described above. It is likely that the pairing proteases have evolved in parallel after divergence into the different species. Otherwise, an inexplicable phylogenetic tree was obtained when the noncoding region sequences were employed (data not shown).

Mathematical analysis of the nucleotide sequences of cDNAs showed that K_N/K_S values are smaller than 1 for any pairs irrespective of comparison within species or between species (Table 1), indicating that nucleotide substitutions in the protein-coding region have occurred more frequently than in the UTRs. Since K_N values are known to be greater than K_S values in general isozyme genes [27,28], the observation made in serine protease cDNAs is unusual. This can be considered in two ways. One is unusually high conservation of the nucleotide sequence in the UTRs and the other is unusually high frequency of nucleotide substitutions in the protein-coding region. Recently, the genes encoding TATA box binding proteins (TBPs), a basal transcription factor, of *T. flavoviridis* and *T. gramineus* were isolated and sequenced

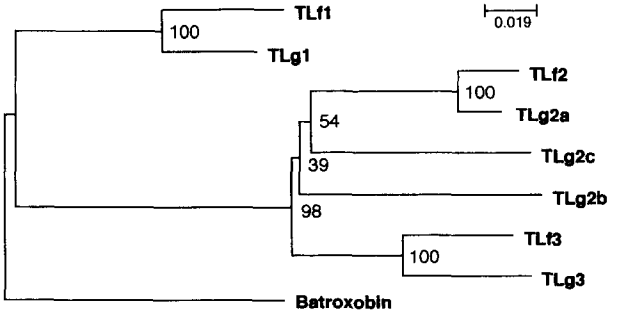

Fig. 5. Phylogenetic tree for mature protein-coding region sequences of serine protease cDNAs of *T. flavoviridis*, *T. gramineus* and *B. atrox*, constructed by the neighbor-joining method [25] based on the number of nucleotide substitutions per site. The numbers on the branches stand for boot-strap probability [29]. The horizontal bars are drawn to scale and represent the numbers of nucleotide substitutions per site.

[30]. Since TBP could be regarded as a general gene which has evolved according to the neutral theory [27], the K_N value for TBP genes was employed as a standard. The K_N values (0.035–0.083) for venom gland serine protease cDNAs (Table 1) were found to be comparable to that for TBP genes (0.0365) [19]. Thus, it became evident that nucleotide substitutions in the UTRs of serine protease cDNAs have occurred as frequent as those of general isozyme genes and that frequency of nucleotide substitutions in the protein-coding region of serine protease cDNAs has been unusually high.

For venom serine protease cDNAs, K_A values are generally greater than K_S values with ratios of 0.67–1.64 (Table 1). This is also in contrast to general isozyme genes in which K_A values are about five times smaller than K_S values [27,28] where nonsynonymous nucleotide substitution has been suppressed possibly to avoid an alteration of protein functions. Such observations suggest that nonsynonymous nucleotide substitutions causing amino acid changes have much more frequently occurred in venom gland serine protease genes than in general isozyme genes.

Several genes encoding venom gland serine proteases were also isolated from *T. flavoviridis* genomic DNA library and their partial sequences were determined (data not shown). They revealed that introns are highly conserved, like their UTRs, suggesting that only the mature protein-coding region has evolved at an unusually higher rate.

Similar evolutionary behavior has previously been observed in the cDNAs [17] and genes [18,19] encoding crotalinae snake venom gland PLA_2 isozymes. This was ascribed to accelerated evolution of these genes [19] accompanied by functional diversification of their products [16]. The acquisition of diverse physiological activities in venom isozymes must be advantageous for disrupting the integrity of prey animals or for defense against predators. The present study provide the evidence that *T. flavoviridis* and *T. gramineus* serine protease isozyme genes have also diverged via accelerated evolution possibly to gain functional diversity in their products. In addition to PLA_2 isozyme and serine protease families, many other isozyme families occur in crotalinae snake venom. It could be assumed that all of them have universally evolved via accelerated evolution.

The nucleotide substitutions having occurred intensively at a specific site of genes have been noted in a few other multigene families. A well-known example is MHC class II genes [31] whose hypervariability is thought to have been produced by gene conversion. In this case, however, nonsynonymous nucleotide substitutions have occurred predominantly in α_1 and α_2 regions which interact with specific antigens, but the other regions of the protein-coding region are highly conserved. In contrast, uniquely in venom isozyme genes, nonsynonymous nucleotide substitutions spread over the entire mature protein-coding region although the UTRs and introns are highly conserved. Mechanism other than gene conversion may be responsible for accelerated evolution of venom gland isozyme genes.

References

[1] Ouyang, C., Teng, C.-M. and Huang, T.-F. (1990) Adv. Exp. Med. Biol. 281, 151–163.
[2] Pirkle, H. and Theodor, I. (1990) Adv. Exp. Med. Biol. 281, 165–175.
[3] Shieh, T.-C., Tanaka, S., Kihara, H., Ohno, M. and Makisumi, S. (1985) J. Biochem. (Tokyo) 98, 713–721.
[4] Iwasaki, A., Shieh, T.-C., Shimohigashi, Y. Waki, M., Kihara, H. and Ohno, M. (1990) J. Biochem. (Tokyo) 108, 822–828.
[5] Nose, T., Shimohigashi, Y., Hattori, S., Kihara, H. and Ohno, M. (1994) Toxicon 32, 1509–1520.
[6] Serrano, S.M.T., Mentele, R., Sampaio, C.A.M. and Fink, E. (1995) Biochemistry 34, 7186–7193.
[7] McMullen, B.A., Fujikawa, K. and Kisiel, W. (1989) Biochemistry 28, 674–679.
[8] Tokunaga, F., Nagasawa, K., Tamura, S., Miyata, T., Iwanaga, S. and Kisiel, W. (1988) J. Biol. Chem. 263, 17471–17481.
[9] Bjarnason, J.B., Barish, A., Direnzo, G.S., Campbell, R. and Fox, J.W. (1983) J. Biol. Chem. 258, 12566–12573.
[10] Ohtani, Y. and Takahashi, H. (1988) Toxicon 26, 181–190.
[11] Shimokawa, K. and Takahashi, H. (1993) Toxicon 31, 1221–1227.
[12] Zhang, Y., Wisner, A., Xiong, Y. and Bon, C. (1995) J. Biol. Chem. 270, 10246–10255.
[13] Markland, F.S., Kettner, C., Schiffman, S., Shaw, E., Bajwa, S.S., Reddy, K.N., Kirakossian, H., Patkos, G.B., Theodor, I. and Pirkle, H. (1982) Proc. Natl. Acad. Sci. USA 79, 1688–1692.
[14] Utaisincharoen, P., Mackessy, S.P., Miller, R.A. and Tu, A.T. (1993) J. Biol. Chem. 268, 21975–21983.
[15] Itoh, N., Tanaka, N., Funakoshi, I., Kawasaki, T., Mihashi, S. and Yamashina, I. (1988) J. Biol. Chem. 263. 7628–7631.
[16] Kihara, H., Uchikawa, R., Hattori, S. and Ohno, M. (1992) Biochem. Int. 28, 895–903.
[17] Ogawa, T., Oda, N., Nakashima, K., Sasaki, H., Hattori, M., Sakaki, Y., Kihara, H. and Ohno, M. (1992) Proc. Natl. Acad. Sci. USA 89, 8557–8561.
[18] Nakashima, K., Ogawa, T., Oda, N., Hattori, M., Sakaki, Y., Kihara, H. and Ohno, M. (1993) Proc. Natl. Acad. Sci. USA 90, 5964–5968.
[19] Nakashima, K., Nobuhisa, I., Deshimaru, M., Nakai, M., Ogawa, T., Shimohigashi, Y., Fukumaki, Y., Hattori, M., Sakaki, Y., Hattori, S. and Ohno, M. (1995) Proc. Natl. Acad. Sci. USA 92, 5605–5609.
[20] Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) Biochemistry 15, 5294–5299.
[21] Shieh, T.-C., Kawabata, S., Kihara, H., Ohno, M. and Iwanaga, S. (1988) J. Biochem. (Tokyo) 103, 596–605.
[22] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
[23] Itoh, N., Tanaka, N., Mihashi, S. and Yamashina, I. (1987) J. Biol. Chem. 262, 3132–3135.
[24] Nei, M. and Gojobori, T. (1986) Mol. Biol. Evol. 3, 418–426.
[25] Saitou, N. and Nei, M. (1987) Mol. Biol. Evol. 4, 406–425.
[26] Jukes, T.H. and Canter C.R. (1969) In: Munro, H.N. (Ed.), Mammalian Protein Metabolism. Academic Press, New York.
[27] Kimura, M. (1983) The Neutral Theory of Molecular Evolution. Cambridge University Press, Cambridge.
[28] Nei, M. (1987) Molecular Evolutionary Genetics. Columbia University Press, New York.
[29] Felsenstein, J. (1985) Evolution 39, 783–791.
[30] Nakashima, K., Nobuhisa, I., Deshimaru, M., Ogawa, T., Shimohigashi, Y., Fukumaki, Y., Hattori, M., Sakaki, Y., Hattori, S. and Ohno, M. (1995) Gene 152, 209–213.
[31] Hughes, A.L. and Nei, M. (1989) Proc. Natl. Acad. Sci. USA 86, 958–962.