Structurally conserved α-neurotoxin genes encode functionally diverse proteins in the venom of *Naja sputatrix*[1]

Kandiah Jeyaseelan*, Siew Lay Poh, Ramkishen Nair, Arunmozhiarasi Armugam

Department of Biochemistry, Faculty of Medicine, National University of Singapore, 8 Medical Drive, Singapore 117597

Received 16 July 2003; revised 10 September 2003; accepted 11 September 2003

First published online 26 September 2003

Edited by Takashi Gojobori

Abstract The structure and organization of the genes encoding the long-chain neurotoxins and four other isoforms of weak neurotoxins in the venom of *Naja sputatrix* are reported. The genes contained three exons interrupted by two introns, a structure similar to other members of the three-finger toxin family. The proteins encoded by these genes, however, show varied affinity towards nicotinic acetylcholine receptors. Phylogenetic analysis of these genes showed that the weak neurotoxin gene is confined to a distinct group. We also observe that specific mutations of the gene provide the diversity in function in these toxins while maintaining a common structural scaffold. This forms the first report where the molecular basis of evolution of postsynaptic neurotoxins from an ancestral gene can be demonstrated using the same species of snake.

Key words: Nicotinic acetylcholine receptor; α-Neurotoxin; Weak neurotoxin; Co-evolution; Accelerated evolution; *Naja sputatrix*

1. Introduction

Venoms of elapids (cobra, kraits, mambas and sea snakes) contain low molecular weight curaremimetic proteins called α-neurotoxins (α-NTXs). The α-NTXs from snake venoms are highly specific for the nicotinic acetylcholine receptors (nAChRs) in the postsynaptic membranes of vertebrate muscle and electric organs of *Torpedo* sp. These toxins prevent the binding of the chemical neurotransmitter acetylcholine, and thereby block nerve impulses leading to paralysis of the animal [1]. α-NTXs show a common 'three-finger' loop scaffold and interact with the same molecular target, nAChR. The primary function of the core is to maintain the overall structure of these toxins [2] while the structural plasticity of the three loops has been reported to accommodate the functional diversity [3]. Based on the receptor/toxin interaction, these postsynaptic NTXs have been classified into four groups [4]. Notably the classical α-NTXs, such as the short-chain NTXs (SNTX) from the Elapidae, which bind with high affinity to muscular nAChR only [5,6], constitute the members of group I. Group II consists of long-chain NTXs (LNTX) that bind with high affinity to both the muscular and some subtypes of neuronal receptors (α7, α8, and α9 [5,7]). α/κ-NTXs, which bind with high affinity to the neuronal-type nAChRs [8], belong to group III, while group IV comprises the unconventional NTXs which include the weak NTXs (WNTX). Although the α-NTXs generally bind to AChRs, the binding efficiency of the short, long and weak NTXs to these receptors varies among the subtypes of toxins [9] as well as their targets at the nAChRs. Of the long- and short-chain NTXs binding with high affinity to muscular-type nAChRs [5], only the long-chain toxins show high affinity to the neuronal α7 receptor [7]. The WNTXs, however, show very low affinity towards both the muscle and neuronal-type nAChRs [10,11]. Although the binding occurs in homologous regions of neuronal and muscular receptors, toxin binding does not involve identical determinants in the receptor [9].

It has been well documented that the primary function of the core region in the three-finger toxin is to maintain the overall structure while allowing structural plasticity at individual loops such as loop II for the recognition of the molecular targets [9,12–14]. Besides the conventional signature of four disulfide bridges in the three-finger toxin proteins, LNTX contains an extra cyclic loop in loop II while WNTX contains a similar cyclic loop in loop I of the protein, thus LNTX and WNTX form a new group, the five-disulfide bond three-finger toxin family. Insertion of an extra cyclic disulfide bond in SNTX has been found to increase its affinity to neuronal nAChRs [15] while removal of the cyclic loop in LNTX makes it behave like SNTX [7].

The toxic components of the venom of *Naja sputatrix* consist of phospholipase A_2 (PLA_2) enzymes, cardiotoxins and postsynaptic NTXs [16]. PLA_2 [17,18], cardiotoxins [19–21] and SNTXs [22,23] have been extensively studied in our laboratory. The cDNA and the functional properties of WNTXs (GenBank accession numbers: AF026891, AF086292, AF098923, AF098924) from *N. sputatrix* [10] have also been reported. In this report, we describe a new five-disulfide bridge-containing LNTX that has been purified from *N. sputatrix* venom. cDNA encoding LNTX has also been cloned and characterized (GenBank accession number: AF0026893). This LNTX has been found to bind with high affinity to the

*Corresponding author. Fax: (65) 67791453.
E-mail address: bchjeya@nus.edu.sg (K. Jeyaseelan).

[1] The genomic DNA sequences (WNTXs and LNTX) reported in this paper can be accessed from the GenBank with accession numbers AY081759 to AY081763 and AF026893 for the LNTX-7 cDNA sequence.

Abbreviations: LNTX, long-chain neurotoxin; WNTX, weak neurotoxin; SNTX, short-chain neurotoxin proteins; Their respective cDNAs are designated Lntx, Wntx and Sntx and their genes *Lntx*, *Wntx* and *Sntx*

0014-5793/03/$22.00
doi:10.1016/S0014-5793(03)01039-1

muscular and neuronal nAChR in comparison to another five-disulfide bridge-containing NTX, WNTX, which even though it contains part of the binding determinant of potent α-NTX, binds weakly to *Torpedo* and α7 neuronal nAChRs. In this report we also include the cloning and characterization of the genes encoding these five-disulfide bridge three-finger toxins and offer some plausible explanations for their functional divergence.

2. Materials and methods

2.1. RT-PCR and subcloning

Reverse transcription (RT) reactions were performed using 3 μg of total RNA prepared from venom glands of *N. sputatrix* using 40 ng of antisense primer X191, 5′-GCGGCGGAATTCTTTTTTTTTTTTT-TTTTT-3′. Reverse transcription followed by polymerase chain reaction (PCR) was carried out as described by Poh et al. [10]. The full-length cDNA was amplified using X191 (antisense primer) and X133 (5′-TCCAGAAAAGATCGCAAGATG-3′; sense primer [22]). The PCR products were ligated to pT7 Blue(R) vector and the recombinant plasmids were transformed into *Escherichia coli* JM109 [24]. Putative positive clones were sequenced using M13 forward and reverse universal primers by the dideoxy chain termination method on an automated DNA sequencer (Applied Biosystems, Model 3100, USA).

2.2. Isolation and characterization of LNTX

Lyophilized crude venom milked from a single *N. sputatrix* was reconstituted in 0.1 M ammonium acetate buffer, pH 4.5, centrifuged and subjected to Superdex G75, HR 10/30 (Pharmacia) for gel filtration. Fractions containing peptides in the range of 6000–7500 kDa were combined and further resolved by FPLC using Mono S HR 5/5 (Pharmacia Biotech), equilibrated with 50 mM ammonium acetate, pH 4.5. Bound peptides were eluted from the column with a linear gradient of NaCl in the same buffer. The final purification step for the peptides was carried out by RP-HPLC using an analytical Vydac 5 μm, C18 column (250×10 mm) with a gradient of 25–100% of solution B for 10 min (solutions A and B are 0.1% trifluoroacetic acid (TFA) in water and 90% acetonitrile in solution A respectively). The flow rate was maintained at 1 ml/min and the protein elution was monitored at 280 nm. All the proteins were analyzed on Tris-Tricine sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE). Mass spectrometric analysis was carried out on an ion spray mass spectral system, Micromass Platform II (Micromass, Altrincham, UK). All mass analyses were carried out in positive mode with ion spray voltage 4000 V, orifice voltage 30 V, and mass scanning range at 500–2000 amu. The amino acid composition of the hydrolyzed toxin was analyzed using an Applied Biosystems Model 130A Automatic analyzer equipped with an on-line 420A derivatizer for the conversion of the free amino acid into phenyl thiocarbamoyl derivatives. The N-terminal amino acid sequence of α-NTX was determined by automated Edman degradation using an Applied Biosystems 477A pulsed liquid phase sequencer equipped with an on-line PTH amino acid analyzer (120A).

2.3. Lethality assay (LD_{50})

Three groups (four mice each) of adult Swiss albino mice (18–20 g) were injected i.v. with 100 μl of isotonic solution as control and at least four doses of purified toxin and crude venom. The mice were observed for up to 48 h, and the LD_{50} (i.v.) values were calculated according to the Spearman–Karber method (WHO).

2.4. AChR binding assay

Binding assays were performed using [^{125}I]α-bungarotoxin ([^{125}I]-α-Bgtx, 210–250 Ci/mmol, Amersham) prepared as described by Saitoh et al. [25]. Chimeric α7 receptors were obtained by expressing the chimeric cDNA (α7-5HT$_3$) in HEK cells [7]. Varying amounts of toxins were incubated with 3 nM active sites of receptors and 5 nM of [^{125}I]α-Bgtx for at least 4 h as described by Poh et al. [10]. The curves were fitted with the empirical Hill equation. K_d values for *Torpedo* AChR were calculated according to Cheng and Prusoff [26], and protection constants (K_p) for α7 receptors were calculated according to Weber and Changeux [27].

2.5. Amplification of five-disulfide bridge α-NTX genes

Intact genomic DNA was isolated from the liver of a single Malayan spitting cobra (*N. sputatrix*). The amplification of NTX genes was carried out on an uncloned library created using the AdvanTAge Genomic PCR kit (BD Biosciences, Clontech, USA). Primers for PCR were designed based on the conserved regions in the structural gene of *N. sputatrix* NTX cDNA (AF026893). The primer sets used to clone the LNTX gene were: X133 (forward) 5′-TCCGAAAAAGATG-GAAGTG-3′ and X491 (reverse) 5′-ACTCAAGGACGTGTTG-GAAATG-3′ and those for WNTX gene(s) were X133 and X477 (reverse) 5′-TCGGTTGCATCTGTATGTTGA-3′. These primers corresponded to the beginning of the signal peptide and the last few amino acid residues at the C-terminus of LNTX and WNTX, respectively. The PCR amplification involved a hot start at 94°C for 1 min followed by 30 cycles of 94°C for 30 s, 50°C for 30 s and 68°C for 3 min and a final extension at 65°C for 10 min. The PCR products were subcloned into pT-Adv vector (BD Biosciences, Clontech) and 256 clones carrying the recombinant plasmids were sequenced.

2.6. 3′ Rapid amplification of genomic ends (RAGE)

The 3′ ends of the *Lntx* and *Wntx* genes were determined using 3′ RAGE. Briefly, adapter-ligated libraries constructed after restriction of the genomic DNA with *Dra*I, *Eco*RV, *Pvu*II, *Sca*I, and *Stu*I were used as templates. Two separate sets of primers were used for amplifying the 3′ ends of the *Lntx* and *Wntx* genes. The *Lntx* gene was amplified using the adapter primer 1 (AP1: 5′-GTAATACGACT-CACTATAGGGC-3′), and the gene-specific primer X492 (5′-GACT-TAGGATACATAAGATGCT-3′) followed by nested PCR using adapter primer 2 (AP2: 5′-ACTATAGGGCACGCGTGGT-3′) and the gene-specific primer X263 (5′-CACCTGATGTTACCTCTACG-3′). The 3′ end of the *Wntx* gene(s) was amplified using AP1 and X476 (5′-CCTTGACATGTCTCAATTGC-3′) primers followed by nested PCR using AP2 and a gene-specific primer, X289 (5′-TGTGCTACTTGCCCTGGAAC-3′). Primary PCRs were carried out for 10 cycles each of 94°C for 30 s and 55°C for 3 min (touchdown) followed by 30 cycles each of 94°C for 30 s and 50°C for 3 min, with a final extension at 65°C for 10 min. The nested PCRs were carried out for 10 cycles each of 94°C for 30 s and 65°C for 3 min followed by 30 cycles each of 94°C for 30 s and 55°C for 3 min with the final extension at 68°C for 10 min. The PCR products were analyzed on agarose gels. The appropriate fragments were purified, subcloned into pT-Adv vector and sequenced.

2.7. Sequence analysis and homology modelling

Nucleotide sequence homology searches were performed using the BLAST program (www.ncbi.nih.nlm.gov). DNA sequence alignments were carried out using MegAlign (DNASTAR, USA). The number of nucleotide substitutions per site (K_N) in the 3′ non-coding regions and introns, the number of nucleotide substitutions per synonymous site (K_S) and non-synonymous site (K_A) in the protein coding region were computed for all pairs of *N. sputatrix Wntx* genes according to the method of Nei and Gojobori [28]. Phylogenetic trees were constructed using MEGA2.1 software [29] following the neighbor-joining (NJ) method using the Poisson-corrected amino acid distance. Homology modelling was carried out by the automatic homology modelling (first approach mode) available at http://www.expasy.ch/swissmod). The Swiss PDB viewer was used for three-dimensional structural comparison and manipulation [30–32].

3. Results

3.1. cDNA cloning and protein analysis of N. sputatrix LNTX

The product of RT-PCR using oligo dT and the 5′ untranslated region (UTR) primer for NTX was ~380 bp. Sequence analysis of the positive clones showed that the open reading frame of the cDNA encoded a long-chain α-NTX (Lntx) designated Lntx-7 (Fig. 1). The deduced amino acid sequence of this NTX showed that the first 63 nucleotides encoded 21 amino acid residues of a highly hydrophobic signal peptide. The remaining nucleotides encoded a mature toxin of 69 amino acid residues. The molecular weight of Lntx-7 was calculated to be 7493.5 Da. It contained 10 cysteine residues. Only

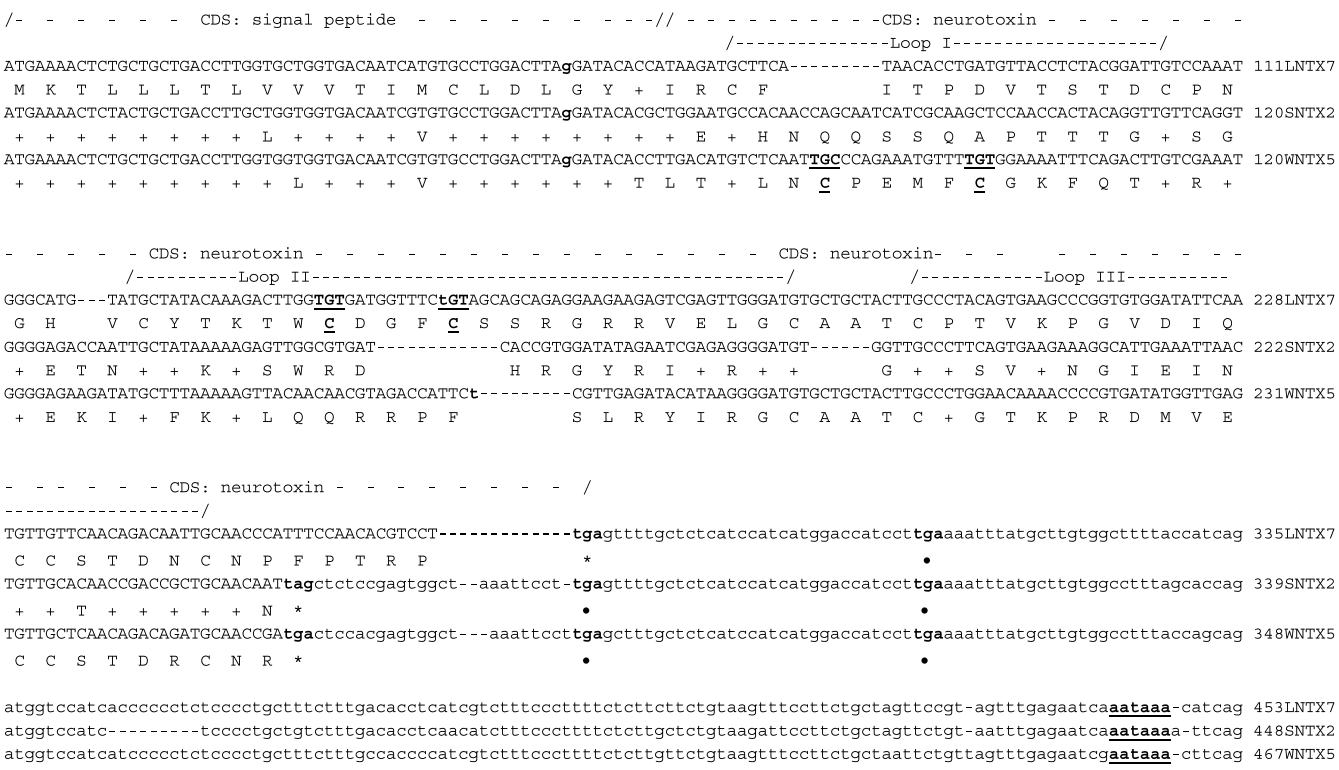


Fig. 1. Alignment of nucleotide and deduced amino acid sequences of α-NTXs identified in *N. sputatrix*. The sequence of Lntx-7 is from this study. NTX2 (SNTX) is from Afifiyan et al. [22] and Wntx is from Poh et al. [10]. The nucleotides in the coding and non-coding regions are indicated in upper and lower case letter respectively. The corresponding amino acid residues are indicated as single code upper case letters directly below the second base of the respective codon. Using Lntx-7 as the consensus sequence, all identical amino acid residues are indicated by '+' and insertions for maximum alignment are indicated by '–'. The numbering on the right refers to the number of nucleotides excluding the insertions. The position of exon/intron splice sites in the corresponding gene is shown in lower case bold letters (nucleotides) while the Cys residues (and their respective codons) responsible for the fifth cyclic loop in both Lntx and Wntx are given in bold letters and underlined. The stop codons are denoted by asterisks while the potential stop codons are denoted by bullets.

one transcript of mRNA was detected for Lntx-7, among the 256 clones that were sequenced from five separate RT-PCRs. The other clones encoded SNTXs as well as WNTXs.

Lntx-7 protein was purified from the crude venom by a combination of sequential gel filtration, cation exchange and reverse phase chromatography (Fig. 2). Mass analysis of the peptide showed it to be 7493.5 Da and N-terminal amino acid sequencing of the first 15 amino acid residues, IRX-FITPDVTSTDXP, suggested that the purified toxin corresponded to Lntx-7. No other variant of Lntx was detected in the venom. This protein was designated nLNTX-7 (native long-chain NTX). Amino acid composition analysis indicated that nLNTX-7 conforms to the structure of a typical potent LNTX having high sequence homology to α-cobratoxin (α-Cbtx) from *Naja kaouthia*.

The LD_{50} of crude venom was found to be 0.9 μg/g while that of nLNTX-7 was 0.14 μg/g. nLNTX-7 was further subjected to competitive receptor binding assays using muscular-type AChR from *Torpedo marmarata* and neuronal-type α7 AChR expressed on the cell surface of HEK as chimeric α7-$5HT_3$ nAChR. LNTX-7 was found to exhibit high affinity to both types of nAChRs (Table 1; $K_d = 54$ pM and 65 nM for muscular- and neuronal-type α7 nAChRs respectively).

3.2. *Structure of* N. sputatrix Lntx *and* Wntx *genes*

Strategies used in cloning the NTX genes are outlined in Fig. 3A. All the primers used were designed based on the *N. sputatrix* LNTX and WNTX cDNAs (Lntx and Wntx respectively). The complete protein coding region was amplified using primer pairs X133-X491 (*Lntx*, 2.2 kb; Fig. 3B, lane 2) and X133-X477 (*Wntx*, 2.0 kb and 2.4 kb; Fig. 3B, lane 1) corresponding to the beginning of the signal peptide and the last few amino acid residues at the C-terminus for the *Lntx* and *Wntx* genes respectively. The 3′ UTR of the *Lntx* gene was elucidated from the 1.2-kb PCR fragment from a *Stu*I (Fig. 3B, lane 3) library. Likewise, the 3′ end of the *Wntx* gene was obtained from the 400-bp 3′ RAGE PCR fragment of the *Pvu*II library (Fig. 3B, lane 4). Sequence analysis of the clones identified the gene sequences of *Lntx* and *Wntx* as 2673 bp and 2698 bp respectively (GenBank accession numbers: AY081759, AY081760, AY081761, AY081762 and AY081763).

The intron/exon junctions of these genes were assigned based on their respective cDNA sequences. The genes encoding both LNTX and WNTX precursors are organized into three exons interrupted by two introns (Fig. 3C). Exon 1 of both genes encoded a large part of the signal peptide and exon 2 encoded the later part of the signal peptide (three amino acid residues) followed by the N-terminal half of mature toxins (amino acids 1–29 in LNTX and 1–34 in WNTXs). Exon 3 encoded the C-terminal end of the toxins and the 3′ UTR. Based on the sequence analysis it is evident that the genome of spitting cobra contains only one *Lntx* gene and four variants for the *Wntx* gene. Since these four variants of *Wntx* were isolated from the same snake and they have been reported to produce four different transcripts [10], they could

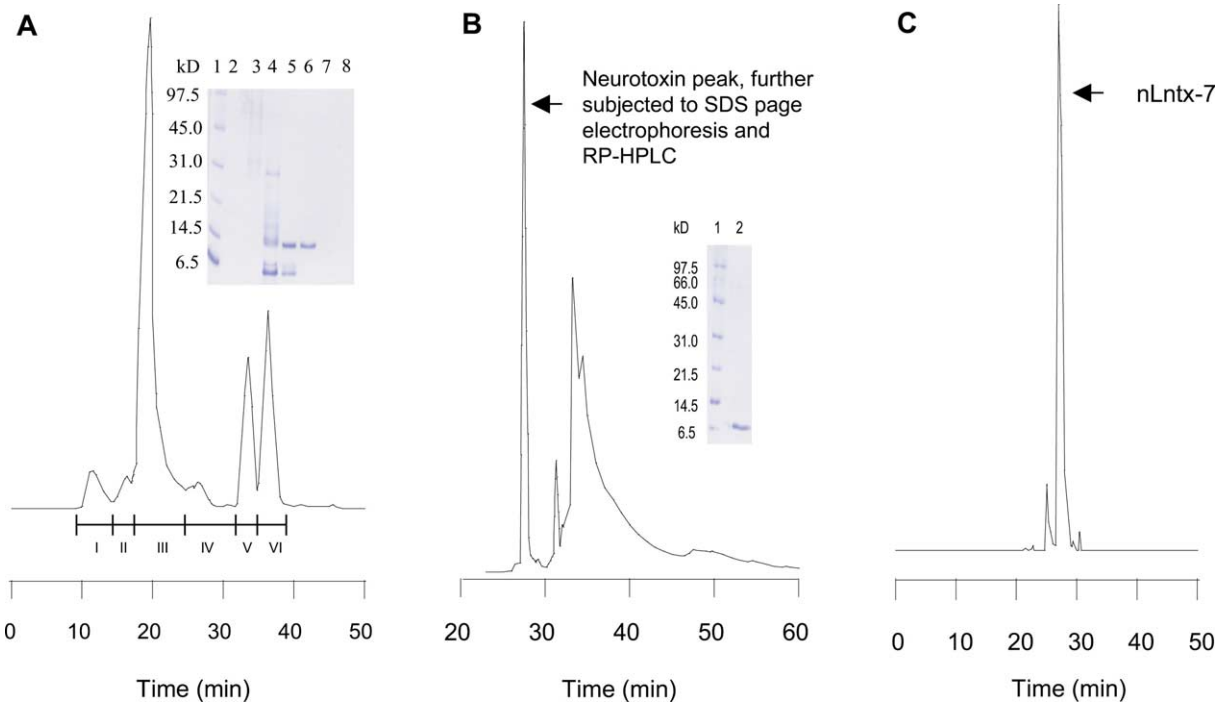


Fig. 2. Purification of nLNTX-7 from the crude venom. A: Gel filtration chromatography of *N. sputatrix* crude venom using Superdex G75 HR 10/30 (Pharmacia). Inset: SDS–PAGE analysis of the fractions. Lane 1: low molecular weight protein markers; lane 2, peak I fraction; lane 3, peak II fraction; lanes 4 and 5, fractions of peak III; lane 6, peak IV fraction; lane 7, peak V fraction; lane 8, peak VI fraction. B: FPLC ion exchange chromatography of pooled fractions from peak III on a Mono-S (5/5 HR) column. Equilibrating buffers: buffer A, 50 mM ammonium acetate, pH 4.5; buffer B, buffer A with 1 M NaCl. Elution of bound peptides was achieved with a two-step gradient that increased the concentration of B by 25% in 10 min followed by increase to 50% in 25 min at a flow rate of 0.5 ml/min. Eluting peaks were collected manually. Inset: SDS–PAGE analysis of the fractions from Mono S HR 5/5. Lane 1: low molecular weight protein markers; lane 2 shows peak 1, indicated by the arrow in the chromatogram. C: Reverse phase HPLC of peak 1 from Mono-S FPLC. An analytical Vydac 5 μm C18 column (250×10 mm) was used with a gradient of 25–100% of solution B in 10 min (solutions A and B are 0.1% TFA in water and 0.1% TFA in 90% acetonitrile respectively). The flow rate was maintained at 1 ml/min and the detection was monitored at 280 nm.

possibly represent four alleles of the same gene. The first intron splice site is within the leader sequence region between nucleotides 74/75 and 73/74 for *Lntx* and *Wntx* respectively. The second intron divides the coding region as a phase 1 intron in both genes. The splice junction (intron 2/exon 3) occurs between the first and second bases of the codons T/GT (Cys^{30}) for *Lntx* and either T/CG or G/GA (Ser^{35} or Gly^{35}) for *Wntx*.

3.3. Phylogenetic analysis and homology modelling

Phylogenetic analysis (Fig. 4A) showed that the NTX genes evolved from a common ancestor and later branched off as genes encoding the five-disulfide bridge three-finger toxins and the gene encoding the four-disulfide bridge toxins. The phylogenetic tree constructed from the protein sequences showed clusters based on the molecular properties of the toxins selected (Fig. 4B) with cardiotoxin and cardiotoxin-like proteins clustering away from the NTX group of proteins. Homology modelling of both LNTX and WNTX showed the signature scaffold for three-finger toxins but with discrepancies at the loop 1 and loop 2 regions of the toxins (Fig. 4C). LNTX showed the fifth cyclic loop in the major loop 2 domain while the cyclic loop in WNTX was seen in the loop 1 domain. These observations are important for further structural studies on these proteins.

4. Discussion

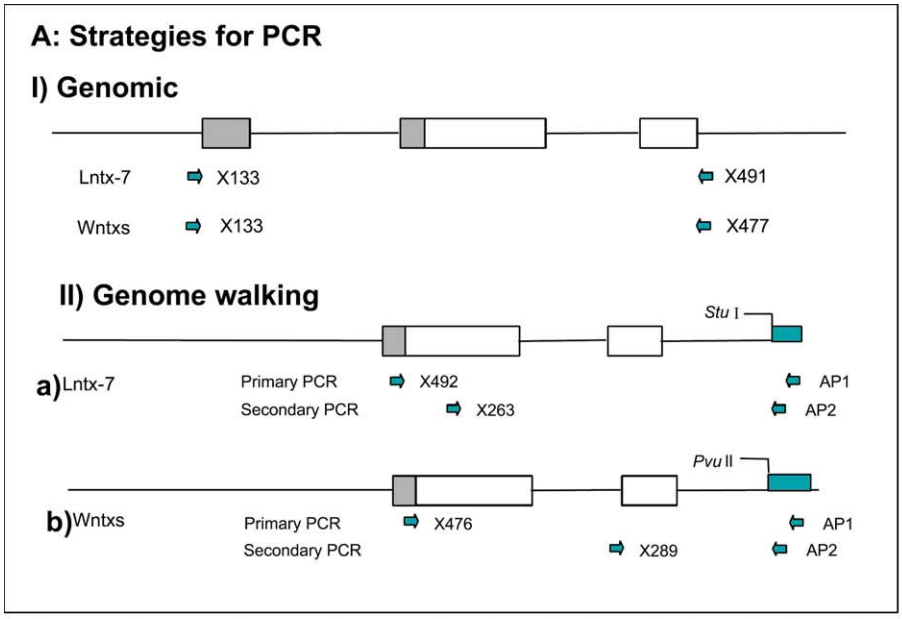
4.1. *LNTX in* N. sputatrix

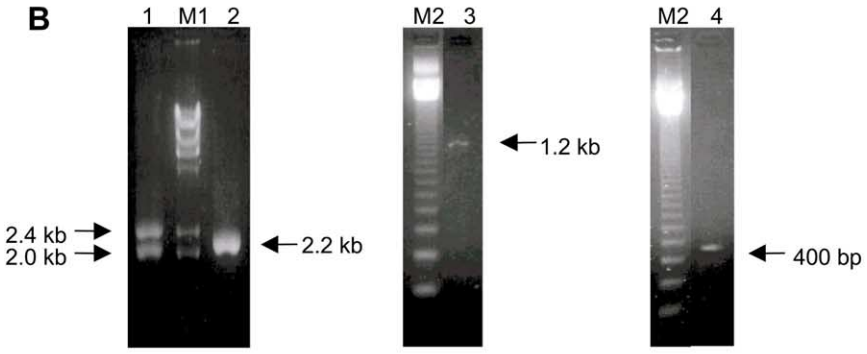
Only one type of LNTX was isolated, in each stage, as a mRNA transcript in the venom gland, a single gene in the genome and as a protein in the venom of *N. sputatrix*. Amino acid sequence comparison showed that LNTX-7 shares 88% homology with α-Cbtx, a long-chain α-NTX isolated from *N. kaouthia*. Functionally important residues involved in the binding of α-Cbtx to α7 nAChR and the 33 conserved residues known to be present in LNTXs [9] are also found in LNTX-7 (Fig. 5). nAChR binding affinity of LNTX-7 is also comparable to α-Cbtx (Table 1).

The neuronal α7-type nAChR binding residue, in α-Cbtx,

Table 1
Receptor binding activity of three-finger toxin (α-NTX) from *N. sputatrix*

	LNTX-7	Synthetic WNTX-5 [10]	SNTX-2
Muscle nAChR (K_d)	0.054 nM	180 nM	0.040 nM
Neuronal α7 nAChR (K_d)	65 nM	90 μM	0.22 nM (synaptosomes)
LD_{50}	0.14 μg/g	non-lethal	0.1 μg/g





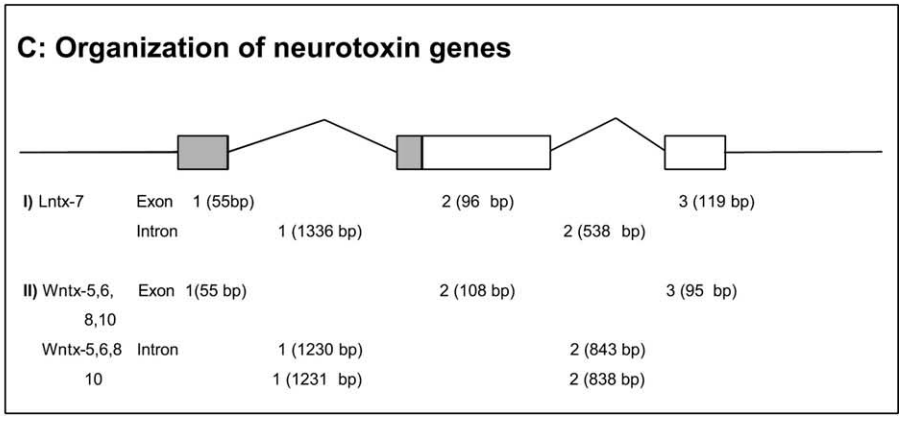


Fig. 3. Strategies for PCR and the gene organization of *Lntx* and *Wntx* of *N. sputatrix*. A: Strategies for PCR. Primers used and their relative positions on the genes are shown. AI: Genomic PCR of coding region. AII: 3′ RAGE. B: Electrophoresis of the products of genomic and genome walking PCR. Lane 1, genomic PCR of the *Wntx* gene between primers X133 and X477; M1: λ *Hin*dIII as DNA marker; lane 2, genomic PCR of the *Lntx* gene between primers X133 and X491; M2: 123-bp DNA ladder; 3′ end of *Lntx* amplified from *Stu*I (lane 3) and 3′ end of *Wntx* amplified from *Pvu*II library (lane 4). C: The structure and organization of the *ntx* genes in *N. sputatrix*. The exons are shown as boxes and introns as thin lines. Closed boxes represent the signal peptide coding region, while open boxes represent the mature protein coding region. The sizes of the exons and introns are shown.

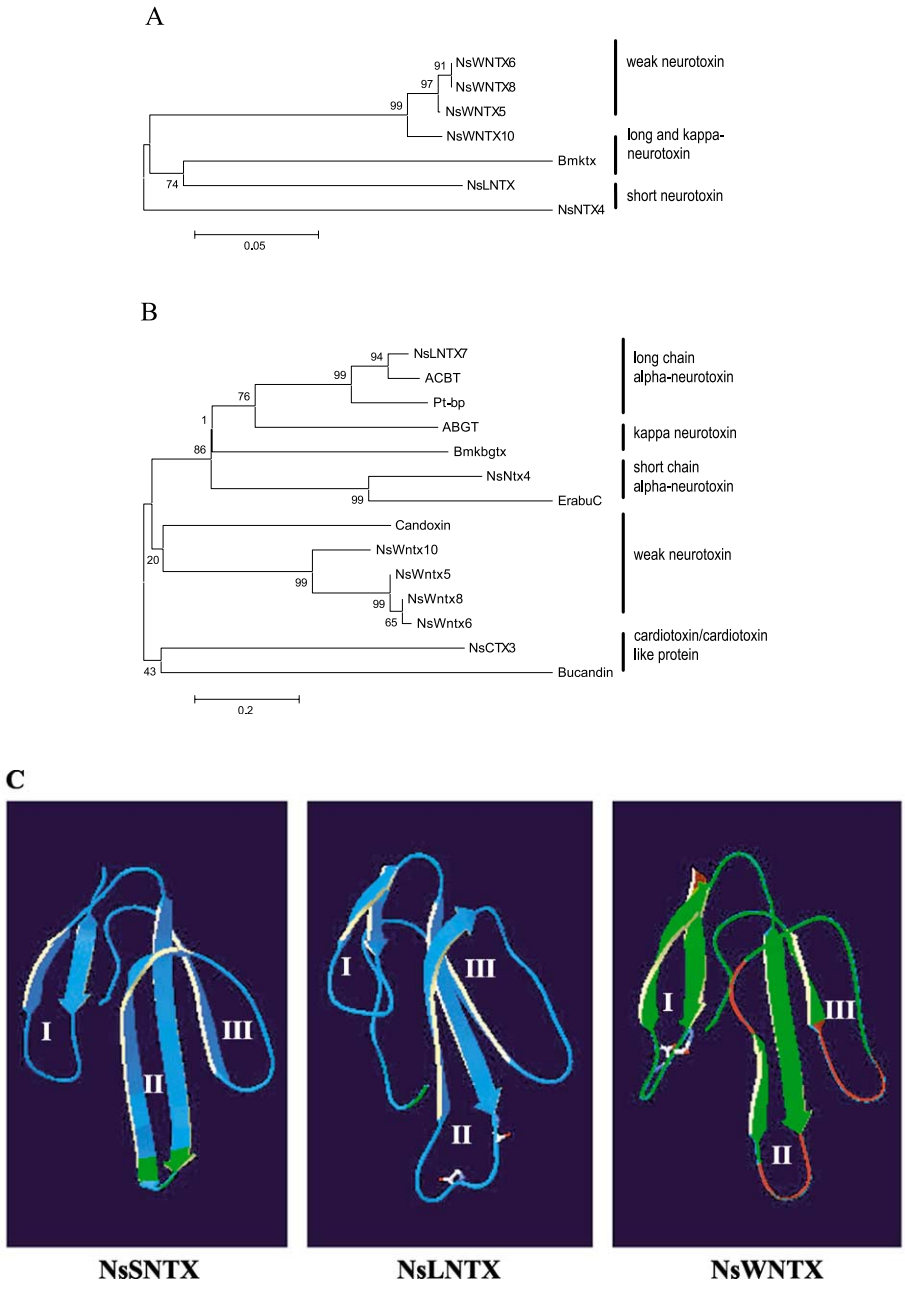


Fig. 4. Phylogenetic analysis and homology modelling. A: Phylogenetic tree of the nucleotide sequences of genes encoding three-finger toxins of *N. sputatrix*. Sequences of NsWNTX-5, 6, 8, 10 and NsLNTX-7, determined in this study, NsNTX4, from Afifiyan et al. [22] and Bmktx κ-Bgtx (Y11768). B: Phylogenetic tree of the amino acid sequences of NTXs, NsLNTX 7, NsWNTX 5, 6, 8 and 10, from this study, NSNTX4, from Afifiyan et al. [22], ACBT (P01391), Pt-bp (P13495), ABGT (P0137), ErabuC (P01435), candoxin (P81783), bucandin (P81782), NsCTX (AAC61314). The tree was constructed using the NJ method, on the basis of the Poisson-corrected amino acid distance. The reliability of clusters was estimated using the interior branch test. C: Homology modelling of three-finger toxins in *N. sputatrix*. The β-strands are indicated as arrows. The cysteine residues forming the cyclic loop are indicated in loop I and loop II for NsWNTX and NsLNTX respectively.

Ala^{31}, is replaced by Gly^{31} and Lys^{38} by Arg^{38} in LNTX-7. The Lys^{14} residue in loop I of α-Cbtx is replaced by Thr^{14} while Thr^{54} is replaced by Pro^{54} in LNTX-7. Since these residues are involved in structural maintenance, these two substitutions could cause minute deviations to the structural conformation of LNTX-7 in comparison to α-Cbtx. The binding affinity of LNTX-7 to α7 nAChR (K_d = 65 nM) is much lower than the affinity of α-Cbtx (K_d = 9 nM [7]) to the same receptor.

Kyte–Doolittle hydropathy profiles of *N. sputatrix* NTXs (LNTX-7, WNTX-5, 6, 8 and 9) show prominent positive hydrophilicity along the sequences [10]. Subtle differences were observed in loop II of LNTX-7 in comparison to α-Cbtx. The net charge of LNTX-7 (+0.8) is lower than that of α-Cbtx (+3.77). Together, these findings may explain the variation in biological potency between the two toxins. LNTX-7 is, however, more potent than the non-toxic WNTX (Table 1), another five-disulfide bridge three-finger toxin, isolated from the same snake.

Apart from the variation in chain length compared to that

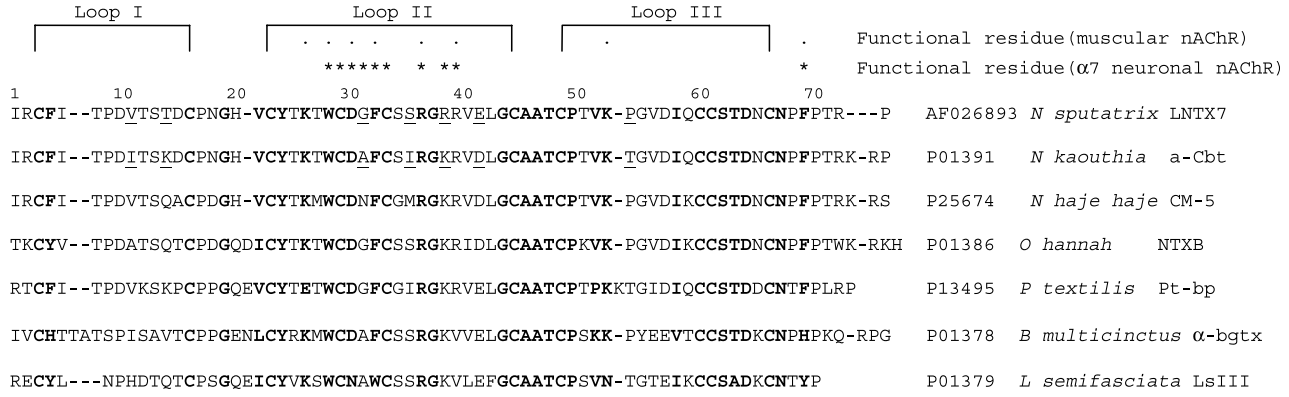


Fig. 5. Amino acid sequence alignment of snake LNTXs. All sequences were denoted according to the accession numbers as provided by the GenBank, Swiss-Prot and TrEMBLE databases. Conserved residues in LNTXs are in bold letters. The variant amino acid residues between LNTX-7 and α-Cbtx are underlined.

of SNTX, a fifth pair of cysteine residues can be seen in both LNTX and WNTX (Fig. 1). These two cysteines (Cys^6 and Cys^{11}) form the fifth cyclic loop within the signature loop I region of WNTX [10]. Loop I of Lntx-7 is four amino acid residues shorter than loop I of SNTX and WNTX. On the other hand, loop II of Lntx-7 contains a fifth cyclic loop formed by Cys^{29} and Cys^{33} (Fig. 1). This cyclic loop, which is characteristic of all LNTXs, is reported to be involved in binding of LNTX to α7 nAChR [7]. Examination of the 3′ non-coding region of the NTX cDNAs (Fig. 1) revealed that there could be more than one possible stop codon. Comparison of the deduced amino acid sequences of these cDNAs shows five additional amino acid residues for LNTX at the C-terminus (Fig. 1). The corresponding stop codons, TAG in Ntx-2 and TGA in Wntx-5, have been mutated to TTT in Lntx-7 cDNAs, resulting in continuation of the message up to the next stop codon (TGA at nt 270–272) as observed for the Ls III isolated from the sea snake *Laticauda semifasciata* [33]. Similar stop codons have been reported for the gene encoding LNTXs found in *Pseudonaja textilis* [34]. Therefore, it is possible that a shift in the stop codon to about 10 nucleotide residues upstream along the 3′ UTR in Lntx-7 could have resulted in a shorter Wntx/Ntx2 transcript in *N. sputatrix*. The Lntx-7 transcript could have been a precursor for the shorter α-NTXs (Fig. 1).

4.2. Genes encoding the five-disulfide bridge three-finger toxins

Cloning and sequence analysis of *Lntx* and *Wntx* genes (GenBank accession numbers: AY081759 to AY081763 for nucleotide sequences) have provided evidence for the presence of one *Lntx* and possibly four alleles of *Wntx* genes in the genome of *N. sputatrix*, which corresponded to the messages and toxins found in the venom gland ([10] and this study). DNA sequence analysis showed that these NTX genes shared identical features with other known 'three-finger' toxin genes, having an overall gene organization of three exons and two introns. Exon 1 is highly conserved among this family of toxins and appeared to be protected by the longer intron 1 in a manner similar to that proposed for PLA_2 genes [18]. The insertion of a snoRNA fragment in this intron has resulted in a longer intron 1 compared to intron 1 in the erabutoxin gene from sea snake [35]. SnoRNA insertion can also be found in intron 1 of other three-finger toxins (NTXs and cardiotoxins) in cobra [20,23,36]. The intron–exon splice junctions show the characteristic GT/AG sites of phase 1 introns [37]. The splice sites of both Lntx and Wntx have been found to be similar to the splice site (phase 1 introns) in genes encoding other LNTXs such as α-Bgtx [38], Pt-bp [34] and SNTXs such as Ec [35], cobratoxin [39], Sntxs from *N. sputatrix* [23] and *P. textilis* [40] and cardiotoxin from *N. sputatrix* [20].

Table 2
Subunit (receptor) specificity of NTXs

Toxin	nAChR subunit
SNTX	High affinity to muscular (α1) and lower affinity to neuronal nAChR (α7)
LNTX	Almost equal affinity to both muscular (α1) and neuronal nAChRs (α7)
κ-NTX	High affinity to neuronal α3 and α4 nAChRs and no affinity to muscular (α1) or neuronal nAChRs (α7)
WNTX	Very low affinity to both muscular (α1) and neuronal (α7) nAChRs

The values for K_N, K_S and K_A indicate that nucleotide substitutions occur more frequently in the mature protein coding regions (exons 2 and 3) than non-coding regions, and that nearly all these nucleotide substitutions result in amino acid changes ($K_A/K_S > 1$ [28]), thus contributing to their diverse functions. All codons encoding the structurally important residues of the toxins have been conserved while the codons for residues located at the surface of the toxin molecule have been found to be highly susceptible to mutation [10]. Interestingly, the codon encoding the extra Cys residues responsible for the formation of the fifth cyclic loop in both toxins can be found in exon 2. In Wntx, these codons can be found around the mid region of exon 2, whereas in Lntx, these codons remain at the end of exon 2, and at the exon 2/intron 2 splice site. This displacement of cysteine residues could have occurred by a deletion mutation at the end of exon 2 in Lntx, followed by an insertion (nt 143–156) and point mutations at nt 138–140 (CGT to TGT; Fig. 1) in the *Wntx* gene.

4.3. Why multiple three-finger toxins in snake venom?

The specificity and selectivity of a particular three-finger toxin towards various nAChRs or other targets depend on the nature and distribution of functional groups along the three-finger scaffold, responsible for the subtle structural deviation and the conformational mobility of these loops (Table

2) [41,42]. The toxin binding affinity of LNTX to neuronal nAChRs is associated with its subtle differences in the nature of the amino acid residues compared to SNTXs as well as the fifth disulfide cyclic loop in the classical, loop II region. Insertion of the extra cyclic loop in loop II of SNTX increases its affinity to neuronal nAChRs [43], while removing this cyclic loop from LNTX increases its affinity to muscular nAChRs [7]. Grant et al. [44] have shown that the removal of the fifth cyclic loop in κ-Bgtx resulted in a 50-fold decrease in neuronal α3β2 AChR binding affinity, thus showing the specific interaction of the toxin with a specific/unique neuronal receptor. Similarly, the absence of the nucleotide sequence encoding the pancreatic loop in the PLA_2 gene resulted in lethal PLA_2 proteins in the venom [18,45]. The biological target of WNTX appears to be AChRs, but the binding affinity of WNTX to nAChR is lower (μM range, Table 1) than that of LNTX and SNTX. The advantage of the low affinity binding of WNTX could be that it is non-toxic to the animal while causing a slow dissociating effect from the receptor [11]. Besides, the toxin binds to homologous regions of neuronal and/or muscular receptors and it does not involve identical determinants [46]. Thus, insertion and removal of the cyclic loop from the proteins (i.e. removal/insertion of codons encoding the residues, from the genes within exon 2) is nature's way of selecting the target site for the 'new' toxins that are needed by the animal, which could therefore explain the presence of multiple LNTXs and SNTXs in venoms.

Naturally, snake NTXs have evolved to paralyze their prey. Paralysis is caused by inactivating muscle AChR, therefore both LNTXs and SNTXs exhibit high affinity to muscle AChR. α-Bgtx, a highly lethal LNTX, shows high affinity to the α1 subunit of frog and chick muscle nAChRs where both frogs and chickens are the natural prey of *Bungarus multicinctus* [46]. The cobra, *N. sputatrix*, feeds mainly on rodents and small mammals. These are fast-moving animals. Thus, the highly lethal and toxic α-NTXs are of importance to rapidly immobilize the prey. On the other hand, predators of snakes, snakes themselves as well as the mongoose and others that are resistant to snake venom in general and α-Bgtx in particular [47,48], show natural mutations in their respective nAChRs to protect them from envenomation. Takacs et al. [49] documented that evolutionary success requires the animal venoms to be targeted against phylogenetically/evolutionarily conserved molecular motifs of fundamental physiological processes. Phylogenetic analysis (Fig. 4A,B) shows that the neurotoxic genes have evolved from a common ancestor. It is noteworthy that all the NTX genes contain an equally sized intron 1 (∼1204–1337 bp). However, intron 2 for the *Wntx* gene is approximately 300 bp longer than intron 2 in other NTXs. Betts et al. [50] reported that in rapid (accelerated) evolution the position of introns and the protein 3D structures are more conserved than that of the amino acid sequences. The authors further emphasized that construction of a chimeric domain between members of the same family could be the function of conservation of intron positions. Therefore, the discrepancies in intron size could reflect functional variations among the toxins as noted in the phylogram constructed using the amino acid sequences (Fig. 4B). Fig. 4B shows distinct clustering based on the molecular properties of the proteins. The highly lethal and high affinity (AChR) ligand-containing NTXs (short, long and κ) form a separate cluster from the non-lethal, low affinity ligand-containing WNTXs. It is interesting to note that CTX and bucandin form an isolated cluster from that of NTXs. The WNTXs are non-lethal five-disulfide bridge toxins with very low affinity towards both neuronal and muscular AChRs. They have been found to preferentially bind to the muscular-type nAChRs from fish and neurons of the freshwater mollusk *Lymnaea stagnalis* and they produce a long lasting binding effect on the nAChR [11,51]. This group of toxins forms a separate cluster (Fig. 4B) with candoxin. Candoxin has been reported to be a poor antagonist of neuronal α7 AChR [52].

Hughes [53] has also reported that functionally novel proteins are the product of duplication of a generalized multifunctional ancestral gene followed by adaptive evolution of the duplicated gene. Thus, it is possible that the precursor for NTX (Fig. 4A,B) could have undergone gene duplication, followed by accelerated adaptive evolution on the surface of the new molecule resulting in (i) ability to bind to both neuronal and muscular nAChRs with high affinity and high lethality (LNTXs), (ii) high affinity binding for muscular-type nAChR while maintaining high lethality (SNTX) and (iii) gradual weakening to receptor binding property, possibly gaining other new functions and loss of lethality [10].

Acknowledgements: This work was supported by Research Grant R-183-000-068-213 from the National University of Singapore.

References

[1] Chang, C.C. (1979) in: Snake Venoms. Handbook of Experimental Pharmacology, Vol. 52 (Lee, C.Y., Ed.), pp. 309–376, Springer-Verlag, Berlin.
[2] Brown, L.R. and Wütrich, K. (1992) J. Mol. Biol. 227, 1118–1135.
[3] Ménez, A., Bontems, F., Roumestand, C., Gilquin, B. and Toma, F. (1992) in: The Advancement of Drug Discovery (Harvey, A.L., Ed.), pp. 83–103, Royal Society, Edinburgh.
[4] Servent, D. and Ménez, A. (2001) in: Neurotoxicological Handbook Vol. I (Massaro, E.J., Ed.), pp. 385–425, Humana Press, Totowa, NJ.
[5] Endo, T. and Tamiya, N. (1991) in: Snake Toxin (Harvey, A.L., Ed.), pp. 165–222, Pergamon Press, New York.
[6] Schmidt, J.J. and Weinstein, S.A. (1995) Toxicon 33, 1043–1049.
[7] Servent, D., Winckler-Dietrich, V., Hu, H.Y., Kessler, P., Drevet, P., Bertrand, D. and Ménez, A. (1997) J. Biol. Chem. 272, 24279–24286.
[8] Grant, G.A. (1998) J. Toxicol. 17, 239–260.
[9] Antil-Delbeke, S., Gaillard, C., Tamiya, T., Corringer, P.J., Changeux, J.P., Servent, D. and Menez, A. (2000) J. Biol. Chem. 275, 29594–29601.
[10] Poh, S.L., Mourier, G., Thai, R., Armugam, A., Molgo, J., Servent, D., Jeyaseelan, K. and Menez, A. (2002) Eur. J. Biochem. 269, 4247–4256.
[11] Utkin, Y.N., Kukhtina, V.V., Kryukova, E.V., Chiodini, F., Bertrand, D. and Methfessel, C. (2001) J. Biol. Chem. 276, 15810–15815.
[12] Hatanaka, H., Oka, M., Kohda, D., Tate, S., Suda, A., Tamiya, N. and Inagaki, F. (1994) J. Mol. Biol. 240, 155–166.
[13] Albrand, J.P., Blackledge, M.J., Pascaud, F., Hollecker, M. and Marion, D. (1995) Biochemistry 34, 5923–5937.
[14] Tsetlin, V.I. (1999) Eur. J. Biochem. 264, 281–286.
[15] Servent, D., Mourier, G., Antil, S. and Ménez, A. (1998) Toxicol. Lett. 28, 199–203.
[16] Tan, N.H. (1991) Trop. Biomed. 8, 91–103.
[17] Armugam, A., Earnest, L., Chung, M.C., Gopalakrishnakone, P., Tan, C.H., Tan, N.H. and Jeyaseelan, K. (1997) Toxicon 35, 27–37.
[18] Jeyaseelan, K., Armugam, A., Donghui, M. and Tan, N.H. (2000) Mol. Biol. Evol. 17, 1010–1021.
[19] Jeyaseelan, K., Armugam, A., Lachumanan, R., Tan, C.H. and Tan, N.H. (1998) Biochim. Biophys. Acta 1380, 209–222.

[20] Lachumanan, R., Armugam, A., Tan, C.H. and Jeyaseelan, K. (1998) FEBS Lett. 433, 119–124.
[21] Ma, D., Armugam, A. and Jeyaseelan, K. (2001) Eur. J. Biochem. 268, 1844–1850.
[22] Afifiyan, F., Armugam, A., Gopalakrishnakone, P., Tan, N.H., Tan, C.H. and Jeyaseelan, K. (1998) Toxicon 36, 1871–1885.
[23] Afifiyan, F., Armugam, A., Tan, C.H., Gopalakrishnakone, P. and Jeyaseelan, K. (1999) Genome Res. 9, 259–266.
[24] Yanish-Perron, C., Vieire, J. and Messing, J. (1985) Gene 33, 103–199.
[25] Saitoh, T., Oswald, R., Wennogle, L.P. and Changeux, J.P. (1980) FEBS Lett. 116, 30–36.
[26] Cheng, Y.C. and Prusoff, W.H. (1973) Biochem. Pharmacol. 22, 3099–3108.
[27] Weber, M. and Changeux, J.P. (1974) Mol. Pharmacol. 10, 15–34.
[28] Nei, M. and Gojobori, T. (1986) Mol. Biol. Evol. 3, 418–426.
[29] Kumar, S., Tamura, K., Jakobsen, I.B. and Nei, M. (2001) Bioinformatics 17, 1244–1245.
[30] Guex, N. and Peitsch, M.C. (1997) Electrophoresis 18, 2714–2723.
[31] Peitsch, M.C. (1995) Bio/Technology 13, 658–660.
[32] Peitsch, M.C. (1996) Biochem. Soc. Trans. 24, 274–279.
[33] Fuse, N., Tsuchiya, T., Nonomura, Y., Ménez, A. and Tamiya, T. (1990) Eur. J. Biochem. 193, 629–633.
[34] Gong, N., Armugam, A., Mirtschin, P. and Jeyaseelan, K. (2001) Biochem. J. 358, 647–656.
[35] Tamiya, T., Ohno, S., Nishimura, E., Fujimi, T.J. and Tsuchiya, T. (1999) Toxicon 37, 181–185.
[36] 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. (Tokyo) 122, 1252–1259.
[37] Breathnach, R. and Chambon, P. (1981) Annu. Rev. Biochem. 50, 349–383.
[38] Chang, L., Lin, S., Huang, H. and Hsiao, M. (1999) Nucleic Acids Res. 27, 3970–3975.
[39] Chang, L.S., Lin, J., Chou, Y.C. and Hong, E. (1997) Biochem. Biophys. Res. Commun. 239, 756–762.
[40] Gong, N., Armugam, A. and Jeyaseelan, K. (2000) FEBS Lett. 473, 303–310.
[41] Ricciardi, A., le Du, M.H., Khayati, M., Dajas, F., Boulain, J.C., Menez, A. and Ducancel, F. (2000) J. Biol. Chem. 275, 18302–18310.
[42] Kumar, T.K., Jayaraman, G., Lee, C.S., Arunkumar, A.I., Sivaraman, T., Samuel, D. and Yu, C. (1997) J. Biomol. Struct. Dyn. 15, 431–463.
[43] Servent, D., Lamthanh, H., Antil, S., Bertrand, D., Corringer, P.J., Changeux, J.P. and Ménez, A. (1998) J. Physiol. (Paris) 92, 107–111.
[44] Grant, G.A., Luetje, C.W., Summers, R. and Xu, X.L. (1998) Biochemistry 37, 12166–12171.
[45] Davidson, F.F (1990) J. Mol. Evol. 31, 228–238.
[46] Samson, A.O., Scherf, T., Eisenstein, M., Chill, J.H. and Anglister, J. (2002) Neuron 35, 319–332.
[47] Duda Jr., T.F. and Palumbi, S.R. (2000) Mol. Biol. Evol. 17, 1286–1293.
[48] Hughes, A.L. (1999) Adaptive Evolution of Genes and Genomes, Oxford University Press, New York.
[49] Takacs, Z., Wilhelmsen, K.C. and Sorota, S. (2001) Mol. Biol. Evol. 18, 1800–1809.
[50] Betts, M.J., Guigo, R., Agarwal, P. and Russell, R.B. (2001) EMBO J. 20, 5354–5360.
[51] Vulfius, C.A., Krasts, I.V., Utkin, Y.N. and Tsetlin, V.I. (2001) Neurosci. Lett. 309, 189–192.
[52] Nirthanan, S., Charpantier, E., Gopalakrishnakone, P., Gwee, M.C., Khoo, H.E., Cheah, L.S., Bertrand, D. and Kini, R.M. (2002) J. Biol. Chem. 277, 17811–17820.
[53] Hughes, A.L. (2002) Trends Genet. 18, 433–434.