Molecular cloning, characterization and evolution of the gene encoding a new group of short-chain α-neurotoxins in an Australian elapid, *Pseudonaja textilis*[1]

NanLing Gong, Arunmozhiarasi Armugam, Kandiah Jeyaseelan*

Department of Biochemistry, Faculty of Medicine, National University of Singapore, 10 Medical Drive, Singapore 119260, Singapore

Received 26 January 2000; received in revised form 21 April 2000

Edited by Takashi Gojobori

Abstract The structure and organization of five genes responsible for the synthesis of six isoforms of short-chain α-neurotoxins in *Pseudonaja textilis* venom are reported in this paper. This also forms the first report which describes the synthesis of two neurotoxin mRNA variants from one of these genes (*Pt-sntx1*) as a result of alternative splicing. Each gene consists of three exons which are separated by two introns and each has a functional promoter. The promoter activity was confirmed by both CAT assay and Real-Time PCR. A transcription initiation site, two putative TATA boxes, one CCAAT box and the transcription factor binding consensus sites for AP-1, GATA-2, c/EBPb were identified in the 5′ non-coding region of each gene. Phylogenetic analysis showed that these five genes from *P. textilis* constituted a distinct group which has evolved by gene duplication followed by accelerated evolution from an ancestral gene.

© 2000 Federation of European Biochemical Societies.

Key words: α-Neurotoxin; Gene structure; Organization; Molecular evolution; *Pseudonaja textilis*

1. Introduction

α-Neurotoxins form some of the major toxic components of the snake venoms and have the capability to block nerve transmission by binding specifically to the nicotinic acetylcholine receptors on the postsynaptic membranes of skeletal muscles and/or of neurons [1]. Based on the amino acid sequences, they can be classified into short and long-chain neurotoxins. Short-chain neurotoxins consist of 60–62 amino acid residues including four disulfide bonds, whereas long-chain neurotoxins contain 66–79 amino acid residues with five disulfide bridges [2]. Although a diversity in their primary structures can be observed, they share a common three-finger-loop structure [2].

Pseudonaja textilis, a common brown snake found in the farmlands of eastern Australia, is one of the most poisonous elapids in the world. A new group of short-chain neurotoxins from *P. textilis* were identified and characterized by us recently [3]. They bind specifically to nicotinic acetylcholine receptors at neuromuscular junction, as expected for any short-chain neurotoxins. However, they possess lower neurotoxic activities and fewer amino acids (57 or 58 residues) than the short-chain neurotoxins found in other snakes. Alignment of deduced amino acid sequences of this group of neurotoxins with the other known short-chain neurotoxins shows around 40% homology between them. Furthermore, phylogenetic analysis based on the amino acid sequences suggests that these neurotoxins from *P. textilis*, while forming a new group in a phylogenetic tree, closely resemble the short neurotoxins from other Australian snakes. Analysis of the nucleotide (nt) sequences of cDNAs corresponding to these short-chain neurotoxins shows that most of the nt substitutions result in amino acid changes, as observed in short-chain neurotoxins [4], cardiotoxins [5] and phospholipases A_2 [6] from other species of snakes. So far, the genomic structures of short neurotoxins from sea snake *Laticauda semifasciata* [7] and land snakes (cobras, *Naja* species; [8,9]) and *Naja naja sputatrix* [10] have been elucidated. Thus, the determination of the gene structure of short-chain neurotoxins from *P. textilis* (a *Pseudonaja* species) will facilitate a better understanding of the evolutionary position of this new group of toxins, while providing more information for future studies on the regulation of their biosynthesis.

2. Materials and methods

2.1. P. textilis

P. textilis specimen was obtained from the Venom Supplies Pte Ltd, Adelaide, S.A., Australia. It was identified by Mr. Peter Mirstchin (Venom Supplies, Australia) based on its anatomical features.

2.2. Preparation of genomic DNA of P. textilis

The frozen liver from a single snake kept at −80°C was gently ground into fine powder using a mortar and pestle. The method of Blin and Stafford [11] was employed to isolate high molecular weight genomic DNA from the liver powder.

2.3. Genomic PCR and genome walking

The amplification of neurotoxin gene(s) from the highly intact genomic DNA of the snake was carried out by using the AdvanTAge Genomic PCR method (Clontech, USA). This PCR employs two thermostable DNA polymerases, *Tth* DNA polymerase as primary polymerase and *Vent* DNA polymerase providing 3′-5′ proofreading activity. Primers used were X403 (forward) 5′-ATgAAAACTCTACTgACCTTg-3′ and X404 (reverse) 5′-gAgCTACTTgTTgCACAggTCCgT-3′, corresponding to the beginning of the signal peptide and the last few amino acid residues at the C-terminus of the short-chain neurotoxins of *P. textilis*, respectively. The amplification conditions involved a hot start at 94°C for 1 min followed by 30 cycles of 30 s at 94°C, 30 s at 60°C and 3 min at 68°C and a final extension at 67°C for 10 min after the final cycle. The PCR product was analyzed on aga-

*Corresponding author. Fax: (65)-7791453.
E-mail: bchjeya@nus.edu.sg

[1] Database submission: The nucleotide sequences of short-chain neurotoxin genes reported in this paper have been submitted to GenBank databases under the accession numbers AF204969–AF204973 and will be released upon citation.

0014-5793/00/$20.00 © 2000 Federation of European Biochemical Societies. All rights reserved.
PII: S0014-5793(00)01549-0

rose gel and the appropriate bands were excised and purified (QIAquick gel extraction method, Qiagen, USA) followed by subcloning into pT-Adv vectors (Clontech, USA).

To determine the 5′ and 3′ ends of the sequenced neurotoxin gene, the Universal GenomeWalker method (Clontech, USA) was used according to the manufacturer's protocols. In brief, the procedure involved the construction of adapter-ligated libraries made by separate restriction digestion of the highly intact genomic DNA with *Dra*I, *Eco*RV, *Pvu*II, *Sca*I and *Stu*I, followed by the ligation of these digested fragments with blunt ends to special adapters provided in the kit. The adapter-ligated libraries were used as templates. To amplify the 5′ end of neurotoxin gene, adapter primer 1 (AP1, 5′-gTAATACgACTCACTATAgggC-3′; forward) and 22 mer gene specific primers X464 (5′-gTCCAggCACATgATTgTCACC-3′; reverse) were used in the primary PCR, and subsequently, the nested adapter primer 2 (AP2, 5′-ACTATAgggCACgCgTggT-3′; forward) and primer X464 were used in the secondary PCR. Likewise, the combination of primers used for amplification of the 3′ end of the neurotoxin gene, included a 24 mer gene specific primer X443 (5′-AggggATgTggTACTTCTTgCCCT-3′; forward) and AP1 (reverse) for primary PCR, followed by X443 and AP2 for secondary PCR. Cycling parameters for primary PCR for the amplification of both 5′ and 3′ ends of neurotoxin gene were: hot start at 94°C for 1 min, then 15 cycles of 94°C for 25 s, 68°C for 3 min; followed by 25 cycles of 94°C for 25 s, 65°C for 3 min; and 65°C for an additional 7 min after the final cycle. The secondary PCR was carried out as hot start at 94°C for 1 min followed by five cycles of 94°C for 25 s, 70°C for 3 min and then 25 cycles of 94°C for 25 s, 65°C for 3 min with a final extension cycle at 65°C for 7 min. The PCR products were analyzed on agarose gel and the appropriate bands excised and purified before subcloning into pT-Adv vectors (Clontech, USA). The inserts were sequenced on both strands with M13 and sequence specific forward and reverse primers.

2.4. Sequence analysis and phylogenetic analysis

All double-stranded sequencing were performed with the ABI Prism Dye Terminator Cycle Sequencing protocol (Perkin-Elmer Biosystems, USA) using an automated DNA sequencer (Applied Biosystems, Model 373). The clones were completely sequenced on both strands using vector specific (M13 forward and reverse primers) and gene specific primers where appropriate.

Nucleotide sequence homology searches on GenBank databases (National Center for Biotechnology Information, USA) were performed using the BLAST program [12]. DNA sequence alignments were carried out using the DNASTAR software package (DNASTAR, USA). The number of nt substitutions per site (K_N) in the 5′, 3′ non-coding regions and introns, the numbers of nt substitutions per synonymous site (K_S) and non-synonymous site (K_A) in the protein-coding region were computed for all the pairs of *P. textilis* short-chain neurotoxin genes according to the method I of Nei and Gojobori [13]. A synonymous site is a site of the codon at which base substitution causes no amino acid change. A non-synonymous site is a site of codon at which base substitution causes an amino acid change.

Phylogenetic analysis was performed by using the program MegAlign [14].

2.5. Primer extension analysis

Primer extension analysis was performed to locate the transcription initiation site (TIS) of the neurotoxin gene. Total RNA was isolated from the *P. textilis* venom glands by using the guanidinium isothiocyanate method [15]. 10 pmol of primer SPEXT (5′-ggCACATgATTgTCACCAT-3′) was labeled at the 5′-end by T4 polynucleotide kinase using [γ-^{32}P]ATP (specific activity > 5000 Ci/mmol). The ^{32}P-labeled primers purified by ethanol precipitation were hybridized with 10 μg total RNA in a total volume of 5 μl of (1×) hybridization buffer (150 mM KCl, 10 mM Tris–HCl, pH 8.3 and 1 mM EDTA) at 65°C for 90 min and then allowed to cool to room temperature slowly. Then 5 μl reverse transcriptase (RT) mixture (1 mM of each of the four dNTPs, 10 U rRNasin and 50 U Mu-MLV RT in 2×Promega RT buffer) was added. The reverse transcription reaction was carried out at 42°C for 1 h, and was terminated by adding 140 μl RNase reaction mixture (100 μg/ml salmon sperm DNA, 20 μg/ml RNase A) and incubating for 15 min at 37°C. Sodium acetate (3 M, 15 μl) was added and the samples were extracted with phenol/chloroform/isoamyl alcohol (25:24:1), precipitated with ethanol and resuspended in 4 μl of sequencing gel loading buffer for analysis in an 8% acrylamide plus 7 M urea gel. A set of dideoxynucleotide DNA sequencing reaction using ^{32}P-labeled primer SPEXT and 30 fmol of the pT-Adv template harboring the 5′ flanking region of the neurotoxin gene were generated by cycle sequencing (Perkin-Elmer-Cetus, USA) and were used as size markers.

2.6. Promoter activity analysis

The promoter activity of the short neurotoxin gene was determined by measuring the activity of chloramphenicol acetyltransferase (CAT) using pMAMneo-CAT reporter vector (Clontech, USA) and by measuring the amount of mRNAs which encode CAT using reverse transcription followed by Real-Time PCR (Perkin-Elmer Biosystem, USA) using CAT gene specific primers (forward: 5′-gCCgCTggCgATTCAg-3′ and reverse: 5′-TTCATTAAgCATTCTgCCgACAT-3′) and the FAM labeled *Taq*man probe (5′-TCATCATgCCgTCTgTgATggCTTC-3′).

2.7. Subcloning of 5′ flanking region into the reporter vector of pMAMneo-CAT

The promoter region of short neurotoxin gene was first amplified by PCR using primers, X532 (forward; 5′-CTgCATATgAAAgTCATCAgTAAg-3′) and X533 (reverse; 5′-CTAgCTAgCCTTgCAATCTTCTCT-3′) containing *Nde*I and *Nhe*I restriction sites respectively. The PCR product digested by *Nde*I and *Nhe*I was ligated to linearized pMAMneo-CAT after the Mouse Mammary Tumour Virus (MMTV) promoter region in the plasmid was removed using *Nde*I and *Nhe*I. The recombinant plasmids were transformed into *Escherichia coli*, DH5α and putative positive clones were examined both by the size of the insert and by DNA sequencing. The neurotoxin promoter-CAT recombinant construct obtained was designated SF-pMAMneo-CAT. Likewise, the recombinant plasmid used for negative control (SR-pMAMneo-CAT) was constructed, by inserting the promoter region of short neurotoxin gene in a reverse direction in the reporter vector. Primers used in this case were: X534 (forward) 5′-CTggCTAgCAAAgTCATCAgTAAg-3′ and X535 (reverse) 5′-CTgCATATgCTTgCAATCTTCTCT-3′.

2.8. CAT enzyme assay

The promoter activity study was carried out using a mammalian host cell, HepG2. An equal amount of constructs SF-pMAMneo-CAT and SR-pMAMneo-CAT were transfected into early log phase HepG2 by using the calcium phosphate method [16]. Transfection efficiencies for all samples tested were normalized by cotransfecting the pSV-βGal plasmid followed by β-galactosidase assay. Appropriate amounts of cell lysates containing the transiently expressed CAT gene were used for CAT enzyme assays according to the procedures supplied by Promega. Briefly, the reaction was carried out by mixing the cell lysate and the substrates, [^{14}C]chloramphenicol and *n*-butyryl coenzyme A in a final volume of 125 μl (in buffer 0.25 M Tris–HCl, pH 8.0). The mixture was incubated at 37°C for 1 h and the reaction was terminated with 500 μl ethyl acetate. The upper, organic phase obtained after spinning was transferred, dried and resuspended in 30 μl ethyl acetate. 10 μl of each sample was spotted onto a silica gel TLC plate and chromatographed in a developing chamber pre-equilibrated for 1 h with chloroform:methanol mixture (97:3). The TLC plate was then subjected to autoradiography.

2.9. Real-Time PCR analysis

A quantitative assay on the expression of mRNAs from the CAT gene under the control of both MMTV (positive control) and the neurotoxin promoters was carried out separately by RT-PCR according to the procedures described by Perkin-Elmer Applied Biosystems (USA), who supplied the reagents and the Sequence Detection system (Model 7700). The C_T values were determined for different dilutions of the template. An amplicon of 67 bp was amplified in each case. The negative controls included a no template control (NTC), RNA from HepG2 cells and the neurotoxin gene promoter subcloned in the reverse orientation (SR-pMAMneo CAT).

3. Results

3.1. Structure and organization of P. textilis short neurotoxin genes

PCR was used to amplify the neurotoxin genes. The strategies used in the cloning are outlined schematically in Fig. 1A.

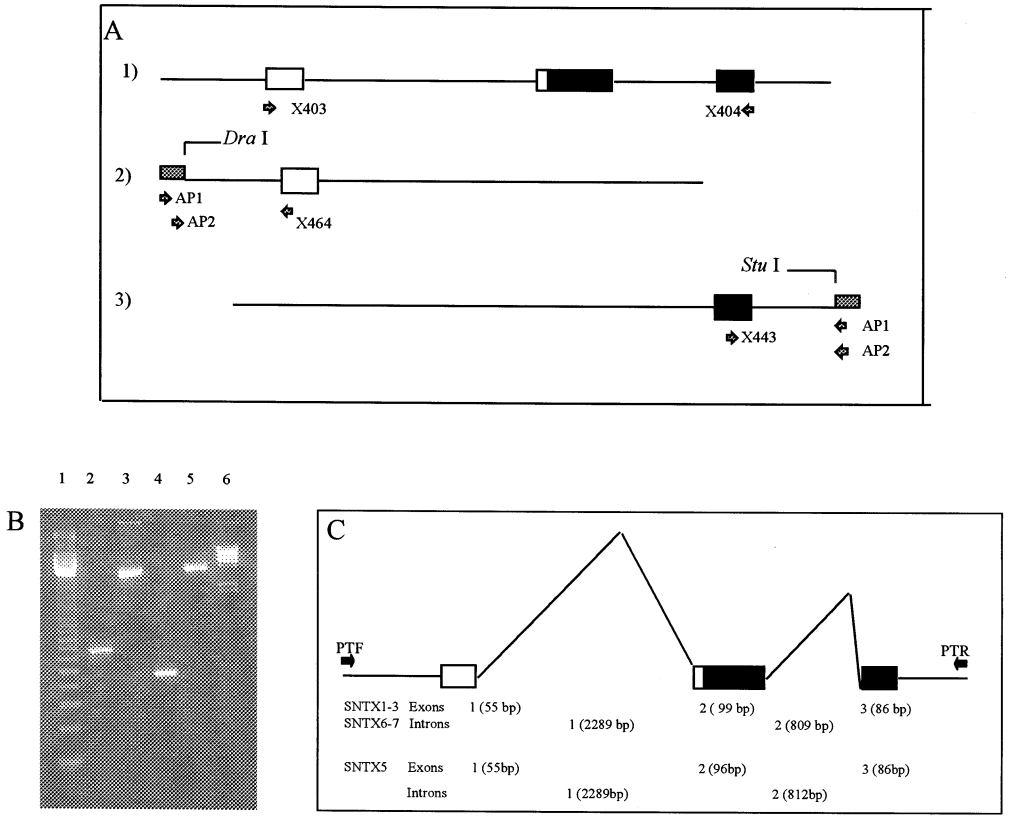

Fig. 1. The structure and strategies for the amplification of the short-chain neurotoxin genes of *P. textilis*. A: Strategies for PCR. Primers used and their relative positions on the genes are shown. A1: Genomic PCR of coding region; A2: genome walking PCR of 5′ non-coding region; A3: genomic PCR of 3′ non-coding region. B: Electrophoresis of the products of genomic and genome walking PCR. Lane 1, 123 bp ladder as size marker; lane 2, 5′ end amplified from *Dra*I library; lane 3, genomic PCR of the gene between X403 and X404 primers; lane 4, 3′ end amplified from *Stu*I library; lane 5, genomic PCR of *Pt-sntx* gene using primers PTF and PTR; lane 6, λ*Hin*dIII DNA size markers. C: The organization of the short-chain neurotoxin genes in *P. textilis*. The exons are shown as boxes and introns as thin lines. The signal peptide-coding region is represented by open boxes while the mature protein-coding region is represented by closed boxes.

All the primers used were designed based on the *P. textilis* short neurotoxin cDNAs. In order to clone the whole protein-coding region, primers X403 and X404 were used, which corresponded to the beginning of the signal peptide and the last few amino acid residues at the C-terminus of the mature short neurotoxin. Amplification using primers X403 and X404 gave a fragment of about 3.9 kb (Fig. 1B) on agarose gel. PCR products were then subcloned into pT-Adv vector. Out of 70 such clones, five short-chain neurotoxin genes were identified, with an almost equal frequency (20%) of appearance for each of them. These genes represented the mRNAs [3] identified in the venom gland of the same snake. Thus, they are functional and do not represent any pseudogene(s).

Nucleotide sequence comparison of each of the genes and the cDNA isoforms resulted in the assignment of intron/exon junctions. They were found to contain two introns around 2.2 kb and 0.8 kb. Both introns started with gT and end with Ag. The first intron was found between nt 87 and 2375 which interrupted the neurotoxin leader sequence region (Fig. 2) at the same position as in Ec [7], Cbt [8], Cbt b [9] and Ntx1 from *N. naja sputatrix* [10]. The second intron occurred between nt 2475 and 3283 within the coding region (Fig. 2). This splice junction was found to split the codon for Asn in the deduced protein sequence between the first and second base (Fig. 2) resulting in three exons. The first exon encoded most of the signal peptide (18 amino acid residues). The second exon encoded the C-terminus of the signal peptide (three amino acid residues) and the N-terminal half of the mature neurotoxin and the third exon encoded the remainder of the mature neurotoxin.

The 5′ end of the short neurotoxin gene was obtained by genome walking using a *Dra*I library of the *P. textilis* genome and primers X464 and AP1 for primary PCR followed by primers X464 and AP2 for secondary PCR. The 0.8 kb fragment obtained was subcloned and sequenced. Fifty such clones were sequenced and all showed identical nt sequences.

Likewise, the 3′ end of the neurotoxin gene was elucidated by subcloning and sequencing of a 0.4 kb PCR fragment from a *Stu*I library using primers X443 and AP1 followed by nested PCR involving X443 and AP2. Once the complete nt sequences of short neurotoxin genes of *P. textilis* were established, two new primers PTF (5′-AAAgTCATCAgTAAgAAgATC-3′) and PTR (5′-CCTgAAgTTgAggTTTCACTg-3′) corresponding to the 5′ and 3′ end of the gene respectively, were synthesized and all the five neurotoxin genes (5 kb each) were amplified by PCR from the genomic DNA of *P. textilis* (Fig. 1B). After subcloning the fragments into vector pT-Adv vector, 12 positive clones were sequenced by using vector specific primers and appropriate gene specific primers. The results reconfirmed the existence of five copies of the short neurotoxin gene in *P. textilis*.

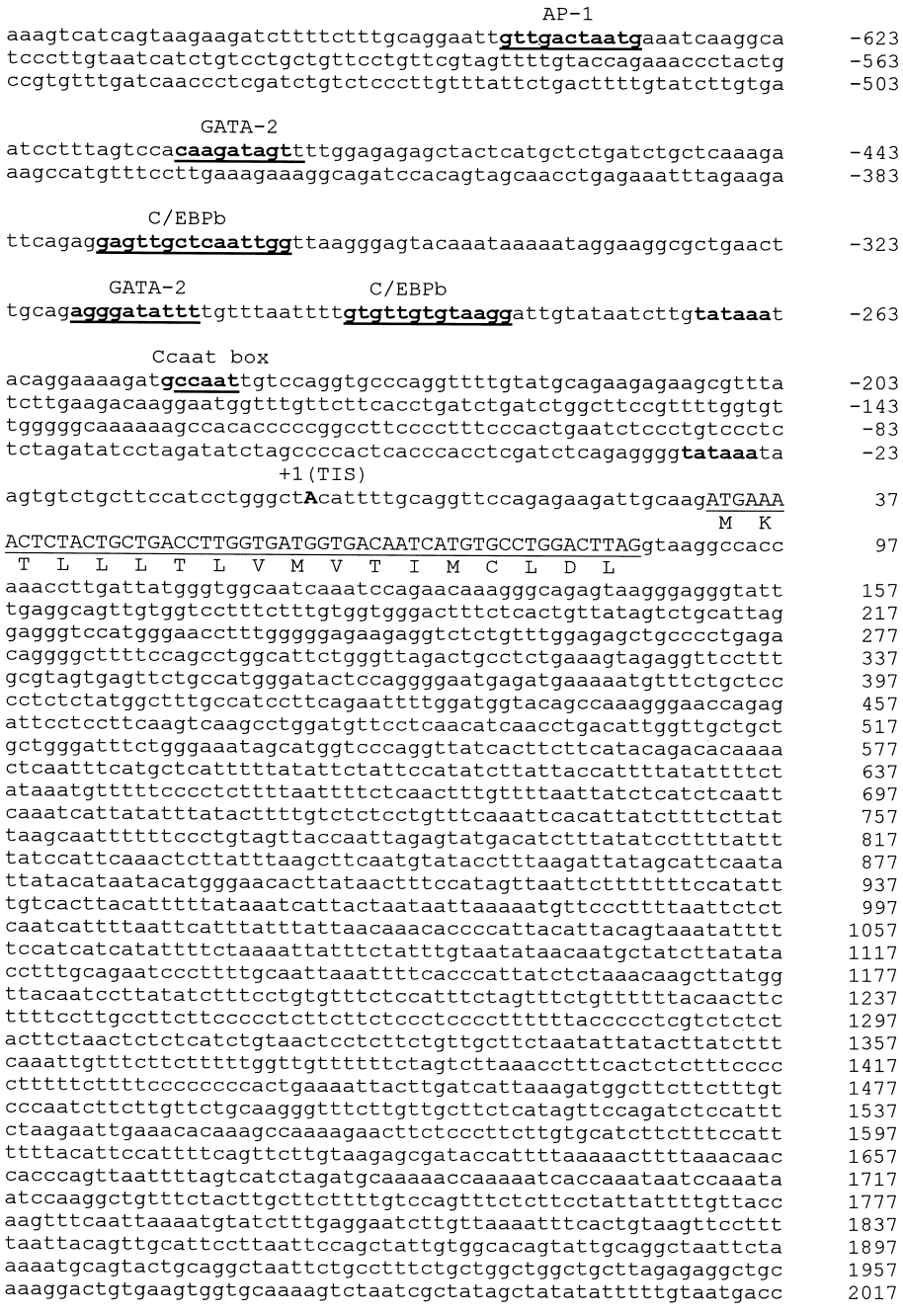

Fig. 2. Complete nt sequences of the *Pt-sntx1* gene. TIS is numbered by +1. Exons are represented in capitals and the deduced amino acid sequences are shown below the nt sequence by single-letter codes in upper case, while the amino acid sequence of the mature neurotoxin is boldface. Intron sequences, 5′ and 3′ non-coding regions of the neurotoxins gene are given in lowercase letters. The stop codon is indicated by an asterisk. The consensus sequence for polyadenylation, AATAAA, at the 3′ flanking region of the gene is also highlighted. Consensus sites for some putative transcription factors are indicated at the 5′ end of the gene. TATA box motifs are also shown. Alignment of the mature toxin coding regions of the *Pt-sntx2, 3, 6* and *7* gene sequences and the Pt-sntx5 cDNA sequence are included in the figure. The three bases missing in the cDNA of *Pt-sntx5* are indicated by ●. The amino acid sequences of the putative protein encoded by the five alleles are represented by uppercase single-letter code. The *Pt-sntx1* was taken as the consensus sequence and any variations from this sequence (both nt and amino acid sequences) are indicated for the other alleles. nt residues that match the *Pt-sntx1* sequence are indicated by a hyphen '-', while amino acid residues that are different from *Pt-sntx1* are indicated in boldface uppercase single-letter code.

3.2. Sequence analysis and phylogenetic analysis

Comparisons of the nt sequences and the deduced amino acid sequences of *P. textilis* short neurotoxins (Fig. 2) clearly showed that the nt substitutions occurred mainly in the protein-coding region except for the signal peptide-coding region. This suggests that the mature protein-coding region is much more variable than the introns and the 5′ and 3′ non-coding regions. The sequences were further analyzed by considering the exons together as a group and the introns with 5′ and 3′ non-coding regions as another group. Tables 1A and 1B shows the data from 10 such possible combinations among the five neurotoxin genes. The K_N/K_S values determined ac-

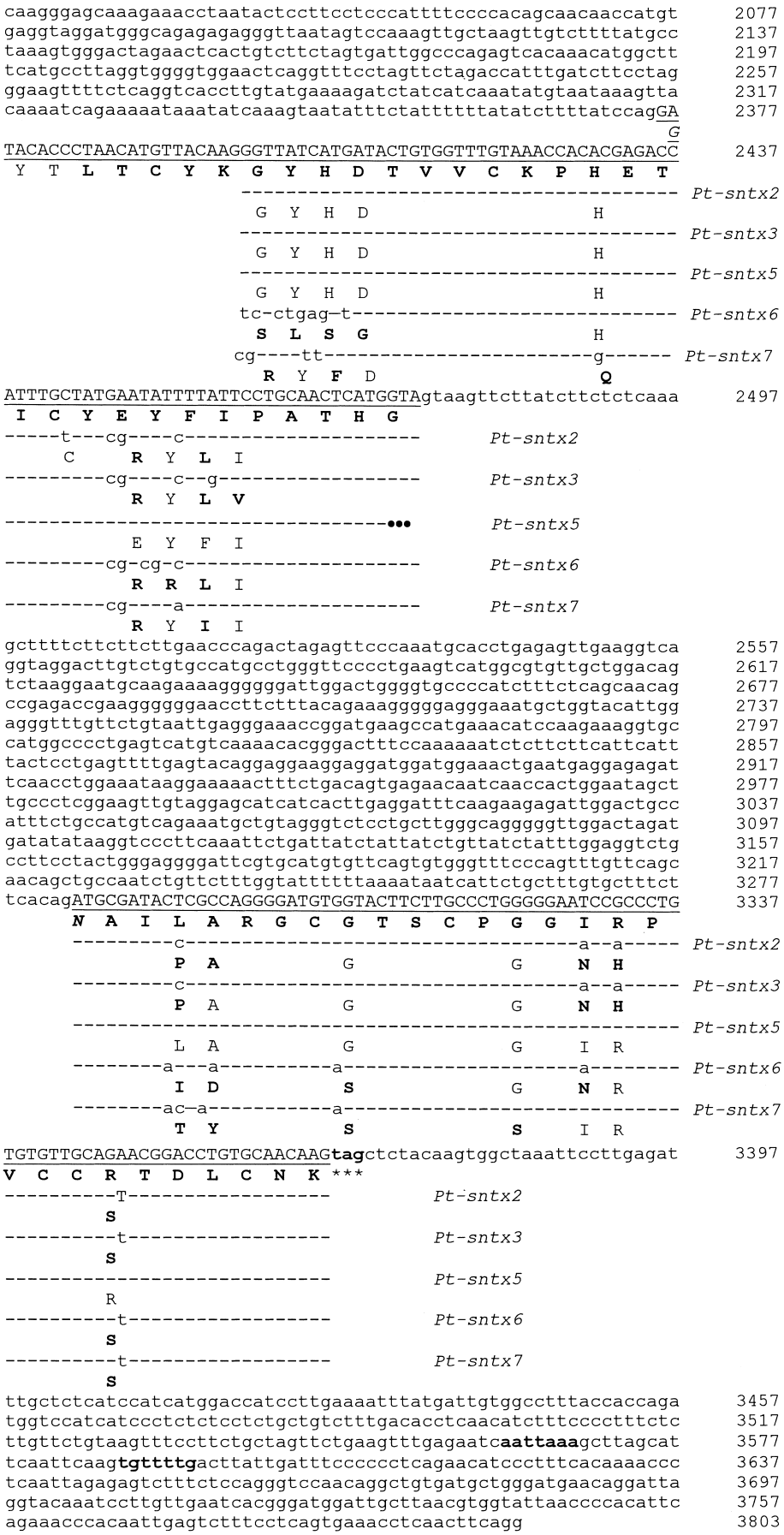

Fig. 2 (*continued*).

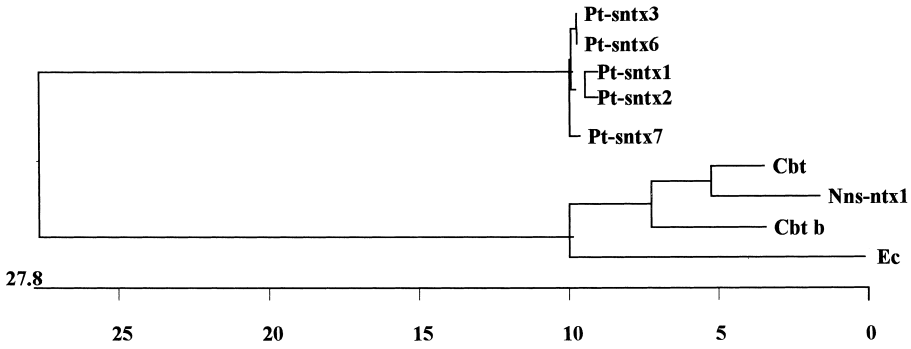

Fig. 3. Phylogenetic analysis of neurotoxin genes. The cladogram was constructed by the program MegAlign. The length of each pair of branches represents the distance between sequence pairs. The units at the bottom of the tree indicate the number of substitution events. The complete gene sequences of Ec, erabutoxin c from *L. semifasciata* [7]; Cbt, cobrotoxin from *N. naja atra* [8]; Cbt b, cobrotoxin b from *N. naja atra* [9]; Nns-ntx1, NTX1 from *N. naja sputatrix* [10]; *Pt-sntx1*, *2, 3, 6* and *7* (this study) were used in this analysis.

cording to Nei and Gojobori [13] were found to be less than 1. This indicates that nt substitutions have occurred more frequently in the protein-coding region than in the other parts of the gene. The K_A/K_S values [13] are however greater than 1. This is due to the fact that nt substitutions in the protein-coding regions tend to cause amino acid changes. These two features are in accordance with those found in snake venom phospholipase A_2 cDNAs and genes as well as in cDNAs encoding serine protease [6,17,18].

Since exons and introns seemed to have undergone changes in different ways, we constructed a phylogenetic tree (Fig. 3) based on the whole gene sequence to examine the overall evolutionary status of the gene. The short-chain neurotoxin genes of *P. textilis* formed a unique group from those of other land and sea snakes and appear to have also evolved earlier than the other short neurotoxins. This observation is consistent with our previous report based on the molecular properties of the proteins [3].

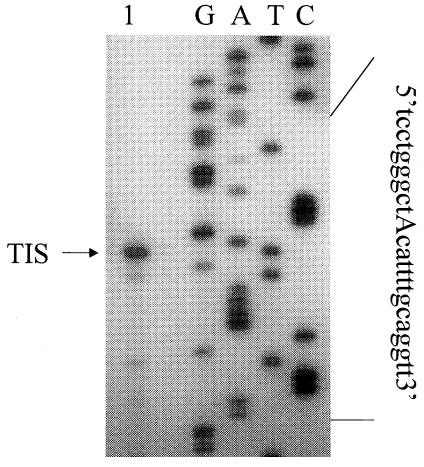

Fig. 4. Determination of the TIS. The primer SPEXT was end-labeled by [γ-^{32}P]ATP and annealed to the total RNA from the venom glands of *P. textilis*. The annealed template/primer was extended with RT. The major product from the venom gland RNA is indicated by an asterisk (lane 1). Lanes G, A, T and C represent a set of dideoxy nt DNA sequencing reactions generated with primer SPEXT and pT7 blue harboring the 5′ non-coding region of the gene as template.

3.3. TIS and regulatory elements of the neurotoxin gene

The determination of the TIS was conducted by primer extension analysis using RT and ^{32}P-labeled primer SPEXT, which is complementary to the 5′ end of the neurotoxin mRNA sequence. One major product arising from adenosine (A) at nt+1 was observed (Figs. 2 and 4). This site corresponded to the TIS previously reported for erabutoxin c [7], cobrotoxin [8], cobrotoxin b [9] and TIS1 in *N. naja sputatrix* ntx1 [10]. This is supported by the presence of a putative TATA box and a CAAT box at 30 and 250 bp upstream of TIS.

A search for the regulatory elements at the 5′ flanking region of neurotoxin gene using TFSEARCH (Yutaka Akiyama-http://www.rwcp.or.jp/papia/) showed the presence of binding sites for the transcription factors, AP-1 [19], C/EBPb [20,21] and GATA-2 [22].

3.4. Promoter activities

Negative control SR-pMAMneo-CAT showed no activity of CAT while SF-pMAMneo-CAT indicated CAT activity. The relative quantitative assay by Real-Time PCR (Table 1B) also indicated that the promoter is active with a C_T value of 18 for the SF-pMAMneo-CAT and 29 for the MMTV-pMAMneo-CAT. Total RNA at 25 ng was required for the optimal determination of C_T values.

4. Discussion

4.1. Structure of the short neurotoxin gene

The DNA sequence analysis of the short neurotoxin gene of *P. textilis* shows a typical eukaryotic gene structure. Two introns (2289 and 809 bp) split the coding region into three exons (86, 99 and 88 bp). The whole sequence of neurotoxin gene contains 4486 bp with 5′ and 3′ regions of 681 and 434 bp respectively. The nt A in the sequences of tAcattt numbered +1 was assigned as TIS by primer extension analysis. Both translation start codon and translation stop codon can be located on the gene sequence, as ATG and TAG, respectively. The 3′ end of the neurotoxin gene contains a polyadenylation signal AATTAAA at nts 3562–3568. Also identified is the sequence TGTTTTG located immediately downstream of the poly A signal, which regulates the half-life of neurotoxin mRNA by acting as mRNA degradation signal [23]. The promoter activity can be ascribed to the pres-

Table 1A
Calculation of the K_N, K_S, K_A, K_N/K_S and K_A/K_S values

Combinations	K_N	K_A	K_S	K_N/K_S	K_A/K_S
Pt-sntx 1 and 2	0.0074	0.0722	0	–	–
Pt-sntx 1 and 3	0.0062	0.0720	0	–	–
Pt-sntx 1 and 6	0.0074	0.1664	0.0117	0.63	14.22
Pt-sntx 1 and 7	0.0070	0.1352	0.0089	0.78	15.19
Pt-sntx 2 and 3	0.0070	0.0086	0.0174	0.40	0.49
Pt-sntx 2 and 6	0.0070	0.1996	0.0294	0.24	6.79
Pt-sntx 2 and 7	0.0070	0.1153	0.0268	0.26	4.30
Pt-sntx 3 and 6	0.0039	0.2000	0.0116	0.33	17.24
Pt-sntx 3 and 7	0.0058	0.1255	0.0088	0.66	14.26
Pt-sntx 6 and 7	0.0062	0.1397	0.0206	0.30	6.78

The 10 possible combinations within the venom short-chain neurotoxin alleles of *P. textilis* were analyzed according to Nei and Gojobori [13].

ence of the TATA box motif found 30 bp immediately upstream of TIS, as well as to the CCAAT box motif located 250 bp away from TIS. Another TATA box motif can also be identified upstream of the CCAAT box. However it does not seem to initiate the transcription of the neurotoxin gene in the venom gland. The promoter also shows a putative binding site, GTTGACTAATG at nt −634 to −644, for transcription factor AP-1 [19] and two binding sites for GATA-2 (nt −489 to −481 and −317 to −308). The GATA family recognizes the GATA motif and among them six members, GATA-1 to GATA-6 were identified. It has been suggested that the GATA motif negatively controls transcription [22]. Two c/EBPb elements are also present in the promoter region, at nt −375 to −361 and −296 to −284.

The similar frequency of appearance of all *Pt-sntx* gene sequences from a pool of 70 clones sequenced, suggests that all *Pt-sntx* genes were identified with the least possibility of missing out any of them. A total of five genes were identified. However, the cDNA cloning resulted in the identification of six isoforms of mRNAs [3]. It is therefore reasonable to assume that one of the six mRNAs arise from any one of the five genes, possibly by alternative splicing. In fact such a possibility does exist on the *Pt-sntx1* gene sequence at nt 2472, 3 bp upstream of the intron 2 splice site (nt 2475). Splicing of intron 2 at nt 2472 and nt 2475 separately has actually produced the two mRNAs named previously Pt-sntx1 and Pt-sntx5, respectively (Fig. 2). Except for these two short neurotoxin mRNAs, all the others have their independent genes in the genome of *P. textilis*. This excludes the possibility of RNA editing as the source of the multiplicity of short neurotoxins in the snake venom, unlike the case in long neurotoxin α-bungarotoxin suggested by Liu et al. [24]. Hence this forms the first report on the existence of alternative splicing for the expression of two types of short-chain neurotoxin mRNAs from a single gene.

4.2. Further analysis of the short neurotoxin gene

The exon–intron organization of the *P. textilis* neurotoxin gene is similar to that of the erabutoxin c from *L. semifasciata* [7], cobrotoxin and cobrotoxin b from *Naja naja atra* [8,9] and NTX-1 from *N. naja sputatrix* [10]. The nt sequence of the exons and introns of the *P. textilis* gene also shows reasonable similarity to them. The 5′ and 3′ non-coding regions and the sequences near the splice donor/acceptor sites in introns were found to be highly conserved whereas the conservation level of sequences in the middle part of each intron was found to be relatively lower. The exons, however, are more variable, with the exception of the signal peptide-coding region.

This is further confirmed by the mathematical analysis in which K_N/K_S is less than 1 and K_A/K_S is more than 1. First, K_N values for introns, 5′ and 3′ non-coding regions are approximately one-fifth to half of the K_S values for the protein-coding regions in these pairs of genes except for pairs of *Pt-sntx* 1 and 2 and 1 and 3 (Table 1A). This shows that both introns and 5′ and 3′ non-coding regions are unusually conserved as compared to protein-coding regions. The protein-coding regions of these short-chain neurotoxins have therefore evolved with higher substitution rates than the introns and 5′ and 3′ non-coding regions. This is in contrast to what was observed for genes that evolve spontaneously under natural selection where K_N values are normally greater than K_S values. The ratio of K_A/K_S is also much higher than 1 with ratios of 4.30–17.24 for most of PT-sntxs except for the pairs of *Pt-sntx* 1 and 2, 1 and 3 and 2 and 3. This shows that the substitutions occurring in non-synonymous sites are much higher than those in synonymous sites. For a gene which has evolved according to neutral theory, the K_A values are expected to be several times lower than their K_S values where non-synonymous nt substitution was suppressed as compared to synonymous nt substitution due to functional constraints [25,26]. Thus it is apparent that the non-synonymous nt substitutions causing amino acid changes have occurred much more frequently in *P. textilis* short-chain neurotoxin genes and it is possible that these genes have evolved through a different mechanism so as to bring about accelerated amino acid substitutions.

Table 1B
P. textilis short-chain neurotoxin gene promoter activity as determined by Real-Time PCR

Sample/promoter	Total RNA quantity (ng)	Mean threshold cycle value (C_T)	S.D.
NTC	–	39.75	±0.18
MMTV-pMAMneo-CAT	25	29.18	±0.17
SR-pMAMneo-CAT	25	38.86	±0.14
SF-pMAMneo-CAT	25	18.05	±0.09
HepG2	25	39.30	±0.16

The mean C_T values and the standard deviation observed for assays carried out for each sample in triplicate are shown.

Similar accelerated evolution was observed in the cDNAs encoding crotalinae venom serine protease isozymes [18], and cDNAs and genes encoding crotalinae snake venom phospholipase A_2 isozymes [6,17]. Ohno et al. [27,28] suggested that the molecular evolution of snake toxins involves a mechanism of accelerated evolution accompanied by functional diversification of their products. All the features observed in *P. textilis* short-chain neurotoxin could be accounted for by this mechanism. The *Pt-sntx* genes could have evolved by gene duplication followed by accelerated evolution at its mature protein coding region to produce proteins with diverse functions under adaptive pressure.

Since the adaptive pressure has always been on the venomous snakes from their habitats, we assume that all the genes encoding the proteins present in the *P. textilis* venom must have been subjected to a similar accelerated evolution as described by Ohno et al. [27]. Daltry et al. [29] also have shown that changes in habitat and their preys could contribute to adaptive evolution.

A recent method of Real-Time PCR introduced in this study provides an alternative, faster and efficient means of determining the promoter activity of genes. In this paper we report a set of CAT gene specific primers and probes which can be used for such studies. We have carried out a simpler relative quantitation assay to determine the *Pt-sntx* promoter activity and the results confirm that the *Pt-sntx* gene contains a functional promoter and that none of its alleles are pseudogenes. It is also interesting to note that the *Pt-sntx* promoter functioned as a stronger promoter (at least 1.5 fold higher) than the MMTV promoter under the experimental conditions used in this study.

Acknowledgements: This work was supported by the research Grant RP991314. Ms N.L.G. received a research scholarship from the National University of Singapore.

References

[1] Servent, D., Winckler-Dietrich, V., Hu, H.Y., Kessler, P., Drevet, P., Bertrand, D. and Menez, A. (1997) J. Biol. Chem. 272, 24279–24286.
[2] Endo, T. and Tamiya, N. (1991) in: Snake Toxins (Harvey, A.L., Ed.), pp. 165–222, Pergamon, New York.
[3] Gong, N.L., Armugam, A. and Jeyaseelan, K. (1999) Eur. J. Biochem. 265, 982–989.
[4] Afifiyan, F., Armugam, A., Gopalakrishnakone, P., Tan, C.H. and Jeyaseelan, K. (1998) Toxicon 36, 1871–1885.
[5] Jeyaseelan, K., Armugam, A., Lachumanan, R., Tan, C.H. and Tan, N.H. (1997) Biochim. Biophys. Acta 1380, 209–222.
[6] 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.
[7] Fuse, N., Tsuchiya, T., Nonomura, Y., Menex, A. and Tamiya, T. (1990) Eur. J. Biochem. 193, 629–633.
[8] Chang, L.-S., Lin, J., Chou, Y.-C. and Hong, E. (1997) Biochem. Biophys. Res. Commun. 239, 756–762.
[9] Chang, L.-S., Chou, Y.-C., Lin, S.-R., Wu, B.-N., Lin, J., Hong, E., Sun, Y.-J. and Hsiao, C.-D. (1997) J. Biochem. 122, 1252–1259.
[10] Afifiyan, F., Armugam, A., Tan, C.H., Gopalakrishnakone, P. and Jeyaseelan, K. (1999) Genome Res. 9, 259–266.
[11] Blin, N. and Stafford, D.W. (1976) Nucleic Acids Res. 3, 2303–2306.
[12] Altschul, S., Thomas, L., Alejandro, A., Jinghui, Z. and Zheng, Z. (1997) Nucleic Acids Res. 25, 3389–3402.
[13] Nei, M. and Gojobori, T. (1986) Mol. Biol. Evol. 3, 418–426.
[14] Saitou, N. and Nei, M. (1987) Mol. Biol. Evol. 4, 406–425.
[15] Chomczynski, P. and Sacchi, N. (1987) Anal. Biochem. 162, 156–159.
[16] 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.
[17] Nakashima, K., Nobuhisa, I., Deshimaru, M., Nakai, M., Ogawa, T., Shimohigashi, Y., Fukumaki, Y., Hattori, M., Sakaki, Y. and Hattori, S. et al. (1995) Proc. Natl. Acad. Sci. USA 92, 5605–5609.
[18] Deshimaru, M., Ogawa, T., Nakashima, K.-I., Nobuhisa, I., Chijiwa, T., Shimohigashi, Y., Fukumaki, Y., Niwa, M., Yamashina, I., Hattori, S. and Ohno, M. (1996) FEBS Lett. 397, 83–88.
[19] Pennypacker, K.R., Hong, J.S. and McMillian, M.K. (1994) FASEB J. 8, 475–478.
[20] Landschulz, W.H., Johnson, P.F., Adashi, E.Y., Graves, B.J. and McKnight, S.L. (1988) Genes Dev. 2, 786–800.
[21] Landschulz, W.H., Johnson, P.F. and McKnight, S.L. (1988) Science 240, 1759–1764.
[22] Imagawa, S., Yamamoto, M. and Miura, Y. (1996) Acta Haematol. 95, 248–256.
[23] Shaw, G. and Kamen, R. (1986) Cell 46, 659–667.
[24] Liu, L.-F., Chang, C.-C., Liau, M.-L. and Kuo, K.-W. (1998) Nucleic Acids Res. 26, 5624–5629.
[25] Kimura, M. (1983) The Neutral Theory of Molecular Evolution, Cambridge University Press, Cambridge.
[26] Nei, M. (1987) Molecular Evolutionary Genetics, Columbia University Press, New York.
[27] Ohno, M., Menez, R., Ogawa, T., Danse, J.M., Shimohigashi, Y., Fromen, C., Ducancel, F., Zinn-Justin, S., Le du, M.H., Boulain, J.C., Tamiya, T. and Menez, A. (1998) Prog. Nucleic Acid Res. Mol. Biol. 59, 307–364.
[28] Kihara, H., Uchikawa, R., Hattori, S. and Ohno, M. (1992) Biochem. Int. 28, 895–903.
[29] Daltry, J.C. (1996) Nature 379, 537–540.