Adaptive evolution in the snake venom Kunitz/BPTI protein family

Vera Župunski[a,*], Dušan Kordiš[a], Franc Gubenšek[a,b]

[a]*Department of Biochemistry and Molecular Biology, Jožef Stefan Institute, Jamova 39, 1000 Ljubljana, Slovenia*
[b]*Department of Chemistry and Biochemistry, Faculty of Chemistry and Chemical Technology, University of Ljubljana, Ljubljana, Slovenia*

Abstract Snake venoms are rich sources of serine proteinase inhibitors that are members of the Kunitz/BPTI (bovine pancreatic trypsin inhibitor) family. However, only a few of their gene sequences have been determined from snakes. We therefore cloned the cDNAs for the trypsin and chymotrypsin inhibitors from a *Vipera ammodytes* venom gland cDNA library. Phylogenetic analysis of these and other snake Kunitz/BPTI homologs shows the presence of three clusters, where sequences cluster by functional role. Analysis of the nucleotide sequences from the snake Kunitz/BPTI family shows that positive Darwinian selection was operating on the highly conserved BPTI fold, indicating that this family evolved by gene duplication and rapid diversification.
© 2003 Federation of European Biochemical Societies. Published by Elsevier Science B.V. All rights reserved.

Key words: Kunitz/BPTI homolog; Snake venom; Multigene family; Adaptive evolution

1. Introduction

Snake venoms are rich sources of many different proteins encoded by functionally diversified multigene families. A relatively small number of these multigene families have been studied including phospholipases A_2 (PLA_2s) [1–4], PLA_2 inhibitors [5,6], serine proteases [7] and three-fingered toxins [8]. These families evolved by gene duplication and were later functionally diversified by positive selection [9].

Many serine protease inhibitors were isolated from the venoms of *Viperidae* and *Elapidae* snakes [10–22]. They belong to the functionally diverse bovine pancreatic trypsin inhibitor (BPTI) family [23,24]. These Kunitz/BPTI inhibitors are present in a variety of animals and are characterized by a conserved fold of approximately 60 amino acids, stabilized by three disulfide bridges. Snake Kunitz/BPTI inhibitors have been divided into two groups on the basis of their function. Trypsin inhibitor (Ti) and chymotrypsin inhibitor (Chi) are referred to as non-neurotoxic snake Kunitz/BPTI inhibitors, whereas the homologs with neurotoxic effects belong to

the neurotoxic snake Kunitz/BPTI group [25]. These homologs have lost their inhibitory role and act as K^+ and Ca^{2+} channel blockers [17–19].

Kunitz/BPTI homologs are, based on their three-dimensional (3D) structure, classified into two families: canonical Kunitz-type inhibitors including BPTI-like toxins, and anticoagulant proteins [26]. The latter have secondary structures whose orientation is clearly similar to that of BPTI, but differ in the folds of some loops and particularly in the orientation of the N-terminal segment [27]. Most of the canonical Kunitz-type homologs inhibit serine proteinases of the chymotrypsin family through their highly conserved antiproteinase site. The specificity towards serine proteases is defined by the P1 amino acid and small sequence differences in the region that interacts with the proteases [24]. Snake venom Kunitz/BPTI inhibitors show a completely different pattern of sequence conservation [25,28]. The most important regions for neurotoxicity are the surface near the C-terminal part of dendrotoxins (Dtxs) and the most exposed parts of the B chains of β-bungarotoxins (β-Btxs). Snake venom non-neurotoxic Kunitz/BPTI inhibitors are most conserved in the core and in the N-terminal surface area and not at the antiproteinase site. Thus, they have retained the same fold but developed a variety of functions [25]. However, the physiological role of these in snakes is not known. It has been proposed that they are involved in the processes of coagulation, fibrinolysis and inflammation through undefined interactions with proteases [12]. Textilinins (Txs), BPTI homologs from *Pseudonaja textilis textilis*, for example, inhibit trypsin and plasmin [29]. Therefore, they might be involved in plasmin-mediated digestion of fibrin clots. In fact, they reduce blood loss in a murine bleeding model [29]. From the known structure–function relationships in BPTI homologs, it is obvious that flexibility, rigidity and variations of the amino acid residues are strongly connected with their biological activity [26].

The venom of *Vipera ammodytes* contains several Kunitz/BPTI inhibitors. Ti and Chi show 80% sequence identity [10,11]. These inhibitors are 61 and 65 amino acids long, respectively, with three conserved disulfide bridges. The level of similarity with other snake BPTI homologs is up to 50%. Only a few Kunitz/BPTI inhibitors from snake venoms have known nucleotide sequences: Kunitz inhibitors from *Bungarus candidus*, Tx [20], B chains of β-Btx [30] and DtxK [31]. We have isolated and sequenced Ti and Chi cDNAs from a *V. ammodytes* venom gland cDNA library. Molecular evolution of snake Kunitz/BPTI inhibitors was analyzed and compared to other multigene families from snake venoms. The results show that Kunitz/BPTI inhibitors from snake venoms are encoded by multigene families, and were diversified by posi-

*Corresponding author. Fax: (386)-1-2573594.
E-mail address: vera.zupunski@ijs.si (V. Župunski).

tive Darwinian selection. The adaptive evolution of venomous proteins has resulted in the advantage of snakes in the predator–prey competition.

2. Materials and methods

2.1. Screening the cDNA library

The venom gland cDNA library from *V. ammodytes* [32] prepared in phage λgt11, was screened with a 155-bp-long fragment of the *V. ammodytes* Ti [11]. This fragment was polymerase chain reaction (PCR) amplified from the same library with sense primer 5′-CAGAATTCCAYCCNAARTTYTGYTA and antisense primer 5′-GAGAATTCRCAYTCRTCCCANGTYTT, encoding HPKFCY and KTWDEC amino acids at the N- and C- terminal part of the Ti, respectively. The probe was labelled with [^{35}S]dATP by the random priming method [33]. The library was screened by the plaque hybridization method [34] under the following conditions: 20 h hybridization at 42°C in hybridization buffer (6×SSC (NaCl/Na-citrate), 50% (v/v) formamide, 5×Denhardt's solution, 0.5% (w/v) sodium dodecyl sulfate (SDS) and 500 μg denatured herring sperm DNA). The filters were washed in 4×SSC for 15 min and 2×SSC for 5 min at room temperature. Positive clones were purified by repeating the screening procedure.

2.2. DNA sequencing

Phage DNA was isolated from plate lysates [34] and digested with *Eco*RI restriction enzyme. cDNA fragments were separated from the phage DNA by electrophoresis on 1.5% agarose gel. They were cloned into pUC19 vector and sequenced with an ABI 310 sequence analyzer using BigDye chemistry (Applied Biosystems). Both strands were sequenced.

2.3. PCR amplification of cDNA

Additional Ti and Chi cDNAs were amplified by PCR from the venom gland cDNA library using sense primer 5′-RGAGAATAAATAGAGCSAGCAG and antisense primer 5′-CACCWGARACCRAGARGGCAGS. The degenerate primers were designed based on 5′ and 3′ conserved untranslated regions (UTRs) of *V. ammodytes* Ti and B1 chain of Btx. cDNAs were amplified from 1 μl of *V. ammodytes* cDNA library with 6×10^9 plaque-forming units (PFU)/ml, 30 pmol of each primer, 200 μM of each dNTP, 1.5 μl MgCl$_2$ and 2.5 U of *Taq* DNA polymerase (Promega). Amplification involved 10 min hot-start, 30 cycles of 1 min denaturation at 94°C, 1 min annealing at 46°C and 45 s elongation at 72°C. The PCR fragments were purified by electrophoresis in 1.5% agarose gel and extraction from the gel using QIAEX (Qiagen). They were cloned into pGEM-T by pGEM-T easy vector system I kit (Promega). Nucleotide sequences were determined as described above.

2.4. Sequence analyses

V. ammodytes Ti and Chi cDNAs were translated with the Translate tool at ExPASy (http://www.expasy.org/tools/dna.html). GenBank databases were searched with snake Ti and Chi using different BLAST programs [35]. Genetic distances, numbers of non-synonymous substitutions per non-synonymous site (d_N) and numbers of synonymous substitutions per synonymous site (d_S) of snake Kunitz/BPTI inhibitors were calculated by the MEGA2 program [36]. The significance of differences between d_N and d_S was determined by using the one-tailed test in the MEGA2 program.

2.5. Phylogenetic analyses

The amino acid sequences of snake Kunitz/BPTI inhibitors were aligned using Clustal X [37] and GeneDoc [38]. A rooted phylogenetic tree was inferred using the neighbor-joining (NJ) method [39] as implemented in Treecon [40]. *Bos taurus* BPTI was used as an outgroup. The significance of various phylogenetic lineages was assessed by bootstrap analysis from 1000 replicates.

3. Results and discussion

3.1. Isolation and PCR amplification of V. ammodytes Ti and Chi cDNAs

Several positive clones, encoding 5′ site truncated Ti and

Chi cDNAs were identified by screening the *V. ammodytes* cDNA library made from a single venom gland. Full-length Ti and Chi cDNAs from *V. ammodytes* (accession numbers AY217781 and AY217782) were amplified with novel degenerate primers designed on the highly conserved 5′ and 3′ UTRs of snake Kunitz/BPTI inhibitors.

3.2. Sequence analyses of V. ammodytes Ti and Chi cDNAs

The cDNA sequences for the inhibitors have an initiation codon for methionine, stop codon, polyadenylation signal (AATAAA) and poly(A) tail. Ti and Chi cDNAs encode precursors of 90 and 93 amino acids, respectively (Fig. 1). The signal peptides, 24 amino acids long, are highly conserved, and differ only in the last amino acid. Both cDNAs code for sequences longer at the C-terminus than the mature proteins [10,11] by five amino acids in Ti (MGRPT) and four amino acids in Chi (IASN). Ti and Chi cDNAs encode 61 and 65 residues in the mature proteins, respectively. The nucleotide and amino acid sequences of mature Ti and Chi are 94 and 81% identical, respectively. Most nucleotide replacements lead to non-synonymous substitutions. Amino acid differences between the inhibitors occur at the N- and C-terminal parts of the mature proteins, and in the loop after the P1 site which interacts with the proteases. These regions are involved in the interactions with their target proteins and define the inhibitors' specificity.

Ti cDNA encodes the same sequence as that in the inhibitor isolated from the *V. ammodytes* venom [11], whereas the Chi cDNA encodes an isoform of the isolated Chi. The amino acid sequence of Chi cDNA differs from the isolated inhibitor by four amino acids in the region following the antiproteinase site of the inhibitor [10]. In addition to two isoforms of the Chi, two different Tis were isolated from the venom of *V. ammodytes*, but only one of them was sequenced [11]. These results suggest that Kunitz/BPTI inhibitors from *V. ammodytes* are encoded by a multigene family.

3.3. Snake venom Kunitz/BPTI inhibitors

Several neurotoxic and non-neurotoxic snake homologs were found by searching databases with *V. ammodytes* Ti and Chi. Nucleotide sequences of six different plasmin inhibitors called Txs from *P. textilis textilis* (AF402324–AF402329), Kunitz inhibitors a, b and c from *B. candidus* (AY057886–AY057888), DtxK from *Dendroaspis polylepis polylepis* (S61886), four B chains of β-Btx from *Bungarus multicinctus* (AJ242991, AJ251223, Y12100, Y12101) and three B chains of β-Btx from *B. candidus* (AY057883–AY057885) were compared with the sequences of *V. ammodytes* Ti and Chi cDNAs. The BLASTP and TBLASTN searching showed the presence of neurotoxic and non-neurotoxic representatives of the Kunitz/BPTI family (Fig. 1). The highest degree of sequence similarity was observed with Txs – more than 60% in 83 amino acids.

3.4. High degree of conservation in signal peptides of snake Kunitz/BPTI inhibitors

V. ammodytes Ti and Chi, B chains of β-Btx, DtxK and Txs are the only snake Kunitz/BPTI inhibitors with known signal peptide sequences, and, moreover, the signal peptide of DtxK is truncated by two amino acids at its N-terminus [31]. The signal peptides are highly conserved (Fig. 1), differing only in five out of 24 amino acids. Highly conserved sequences of the

```
                    *         20        *         40        *         60        *         80        *
DtxK     :----SGHLLLLLGLLTLWAELTPVSGAAKYCKLLLRICPCKRKIPSFYWKAKQCLPFDYSGCGGNANRFKTIEECRRTCVG----------79
d-Dtx    :-------------------------AAKYCKLLVRYFPCKKKIPSFYWKAKQCLPFDYSGCGGNANRFKTIEECRRTCVG----------57
DanCal   :-----------------------WQPPWYCKEEVRICSCKKQFSSFFKWTAKKCLPFLFSGCGGNANRFQTIGECRKKCLGK---------60
DtxI     :-----------------------QPLRKLCILHRNPGRCYQKIPAFYNQKKQCEGFTWSGCGGNSNRFKTIEECRRTCIRK---------60
a-Dtx    :-----------------------QPRRKLCILHRNPGRCYDKIPAFYNQKKQCERFDWSGCGGNSNRFKTIEECRRTCVG----------59
DtxE     :----------------------LQHRTFCKLLAEPGPCKASIPAFYNWAAKKCQLFHYGGCKGNANRFSTIEKCRHACVG----------59
e-Dtx1   :-------------------------HRTFCKLLAEPGPCKASIPAFYNWAAKKCQLFHYGGCKGNANRFSTIEKCRHACVG----------57
e-Dtx2   :----------------------LQHRTFCKLLAEPGPCKASIPAFYNWAAKKCQLFHYGGCKGNANRFSTIEKCRRACVG----------59
DtxB     :-------------------------RPYACELIVAAGPCMFFISALSKGANKCYPFTYSGCRGNANRFKTIEECRRTCVV----------57
NnaChi   :-------------------------RPRFCELAPSACSCFGFVSSYYNRYSNTCHSFTYSGCGGNANRFKTIDECNRTCVV----------57
BPTI     :-------------------------RPDFCLEEPYTGPCKARIIRYFYNAKAGLCQTFVYGGCRAKRNNFKSAEDCMRTCGGAIGPWENL--65
NnaTi    :-------------------------RPGFCELPAAKGLCKAHKPAFYNKDSHRCQKFIYGGCGGNANRFRTIDECNRTCVG----------57
OhaChi   :------------------------GRPKFCELLPEPGLCNARKTFYYSLHSHACQKFIYGGCGGNANRFKTIDECHRTCVG----------58
NniTi    :-------------------------RPRFCELLAETGLCKARIRSFHYNRAAQQCLEFIYGGCGGNANRFKTIDECHRTCVG----------57
HhaTi    :-------------------------RPDFCELLAETGLCKAYIRSFHYNLAAQQCLQFIYGGCGGNANRFKTIDECRRTCVG----------57
BfaChi   :------------------------KNRPTFCNLLPETGRCNALIPAFYNSHLHKCQKFNYGGCGGNANNFKTIEECQRTCAAKYGRSS----65
VamChi   :-------------------------RDRPKFCYLLADPGRCLAYMPRFYNPASNKCEKFIYGGCRGNANNFKTWDECRHTCVASGIQPR----65
cVamChi  :MSSGGLLLLLGLLTLWAELTPVSTRDRPKFCYLLADPGRCLAYMPSFYDSASNKCKKFIYGGCRGNANNFKTWDECRHTCVASGIQPRIASN93
cVamTi   :MSSSGGLLLLLGLLTLWAELTPVSGQDHPKFCYLLADPGRCKAHIPRFYDSASNKCNKFIYGGCPGNANNFKTWDECRQTCGASAMGRPT---90
VamTi    :-------------------------QDHPKFCYLLADPGRCKAHIPRFYDSASNKCNKFIYGGCPGNANNFKTWDECRQTCGASA-------61
EmaTi    :-------------------------FCYLLDDPGVCKAHIPRFYNPASNKCKNFIYGGCGGNANNFERAECRHTCVASGKGGPRP--62
DruTi    :-------------------------HDRPTFCNLLAPESGRCRGHLRRIFYNLESNKCKVFFYGGCGGNANNFKTRDECRETCGGK---------60
BmuBtxB1a:MSSGGLLLLLGLLTLSAELIPVSSRQRHRDCDKLPDKCNCGPVRRAFYDTRLKTCKAFQYRGCNGNGNHFKTETLCRCECLVYP--------85
BmuBtxB1 :MSSGGLLLLLGLLTLCAELIPVSSRQRHRDCDKLPDKCNCGPVRRAFYDTRLKTCKAFQYRGCNGNGNHFKTETLCRCECLVYP--------85
BcaBtxB1 :MSSGGLLLLLGLLTLCAELTPVSSRQRHRDCDKLPDKCNCGSVRRAFYDTRLKTCKAFPYRGCNGNGNHFKTETLCRCECLVYP--------85
BmuBtxB3 :MSSGGLLLLLGLLTLCAELIPVSSRQRHRDCDKLPDKCNCGPVRRAFYDTRLKTCKAFQYRGCNGNGNHFKSDHLCRCECLEYS--------85
BmuBtxB2 :MSSGGLLLLLGLLTLCAELTPVSSRKRHPDCDKLPDTKICQTVVRAFYKPSAKRCVQFRYGGCNGNGNHFKSDHLCRCECLEYR--------85
BcaBtxB2a:MSSGGLLLLLGLLTLWAELTPVSSRKRHPDCDKLPDTKICQTVVRAFYKPSAKRCVQFRYGGCNGNGNHFKSDHLCRCECLEYP--------85
BcaBtxB2b:MSSGGLLLLLGLLTLWAELTPVSSRKRHPDCDKLPDTRICQTVVRAFYKPSEKRCVGFRYGGCKGNGNHFKSDHLCRCECLEYR--------85
BcaKiA   :MSSGGLLLLLGLLTLCAELTPVSSKDRPKFCNVLPEPGRCNANVRAFYNPRLRKCIEFTYGGCGGNANNFKSRGECKRTCAE---------83
BcaKiB   :MSSGGLLLLVGLLTLCAELTPVSSKDRPKFCNVLPEPGRCNANVRAFYNPRLRKCIEFTYGGCGGNANNFKSGGECKRACGE---------83
BcaKiC   :MSSGGLLLLLGLLTLWTELTPVSSKNRPPFCNLLPEPGRCNAIVRAFYNSRLRKCLEFPYGGCGGNANNFKTIDECQRTCAG---------83
PteTx1   :MSSGGLLLLLGLLTLWEVLTPVSSKDRPDFCELLADTCPCRVRFPSFYNPDEKKCLEFIYGGCEGNANNFITKEECESTCAA---------83
PteTx2   :MSSGGLLLLLGLLTLWEVLTPVSSKDRPELCELLPDTCPCRVRFPSFYNPDEQKCLEFIYGGCEGNANNFITKEECESTCAA---------83
PteTx5   :MSSGGLLLLLGLLTLWEVLTPVSSKDRPKFCELLPDTCSCEDFTGAFHYSTRDRECIEFIYGGCGGNANNFITKEECESTCAA---------83
PteTx6   :MSSGGLLLLLGLLTLWEVLTPVSSKDRPKFCELLADICPCDDFTGAFHYSPREHECIEFIYGGCKGNANNFNTQEECESTCAA---------83
PteTx4   :MSSGGLLLLLGLLTLWEVLTPVSSKDHPKFCELLADTCSCKGNVPRFYNADHHQCLKFIYGGCGGNANNFKTIEECKSTCAA---------83
PteTx3   :MSSGGLLLLLGLLTLWEVLTPVSSKDRPNFCKLLAETCRCNAKIPRFYNPRQHQCIEFLYGGCGGNANNFKTIKECESTCAA---------83
```
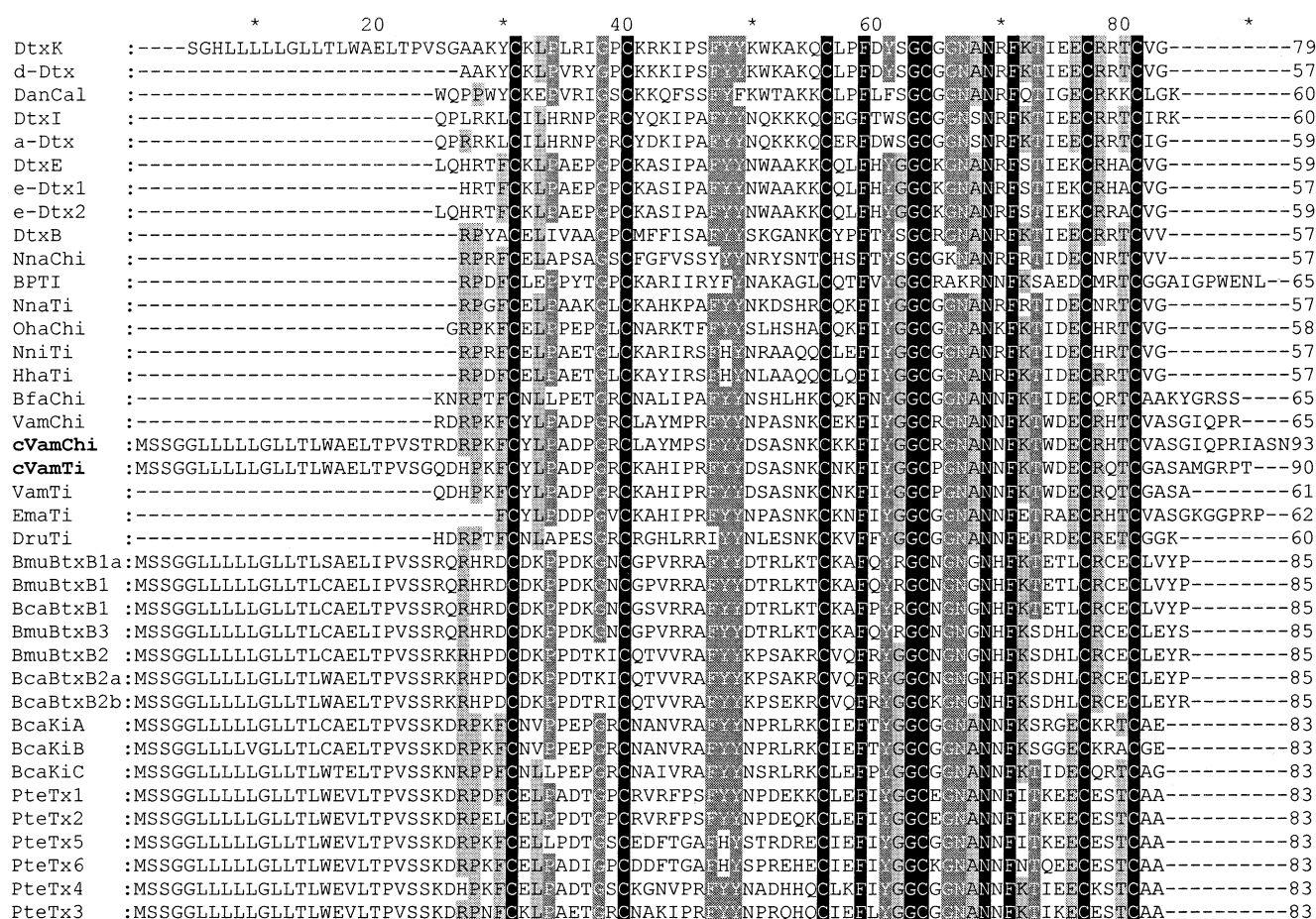
Fig. 1. Alignment of the snake Kunitz/BPTI protein family. The alignment was constructed with programs Clustal X [37] and GeneDoc [38]. Identical residues are in black, more than 80% of identical residues are in dark gray and more than 60% of identical residues in light gray. Snake venom Kunitz/BPTI inhibitors in the alignment are: two *B. multicinctus* β-Btxs B1 chains (BmuBtxB1a, AJ251223; BmuBtxB1, Y12100), *B. multicinctus* β-Btxs B2 chain (BmuBtxB2, Y12101), *B. multicinctus* β-Btx B3 chain (BmuBtxB3, AJ242991), *B. candidus* β-Btx B1 chain (BcaBtxB1, AY057883), *B. candidus* β-Btx B2a chain (BcaBtxB2a, AY057884), *B. candidus* β-Btx B2b chain (BcaBtxB2b, AY057885), Dtx from *D. polylepis* (Dpo) DtxK (S61886), DtxI (P00979), DtxB (P00983), DtxE (P00984), Dtxs from *D. angusticeps* (Dan) α-Dtx (P00980), δ-Dtx (P00982), ε-Dtx1 (B59399), ε-Dtx2 (A59399), calcicludine from *D. angusticeps* (DanCal, A36989), Tx from *P. textilis* (Pte) PteTx1 (AF402324), PteTx2 (AF402325), PteTx3 (AF402326), PteTx4 (AF402327), PteTx5 (AF402328), PteTx6 (AF402329), Ti from *V. ammodytes* (VamTi, P00991), Chi from *V. ammodytes* (VamChi, P00992), Ti cDNA from *V. ammodytes* (cVamTi, AY217781), Chi cDNA from *V. ammodytes* (cVamChi, AY217782), Ti from *Naja naja* (NnaTi, S13846), Chi from *N. naja* (NnaChi, S12957), Ti from *Eristocophis macmahonii* (EmaTi, S19327), Chi from *Ophiophagus hannah* (OhaChi, P82966), Chi from *Bungarus fasciatus* (BfaChi, P25660), Ti from *Daboia russeli* (DruTi, P00990), Ti from *Naja nivea* (NniTi, P00986), Ti from *Hemachatus haemachatus* (HhaTi, P00985), *B. candidus* Kunitz inhibitor a (BcaKiA, AY057886), *B. candidus* Kunitz inhibitor b (BcaKiB, AY057887), *B. candidus* Kunitz inhibitor c (BcaKiC, AY057888), and BPTI (AAD13685).

signal peptide clearly show their common origin from an ancestral gene.

3.5. Phylogenetic analysis of the snake Kunitz/BPTI family of proteins

The evolutionary relationships of snake Kunitz/BPTI proteins were inferred by constructing the phylogenetic tree using the NJ algorithm [39]. Amino acid sequences of mature proteins from different *Elapidae* and *Viperidae* species were analyzed using *B. taurus* BPTI as an outgroup (Fig. 2).

Snake venom Kunitz/BPTI homologs form three clusters. The first cluster contains non-neurotoxic *Elapidae* and *Viperidae* representatives, while the second cluster contains Dtxs and additional *Elapidae* non-neurotoxic homologs. B chains of β-Btxs from *Bungarus* species comprise the third cluster. This analysis shows that clustering is by protein function.

Currently available data show that the Kunitz/BPTI homologs in *Elapidae* snakes are more diverse than those in *Viperidae* snakes. The presence of several diverse proteins in the same species, as shown by Txs, Dtxs and β-Btxs, indicates that they are indeed members of multigene families.

DtxB and DtxE from *D. polylepis polylepis* and their homologs from *Dendroaspis angusticeps* (ε-Dtxs) show strong protease inhibitory activity and weak K^+ channel inhibition [41–43]. The separate position of DtxB from the Dtx group indicates that the neurotoxic lineage evolved after gene duplication. On the basis of the biological activity and structural differences it has been proposed that ε-Dtxs have structural and functional characteristics that are intermediate between protease inhibitors and their neurotoxic homologs [43]. Calcicludine, which is a Ca^+ channel blocker in *D. angusticeps*, evolved after gene duplication and functional diversification.

B chains of β-Btxs form a separate cluster of snake Kunitz/BPTI inhibitors. B chains exhibit no inhibitory activity or neurotoxicity by themselves, whereas the heterodimer of B chain covalently linked to a PLA_2 is a potent neurotoxin
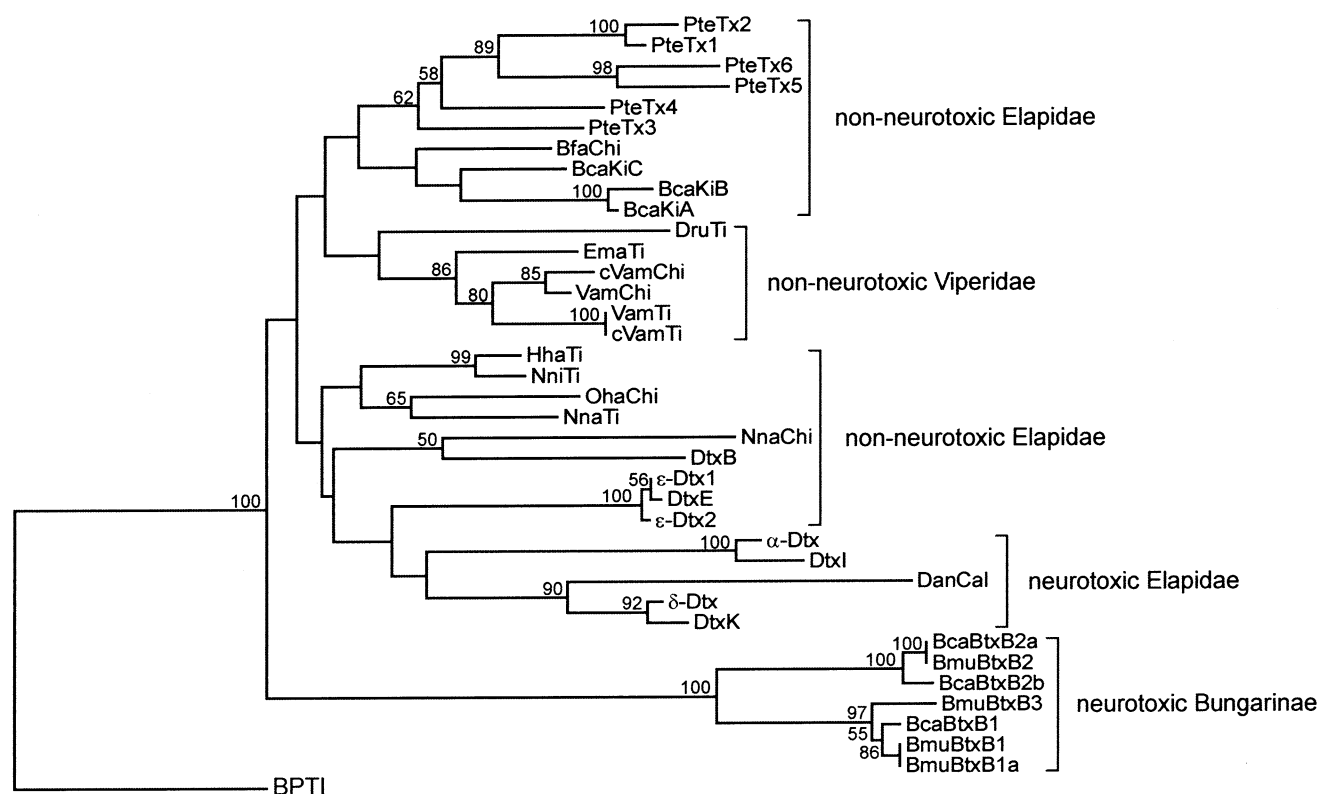
Fig. 2. NJ phylogenetic tree of the snake Kunitz/BPTI protein family. The NJ phylogenetic tree was made using the Poisson correction model with BPTI as an outgroup. The NJ tree was drawn by Treecon program [40] and represents the bootstrap consensus following 1000 replicates. Nodes with confidence values greater than 50% are indicated. See Fig. 1 for species and protein abbreviations.

[18]. Therefore, their separate position in the NJ tree is imposed by specific functional constraints.

3.6. Adaptive evolution of the snake Kunitz/BPTI family of proteins

19 nucleotide sequences of snake Kunitz/BPTI inhibitors are currently known. Genetic distances, number of synonymous substitutions per synonymous site (d_S) and number of non-synonymous substitutions per non-synonymous site (d_N) for each pair of these homologs have been computed (Table 1). Patterns of nucleotide substitutions within signal peptide and mature protein have been examined separately. The high amino acid conservation of signal peptides is not accompanied by conservation in the mature protein sequences. Interspecies identity of snake Kunitz/BPTI inhibitors is 50–70%, while intraspecies identity exceeds 70%. d_N is significantly greater

Table 1
d_N/d_S values between pairs of snake Kunitz/BPTI homologs

pairs of homologs	d_N	d_S	d_N/d_S	P values	d (M)	d (3′ UTR)
Vam Chi vs. Ti	0.089	0.018	4.94	0.001	0.072	0.037
PteTx1 vs. Tx3	0.207	0.097	2.13	0.048	0.180	0
PteTx1 vs. Tx5	0.233	0.076	3.07	0.012	0.195	0
PteTx2 vs. Tx5	0.213	0.048	4.44	0.002	0.173	0
PteTx2 vs. Tx6	0.215	0.068	3.16	0.011	0.180	0
PteTx2 vs. Tx3	0.227	0.040	5.68	0	0.180	0
PteTx3 vs. Tx5	0.264	0.047	5.62	0	0.210	0
PteTx3 vs. Tx6	0.222	0.078	2.85	0.009	0.187	0
BcaKiA/BmuB2	0.579	0.261	2.22	0.011	0.493	/
BcaKiA/BcaB2a	0.579	0.261	2.22	0.011	0.493	0.050
BcaKiA/BcaB2b	0.575	0.271	2.12	0.014	0.491	0.038
BcaKiB/BmuB2	0.553	0.285	1.94	0.021	0.480	/
BcaKiB/BcaB2a	0.553	0.285	1.94	0.021	0.480	0.077
BcaKiB/BcaB2b	0.529	0.270	1.96	0.020	0.458	0.065
BmuB1/BmuB3	0.030	0	∞	0.044	0.023	/
BcaB1/BmuB3	0.046	0	∞	0.014	0.035	/
BcaKiB/BcaKiC	0.132	0.024	5.50	0.010	0.104	0.051

The Kunitz/BPTI family of proteins was analyzed with MEGA2 using the Nei–Gojobori method. Pairs of homologs with d_N/d_S value >1 and P value <0.05 are shown. Kimura two-parameter distances (d) were calculated for the mature protein coding regions (M) and 3′ UTR. d_S, number of synonymous substitutions per synonymous site; d_N, number of non-synonymous substitutions per non-synonymous site; Vam, *V. ammodytes*; Pte,- *P. textilis*; Bmu, *B. multicinctus*; Bca, *B. candidus*.

than d_S in intraspecies pairwise comparisons of mature proteins from *P. textilis*, *V. ammodytes*, *B. multicinctus* and *B. candidus*, and in interspecies pairwise comparison of *Bungarus* species (Table 1). The ratio d_N/d_S for homologs from different snake species is around 1, indicating neutral evolution. The rates of substitution are different for signal peptide and mature protein. d_N and d_S are significantly higher in mature proteins and the number of nucleotide differences is smaller in signal peptides. The d_N/d_S ratios of some signal peptides are greater than 1 with all *P* values over 0.05, showing neutral evolution of signal peptides. On the other hand, the *P* values for mature protein pairs, shown in Table 1, are below 0.05. Genetic distances of UTRs are lower than the distances of mature protein coding sequences (Table 1), as previously determined for animal toxin multigene families that evolved by positive selection (reviewed in [9]). Significant differences between d_N and d_S for mature *V. ammodytes* Ti and Chi, Txs and B chains of β-Btxs indicate diversifying selection and adaptive evolution, whereas the evolution of their signal peptides is functionally constrained.

The increase in evolutionary rate could be a result of positive Darwinian selection or the relaxation of functional constraints [44]. Very high d_N/d_S values are evidence that positive Darwinian selection occurred in the snake venom Kunitz/BPTI family (Table 1). The function of the snake venom is to kill and digest prey [45]. New protein functions are therefore beneficial if they result in more effective venom. Non-neurotoxic snake venom BPTI homologs inhibit a variety of serine proteases from the chymotrypsin family [10–16], but their physiological target proteins are not known. Most amino acid substitutions in the snake venom BPTI-like proteins occur in the proximity of the scissile bond, defining their specificity toward serine proteases. It seems that the pattern of conserved core and variable active site loop is a characteristic of the BPTI fold of snake venom non-neurotoxic inhibitors, as shown by Cardle and Dufton [25]. Serine proteases in diverse prey such as rodents, birds and lizards differ and are also encoded by multigene families [46]. Therefore, it is reasonable to suppose that the driving force for the variability of the inhibitor binding loop was selective pressure driven by diverse prey proteases.

The diversification, by positive Darwinian evolution, of the Kunitz/BPTI inhibitors from *V. ammodytes*, *P. textilis*, *Bungarus* and *Dendroaspis* species, and especially their inhibitor binding loop, resulted in new protein functions (Dtxs) or changed specificities toward proteases in prey. The expression of Kunitz/BPTI inhibitors in snake venom indicates that the selective pressure operating on their mature proteins may be due to the arms race between a predator and a prey. Functionally diverse snake Kunitz/BPTI inhibitors acting as neurotoxins or potential inhibitors in the coagulation cascade and fibrinolysis, have increased the potency of the snake venoms against diverse prey species.

References

[1] Nakashima, K., Ogawa, T., Oda, N., Hattori, M., Sakaki, Y., Kihara, H. and Ohno, M. (1993) Proc. Natl. Acad. Sci. USA 90, 5964–5968.

[2] Nakashima, K., Nobuhisa, I., Deshimaru, M., Nakai, M., Ogawa, T., Shimohigashi, Y., Fukumaki, Y., Hattori, M., Sakaki, Y. and Hattori, S. (1995) Proc. Natl. Acad. Sci. USA 92, 5605–5609.

[3] Kordiš, D. and Gubenšek, F. (1996) Eur. J. Biochem. 240, 83–90.

[4] Kordiš, D. and Gubenšek, F. (1997) Eur. J. Biochem. 246, 772–779.

[5] Nobuhisa, I., Nakashima, K., Deshimaru, M., Ogawa, T., Shimohigashi, Y., Fukumaki, Y., Sakaki, Y., Hattori, S., Kihara, H. and Ohno, M. (1996) Gene 172, 267–272.

[6] Nobuhisa, I., Deshimaru, M., Chijiwa, T., Nakashima, K., Ogawa, T., Shimohigashi, Y., Fukumaki, Y., Hattori, S., Kihara, H. and Ohno, M. (1997) Gene 191, 31–37.

[7] Deshimaru, M., Ogawa, T., Nakashima, K., Nobuhisa, I., Chijiwa, T., Shimohigashi, Y., Fukumaki, Y., Niwa, M., Yamashina, I., Hattori, S. and Ohno, M. (1996) FEBS Lett. 397, 83–88.

[8] Gong, N., Armugam, A. and Jeyaseelan, K. (2000) FEBS Lett. 473, 303–310.

[9] Kordiš, D. and Gubenšek, F. (2000) Gene 261, 43–52.

[10] Ritonja, A., Meloun, B. and Gubenšek, F. (1983) Biochim. Biophys. Acta 746, 138–145.

[11] Ritonja, A., Meloun, B. and Gubenšek, F. (1983) Biochim. Biophys. Acta 748, 429–435.

[12] Shafqat, J., Zaidi, Z.H. and Jornvall, H. (1990) FEBS Lett. 275, 6–8.

[13] Shafqat, J., Beg, O.U., Yin, S.J., Zaidi, Z.H. and Jornvall, H. (1990) Eur. J. Biochem. 194, 337–341.

[14] Takahashi, H., Iwanaga, S., Kitagawa, T., Hokama, Y. and Suzuki, T. (1974) J. Biochem. (Tokyo) 76, 721–733.

[15] Siddiqi, A.R., Zaidi, Z.H. and Jornvall, H. (1991) FEBS Lett. 294, 141–143.

[16] Hokama, Y., Iwanaga, S., Tatsuki, T. and Suzuki, T. (1976) J. Biochem. (Tokyo) 79, 559–578.

[17] Harvey, A.L. (2001) Toxicon 39, 15–26.

[18] Harris, J.B. (1991) in: Snake Toxins (Harvey, A.L., Ed.), pp. 91–129, Pergamon Press, New York.

[19] Schweitz, H., Heurteaux, C., Bois, P., Moinier, D., Romey, G. and Lazdunski, M. (1994) Proc. Natl. Acad. Sci. USA 91, 878–882.

[20] Filippovich, I., Sorokina, N., Masci, P.P., de Jersey, J., Whitaker, A.N., Winzor, D.J., Gaffney, P.J. and Lavin, M.F. (2002) Br. J. Haematol. 119, 376–384.

[21] Chang, L., Chung, C., Huang, H.B. and Lin, S. (2001) Biochem. Biophys. Res. Commun. 283, 862–867.

[22] Liu, C.S., Wu, T.C. and Lo, T.B. (1983) Int. J. Pept. Protein Res. 21, 209–215.

[23] Laskowski, M. (1986) Adv. Exp. Med. Biol. 199, 1–17.

[24] Laskowski, M. and Kato, I. (1980) Annu. Rev. Biochem. 49, 593–626.

[25] Cardle, L. and Dufton, M.J. (1997) Protein Eng. 10, 131–136.

[26] Chen, C., Hsu, C.H., Su, N.Y., Lin, Y.C., Chiou, S.H. and Wu, S.H. (2001) J. Biol. Chem. 276, 45079–45087.

[27] Antuch, W., Guntert, P., Billeter, M., Hawthorne, T., Grossenbacher, H. and Wuthrich, K. (1994) FEBS Lett. 352, 251–257.

[28] Pritchard, L. and Dufton, M.J. (1999) J. Mol. Biol. 285, 1589–1607.

[29] Masci, P.P., Whitaker, A.N., Sparrow, L.G., de Jersey, J., Winzor, D.J., Watters, D.J., Lavin, M.F. and Gaffney, P.J. (2000) Blood Coagul. Fibrinolysis 11, 385–393.

[30] Wu, P.F. and Chang, L.S. (2000) Eur. J. Biochem. 267, 4668–4675.

[31] Smith, L.A., Lafaye, P.J., LaPenotiere, H.F., Spain, T. and Dolly, J.O. (1993) Biochemistry 32, 5692–5697.

[32] Pungerčar, J. (1994) Ph.D. Thesis, University of Ljubljana, Slovenia.

[33] Feinberg, A.P. and Vogelstein, B. (1983) Anal. Biochem. 132, 6–13.

[34] 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.

[35] Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) J. Mol. Biol. 215, 403–410.

[36] Kumar, S., Tamura, K., Jakobsen, I.B. and Nei, M. (2001) Bioinformatics 17, 1244–1245.

[37] Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G. (1997) Nucleic Acids Res. 25, 4876–4882.

[38] Nicholas, K.B., Nicholas Jr., H.B. and Deerfield, D.W. (1997) EMBNEW.NEWS 4, 14.

[39] Saitou, N. and Nei, M. (1987) Mol. Biol. Evol. 4, 406–425.

[40] Van de Peer, Y. and De Wachter, R. (1994) Comput. Appl. Biosci. 10, 569–570.

[41] Strydom, D.J. and Joubert, F.J. (1981) Hoppe Seylers Z. Physiol. Chem. 362, 1377–1384.

[42] Tytgat, J., Vandenberghe, I., Ulens, C. and Van, B. (2001) FEBS Lett. 491, 217–221.

[43] Sigle, R., Hackett, M. and Aird, S.D. (2002) Toxicon 40, 297–308.

[44] Hughes, A.L. (2002) Trends Genet. 18, 433–434.

[45] Kardong, K.V. (1996) Herpetologica 52, 36–46.

[46] Roach, J.C., Wang, K., Gan, L. and Hood, L. (1997) J. Mol. Evol. 45, 640–652.