Molecular characterization of a possible progenitor sodium channel toxin from the Old World scorpion *Mesobuthus martensii*

S. Zhu*, B. Gao

Group of Animal Innate Immunity, State Key Laboratory of Integrated Management of Pest Insects & Rodents, Institute of Zoology, Chinese Academy of Sciences, No. 25 Beisihuan-Xi Road, Beijing 100080, PR China

Abstract Toxins affecting sodium channels widely exist in the venoms of scorpions throughout the world. These molecules comprise an evolutionarily related peptide family with three shared features including conserved three-dimensional structure and gene organization, and similar function. Based on different pharmacological profiles and binding properties, scorpion sodium channel toxins are divided into α- and β-groups. However, their evolutionary relationship is not yet established. Here, we report a gene isolated from the venom gland of scorpion *Mesobuthus martensii* which encodes a novel sodium channel toxin-like peptide of 64 amino acids, named Mesotoxin. The Mesotoxin gene is organized into three exons and two introns with the second intron location conserved across the family. This peptide is unusual in that it has only three disulfides and a long cysteine-free tail with loop size and structural characteristics close to β-toxins. Evolutionary analysis favors its basal position in the origin of scorpion sodium channel toxins as a progenitor. The discovery of Mesotoxin will assist investigations into the key issue regarding the origin and evolution of scorpion toxins.

1. Introduction

Scorpions belong to venomous arachnids and exist at an intermediate level in food chains [1]. As generalist predators, they feed on insects, spiders, and other small animals. In the mean time, these organisms themselves are prey to a variety of larger predators including both invertebrates and vertebrates. Evolutionary emergence of lethally toxic peptides (toxins) in scorpion venoms provides highly efficient means for capturing prey and protecting themselves [2–4]. This could compensate their physical limitation and benefit their survival in a competitive environment [5]. Remarkably, despite having different evolutionary histories and food sources, and occupy-

ing diverse ecological niches, scorpions and other predatory organisms such as spiders, cone snails and snakes have developed a large number of toxins commonly targeting sodium channels present in the excitable cells of preys and competitors [2,6–8]. This may be a consequence of natural selection if we consider crucial roles of sodium channels in controlling the electrical activity of nerve and muscle systems [9]. Not surprisingly, modifying the pharmacological activities of these channels is capable of causing rapid immobilization of their preys.

Mechanically, these sodium channel toxins affect both permeation and gating properties of the channels by targeting distinct receptor sites and inducing conformational changes of the channel protein [10]. Despite very different sequences and fold types, some toxins from phylogentically distant venomous animals convergently target the same receptor site. In this respect, scorpion α-toxins, sea anemone toxins and spider toxins provide good examples in which three unrelated structural types commonly target the site 3 of the channel [10]. Divergent evolution stemming from gene duplication and speciation represents another common strategy in toxin evolution in which structural cores conserves but function changes [11,12].

In the past few years, due to the combination of gene cloning techniques and traditional biochemical methods, the family of scorpion toxins affecting sodium channels (NaScTx) has been enlarged to more than 200 members [13]. A typical NaScTx is composed of about 64 residues cross-linked by four disulfide bridges. On the basis of different pharmacological profiles and binding properties, the NaScTxs are classified into two distinctive groups: α- and β-toxins [14]. The α-toxins prolong the inactivation of the channel whereas β-toxins affect activation of the channel. This disparate pharmacological property roughly reflects the separation of Old World and New World scorpion species: the α-toxins affecting the channel inactivation extensively distribute in Old World scorpions, whereas the β-toxins affecting activation are mainly present in New World scorpions.

Studies on the origin and evolution of NaScTxs have attracted extensive attention in the recent years. A subfamily of NaScTx, represented by Birtoxin, is especially interesting in this aspect. All the members in this subfamily contain only three disulfides with a slightly smaller size (about 58 residues) relative to the NaScTxs with four disulfides [15–18]. They share 40–60% sequence identity to β-toxins but display more diverse pharmacological activity. Some act as modulators affecting sodium channel activation with a characteristic β-toxin effect. Others serve as potassium channel blockers due to the development of a putative functional dyad in their α-helical

*Corresponding author. Fax: +86 10 62579091.
E-mail address: Zhusy@ioz.ac.cn (S. Zhu).

region. Previous phylogenetic analysis assigned the Birtoxin subfamily as the ancestor of all the NaScTxs [13,19]. Recently, a new Birtoxin-like peptide, named BmKBT, was characterized from the Old World scorpion *Mesobuthus martensii* [20]. Genomic data indicates that BmKBT is in fact encoded by a transcript variant of the lipolysis activating peptide α-subunit gene, both differing in one base deletion only (Zhu and Gao, unpublished data). The α-subunit contains seven cysteines forming three intramolecular disulfides and one intermolecular disulfide [21]. The latter links the α-subunit to the β-chain by the last cysteine. Such a deletion led to the production of BmKBT due to truncation and loss of the last cysteine. It thus appears that the Birtoxin subfamily is a mutated version of the peptide with seven cysteines. Another study on the basis of a combination of sequence, structural and functional data hypothesized a functionally unrelated but structurally conserved peptide – antifungal defensin as the ancestor of NaScTxs. Some evolutionary events in terms of the origin were elucidated which included the adding of five-residue turn close to the N-terminus and the extension of C-terminus, and reorganization of the fourth disulfide bridge [22].

Here, we report a full-length cDNA and its complete gene which encodes a unique peptide of 64 residues with only three disulfide bridges. Sequence and structural analysis places it to the β-toxin group. Several characteristics together with a basal position in the evolutionary tree suggest that Mesotoxin might be a putative progenitor of the NaScTx family.

2. Materials and methods

2.1. Preparation of total RNA and genomic DNA
Scorpions (*Mesobthus martensii*) were kindly provided by Dr. Qilian Qin (Institute of Zoology, Beijing). Total RNA and genomic DNA were prepared according to previously described methods [23].

2.2. Isolation of cDNA and genomic clones
According to the nucleotide sequence of BmTXLP3 (AF159977), we designed two reverse gene-specific primers (VP3-1 and VP3-2) for 5′ RACE to obtain its 5′ cDNA end sequence. Briefly, total RNA was reverse-transcribed into the first-strand cDNAs using RT-PreMix kit (SBS Genetech, Beijing) and a universal oligo(dT)-containing adaptor primer (dT3AP). The purified first-strand cDNA mixture was tailed with terminal transferase and dCTP (Takara, Dalian). A PCR amplification of the tailed first-strand cDNAs was carried out using primers dG and VP3-1 and TaKaRa LA Taq, a DNA polymerase with 3′ → 5′ exonuclease proofreading activity. After 35 cycles, 1 μl of diluted PCR product was taken as template for another 35 cycles of amplification with primers dG and VP3-2. To determine the exon–intron organization of the Mesotoxin, the genomic DNA was amplified using primers VP3-F and VP3-2 under standard PCR conditions. PCR products were ligated into the pGEM-T Easy Vector following purification using PCR purification kit (Tiangen Biotech, Beijing). *Escherichia coli* DH5α was used for plasmid propagation.

2.3. DNA sequencing
Recombinant clones were analyzed by PCR using two vector primers (SP6 and T7) and gel electrophoresis. Positive clones were sequenced through the chain termination method using the primers SP6 and T7. The nucleotide sequence of the Mesotoxin gene has been deposited in the GenBank database (http://www.ncbi.nlm.nih.gov) under the accession number of DQ872676.

2.4. Primer sequences
All primers used in this study are synthesized by SBS Genetech, Beijing and Takara, Dalian, and listed in Table 1.

2.5. Searching TF binding sites
Potential transcription factor (TF) binding sites in the upstream of transcriptional start site 2 (TSS2) of Mesotoxin gene was searched using the AliBaba2 program against the TRANSFAC database (http://www.cs.uni-magdeburg.de/grabe/alibaba2).

2.6. RNA secondary structure prediction
Mfold, which predicts RNA secondary structure by free energy minimization (http://www.bioinfo.rpi.edu/applications/mfold/), was used to identify possible RNA hairpins in the non-coding regions of pre-transcripts (5′ and 3′ UTRs and introns).

2.7. Evolutionary analysis
Multiple sequence alignment of scorpion sodium channel toxins and related peptides was carried out using the CLUSTAL W program (http://www2.ebi.ac.uk/clustalw/) and further refined by hand with reference to the cysteine residue position. The aligned sequences were employed to make phylogenetic analysis using the neighbor-joining method implemented in MGEA 3.1 (http://www.megasoftware.net).

2.8. Homology modeling
Fold compatibility and template selection for comparative modeling were performed through GenTHREADER [24]. The experimental structure of scorpion neurotoxin CsEv2a from *Centruroides sculpturatus* (PDB entry 1JZA) [25] was selected as a template for modeling the Mesotoxin1 structure. Sequence alignment was undertaken using the CLUSTAL W program and further refined by hand to remove gaps within α-helix and β-strands. Once an accurate alignment was determined, three-dimensional (3D) models were generated with programs TITO and MODELLER (http://bioserv.cbs.cnrs.fr/). Models were evaluated by the Verify3D [26] and PROSA [27]. Structural superimposition and rmsd (root mean square deviation) calculation were performed using Swiss-PdbViewer program [28]. All structures were displayed using the program MOLMOL [29]. The 3D protein model of Mesotoxin1 has been submitted to the Protein Model database (http://www.caspur.it/PMDB/) under the id number of PM0074674.

3. Results

3.1. Identification of cDNAs encoding Mesotoxins
The present study was initiated by our analysis of the EST clone AF159977 from *M. martensii*. This clone, initially submitted to GenBank database by Shunyi Zhu and Wenxin Li, encodes a peptide of 46 residues (ENLGEDCENLCKQQ-

Table 1
The PCR primers

Name	Sequences	Usage
dT3AP	5′-CTGATCTAGAGGTACCGGATCCTTTTTTTTTTTTTTTTTTT-3′	Reverse-transcribed reaction
dG	5′-ATGAATTCGGGGGGGGGGGGGGGG-3′	5′RACE
VP3-1	5′-TTATTACATTATCGGGTCTACA-3′	5′RACE
VP3-2	5′-GTATCGGGTCTAGTTGCATA-3′	5′RACE, genomic amplification
VP3-F	5′-TTTTAATACAACTAGCTAGAACT-3′	Genomic amplification
3SSF	5′-ATGGATCCGATGACGATGACAAGGTAAAAGACCGTTTCTTGAT-3′	Constructing expression vector
3SSR	5′-ATGTCGACTTACATTATCGGGTCTACAGTAT-3′	Constructing expression vector

KATDGFCRQPHCFCTDMPDNYATRPDTVDPIM). The truncated product contains a cysteine arrangement pattern C......CX₃CX₇GX₁CX₄CX₁C which well matches the consensus motif of the CSαβ superfamily. Lack of the first cysteine (C) may be due to mRNA degradation at the 5′-end during RNA preparation. A BLAST search of the GenBank database using this sequence as a query with a default parameter identified two scorpion β-toxins (Tco 38.32-2 and Tf4) [30,31] with some amino acid similarity within the cysteine frame. However, this truncated peptide is unusual in that relative to the NaScTxs, it carries a long cysteine-free tail comprising 18 residues. Given its possible evolutionary importance, we decided to determine its full-length sequence. Sequencing two positive clones (VP3-4 and VP3-7) derived from 5′ RACE-PCR allowed us to assemble two full-length cDNAs with nearly iden-

tical sequence in the open reading frame (ORF) and 3′UTR but variable 5′ UTRs in length. Both 5′ and 3′UTRs are rich in A + T content, respectively up to 80% and 88%, whereas the ORF only contains 65% A + T. The ORFs of these two clones code for a precursor of 86 residues composed of an amino-terminal signal peptide of 22 residues and a carboxyl-terminal mature peptide of 64 residues which contains the truncated sequence of 46 residues. Only one polymorphic site was found between these two clones, occurring at position 712, which led to an amino acid change from Asn to Asp due to base substitution (A–G) (Fig. 1).

3.2. Two transcriptional start sites in the Mesotoxin gene

The 5′UTR of the large transcript is the longest one characterized to date which is composed of 151 nucleotides (nt), 89 nt
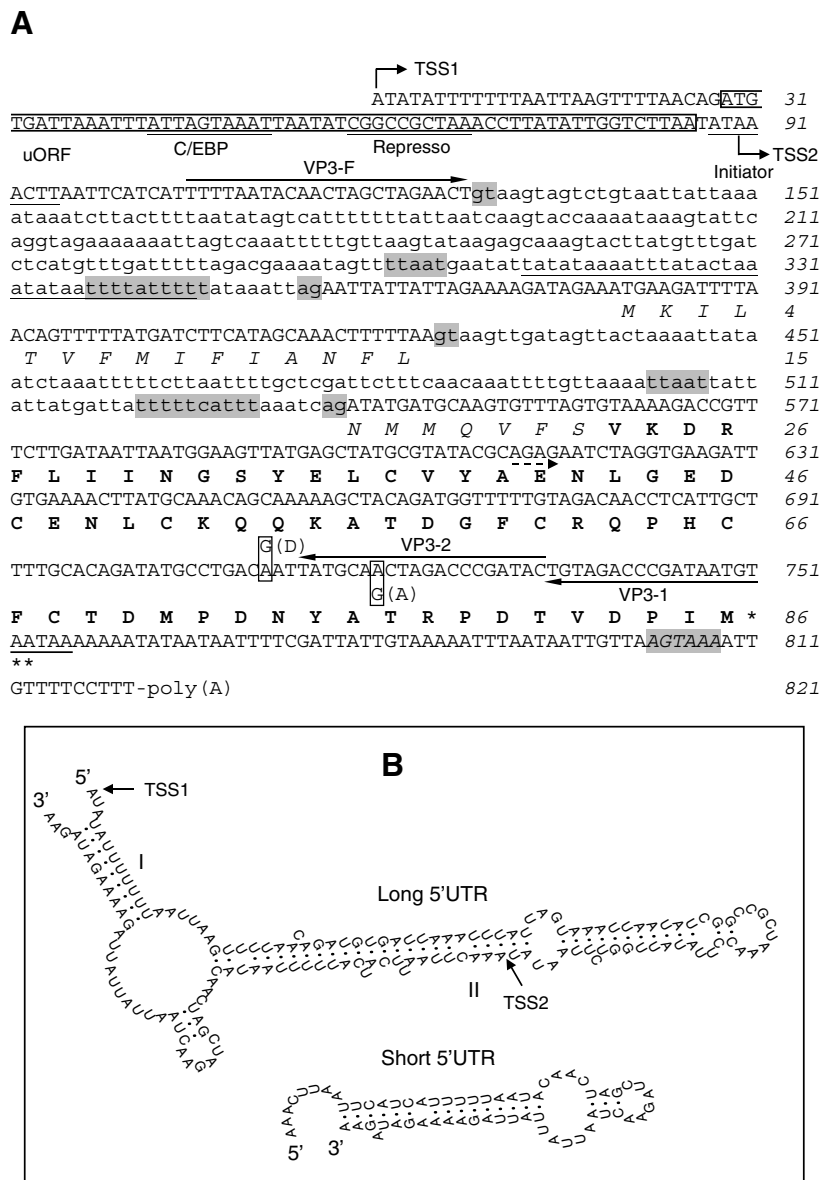

Fig. 1. The Mesotoxin gene. (A) Nucleotide and deduced amino acid sequences. Signal and mature peptides are respectively italicized and boldfaced. The putative polyadenylation signal (AGTAAA) is italicized and shadowed. Cleavage-poly (A) sites and intron splicing signals including splice sites, putative branch sites and the U2AF factor binding region are shadowed. Two non-synonymous replacement sites are boxed. Arrows labels the position of primers. Dotted arrow indicates the 5′ end of the truncated cDNA. The short stem–loop structure of 34 nucleotides in the intron 1 is underlined once. (B) Secondary structures of the long and short 5′UTRs.

longer than the small one. The small transcript is most likely due to alternative usage of the transcriptional start site (TSS) rather than a degraded product. Evidence comes from: (1) despite lacking of a TATA box at -30 nt relative to its TSS, the small transcript contains an initiator element (ATAAACTT) which perfectly matches the eukaryotic consensus $[(TC)(TC)A^{+1}N(TA)(TC)(TC)(TC)]$ (where A^{+1} is the base at which transcription starts, N is any of the four bases) [32]. Because the initiator is an alternative promoter element which is capable of replacing the function of the TATA box, it serves as the basal element determining the location of the TSS2; (2) In addition to the initiator, two other TF binding sites (Represso and C/EBPα) are also present in the upstream region of the TSS2, respectively located at 22 and 38 nt upstream; (3) The 5′UTR size of the small transcript is highly compatible with those of other scorpion venom peptide transcripts [33,34].

What is the biological significance of this alternative transcriptional initiation? A possibility is that this manner represents a regulatory strategy to control the expression level of Mesotoxin gene. Considering the existence of an upstream open reading frame (uORF) [35,36] and stable secondary structure in the 5′ UTR of the large transcript (ΔG of -15.97 kcal mol^{-1} vs -6.9 kcal mol^{-1} of the small one) [37], we hypothesize that this transcript probably have lower trans-lational efficiency than the small one. Thus, by producing two transcripts with different 5′UTRs, the expression level of the Mesotoxin can be efficiently regulated.

3.3. Gene structure of Mesotoxin

To determine the structure of the Mesotoxin gene, we undertook a PCR amplification of *M. martensii* genomic DNA. Comparison of the cDNA and genomic sequences revealed that Mesotoxin gene contains three exons and two introns, one being located within 5′UTR with a length of 228 nt, another at the end of signal peptide coding region comprising 113 nt. The latter is a phase I intron (an intron located between the first and second nucleotides of a codon) which split a small amino acid Asn and its location is conserved across all the members of scorpion toxins with α/β fold. Two introns have a consensus GT–AG splice junction and are also rich in A + T content (82%) (Fig. 1). Putative branch sites for introns 1 and 2 are respectively located upstream 51 nt and 32 nt of the 3′ splicing sites, and both have a pyrimidine tract between the branch site and 3′ splicing site, which likely binds a putative U2AF splicing factor. Mfold predicted that both introns 1 and 2 form stable secondary structures containing 2 to 4 hairpins with ΔG of -23.65 kcal mol^{-1} and ΔG of -10.6 kcal mol^{-1}, respectively (Fig. S1). Interestingly, the in-
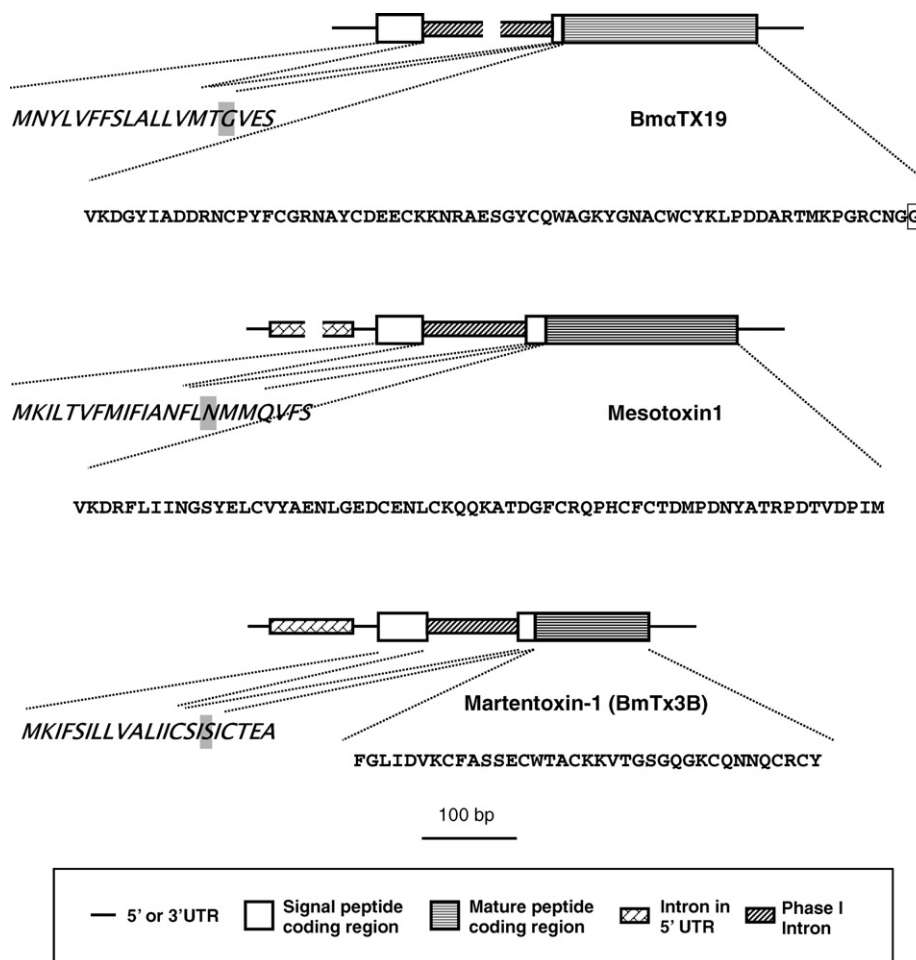

Fig. 2. Schematic representation of gene structures of different types of scorpion neurotoxins. Residues shadowed are the sites for the phase I intron insertion. BmαTX19 gene encoding an α-toxin-like peptide was isolated by us (GenBank accession number DQ872675). Data of the Martentoxin-1 gene is adopted from Ref. [45].

tron 1 contains a short stem-loop structure of 34 nt, which also tandem repeats seven times in the human genome (Fig. S1).

Generally, genes encoding the NaScTxs contain large introns ranging from 307 to 988 nt, whereas introns of scorpion toxins affecting potassium channels (KScTxs) vary from 74 to 125 nucleotides [38]. However, the gene organization of Mesotoxin1 is unique in size and number. In comparison with other scorpion neurotoxin genes, Mesotoxin1 shares the most similar gene organization to Martentoxin-1 (also called BmTx3B), a typical potassium channel blocker isolated from *M. martensii* [45] (Fig. 2).

3.4. Protein sequence analysis

Using the mature Mesotoxin sequence as a probe to carry out BLAST search of nonredundant GenPept database with default parameters, we retrieved 65 hits with E value ranging from 0.047 to 9.8 (E value means the expected number of distinct segment pairs that would obtain a score $\geqslant S$ by chance in a database search) [39], in which all are NaScTxs. BLAST ranked AaHVI, an anti-insect β-toxin, as the first hit with 33% identity and 59% similarity. Use of the full-length amino acid sequence gives similar results in spite of an α-toxin (BmαTx9) [34] as the first hit. Slightly different results obviously are due to the effect of signal peptides. Instead of β-toxins, remarkable similarity was found between signal peptides of Mesotoxins and α-toxins. Whereas the number of identical residues is similar between Mesotoxins and α- or β-toxins, more gaps are needed to make a suitable arrangement for Mesotoxins and α-toxins (Fig. 3A). Mesotoxins retain six conserved cysteines that can form three disulfide bridges. However, the wrapper bridge C1–C8 putting the N- and C-terminal regions together was lost during evolution. To provide insight into the evolutionary relationship between Mesotoxins and typical NaScTxs with four disulfides, we constructed a neighbor-joining tree based on mature sequences. In this tree, Mesotoxins are at the basal position with similar distance to both α- and β-toxins, suggesting its ancestral position in evolution. Two insect-specific toxins (AaHVI and AaHSTR1) from the Old World scorpion with structural and functional similarity to the β-toxins cluster together with the α-toxins from the Old World scorpion, favoring them as an evolutionary link between α- and β-toxins. The tree also supports monophyly of these different pharmacological groups, as previously revealed (Fig. 3B) [11–13].

3.5. Structural characteristic

Mesotoxin displays a typical β-toxin structural feature as characterized by its long J-loop and short B-loop (Fig. 3A). Fold compatibility of Mesotoxin with known 3D structures was analyzed using GenTHREADER, which first ranked the β-toxin CsEv2a (PDB entry 1JZA) as the most compatible structure with E-value 0.007 despite of only 24% sequence identity. In addition to six cysteines, some residues associated with structural stability of β-toxins are also conserved between Mesotoxins and CsEv2a (Fig. 4A). These residues either participate in the formation of hydrophobic core (e.g. Leu5, Val(Ile)6, Trp(Phe)47 and Leu(Met)51, residues are numbered according to CsEv2a) or are located in a turn to minimize the steric hindrance (e.g. Gly11 and Gly39) [25]. These conservations provide rationality for choosing CsEv2a as a template to build Mesotoxin1 structure. The structure modeled

shows satisfactory quality when checked using Verify3D (scoring above 0.2 (0.308)) and PROSA (scoring below −0.3 (−1.366)).

Overall, the structure of Mesotoxin1 is very similar to that of CsEv2a with a root-mean-squared deviation (rmsd) of 0.66 Å for 61 Cα atoms (Fig. 4B). The model structure of Mesotoxin1 also revealed the presence of a three-amino acid cluster (Fig. 4C) which is a conserved structural motif in the majority of NaScTxs. This cluster comprises one basic residue and two aromatic residues contributed by slightly different regions between α- and β-toxins. In the α-toxins, one aromatic residue is provided from the B-loop rather than the turn close to the N-terminus in the β-toxins [25]. In the Mesotoxin1 structure, Phe5, Tyr12 and Arg56 constitute this motif with equivalent positions Tyr4, Lys13 and Tyr58 in CsEv2a. In the α-toxin BmKM1, the motif is composed of Tyr5, Tyr42 and Arg58. Again, in this motif Mesotoxin1 much more resemble β-toxins. It has shown that modification of the basic residue of this cluster causes a marked loss of toxicity in some toxins [25]. For example, the Arg58 residue of BmKM1 located in a cavity surrounded by a hydrophobic gasket, as a key functional determinant, is responsible for the mammal toxicity of the toxin. A conserved replacement Arg58Lys in led to a large reduction of toxicity towards mice [40].

4. Discussion

Emergence of diverse bioactivity by reshaping functional surface in a conserved structural scaffold is a common theme for protein evolution [5]. An interesting structural feature in the NaScTx family is that the three buried disulfide bridges for stabilizing the scaffold are conserved across the family whereas the exposed wrapper disulfide bridge (WDB) varies in position among different toxins [15,22]. There is increasing awareness that alteration of the WDB linkage pattern can lead to functional switch of the NaScTxs via adjusting the conformation of key residues associated with toxin function [22]. Functional role of this unique disulfide has been highlighted in the scorpion α-toxin BmKM1 where mutation of its WDB resulted in a dramatic functional loss but kept the global structure unchanged. This effect likely stems from the local structural change of the NC functional domain due to the removal of the WDB [41].

Supported by these facts, it seems to be reasonable to infer that an ancient peptide with the three disulfide bridges might be the ancestor of the NaScTxs given its scaffold can easily serve as a platform to assembly different types of toxins by evolving the WDB in different positions. This is consistent with the prevailing view that disulfide bridges have been added during evolution to enhance the stability of proteins that functions in a fluctuating cellular environment [42]. Mesotoxin is the first NaScTx-like peptide characterized by its three disulfide bridges with compatible size to typical NaScTxs. Although functional data is not available at present, its evolutionary significance is obvious. The typical structural feature including J- and B-loop sizes and three-amino acid cluster location allow us to classify it into β-toxin group. Compatible size and disulfide pattern together with evolutionary position support its progenitor position in the origin of the NaScTx family. As discussed above, Mesotoxin of 64 residues may serve as a structural platform for diversifying into different types of NaScTxs.
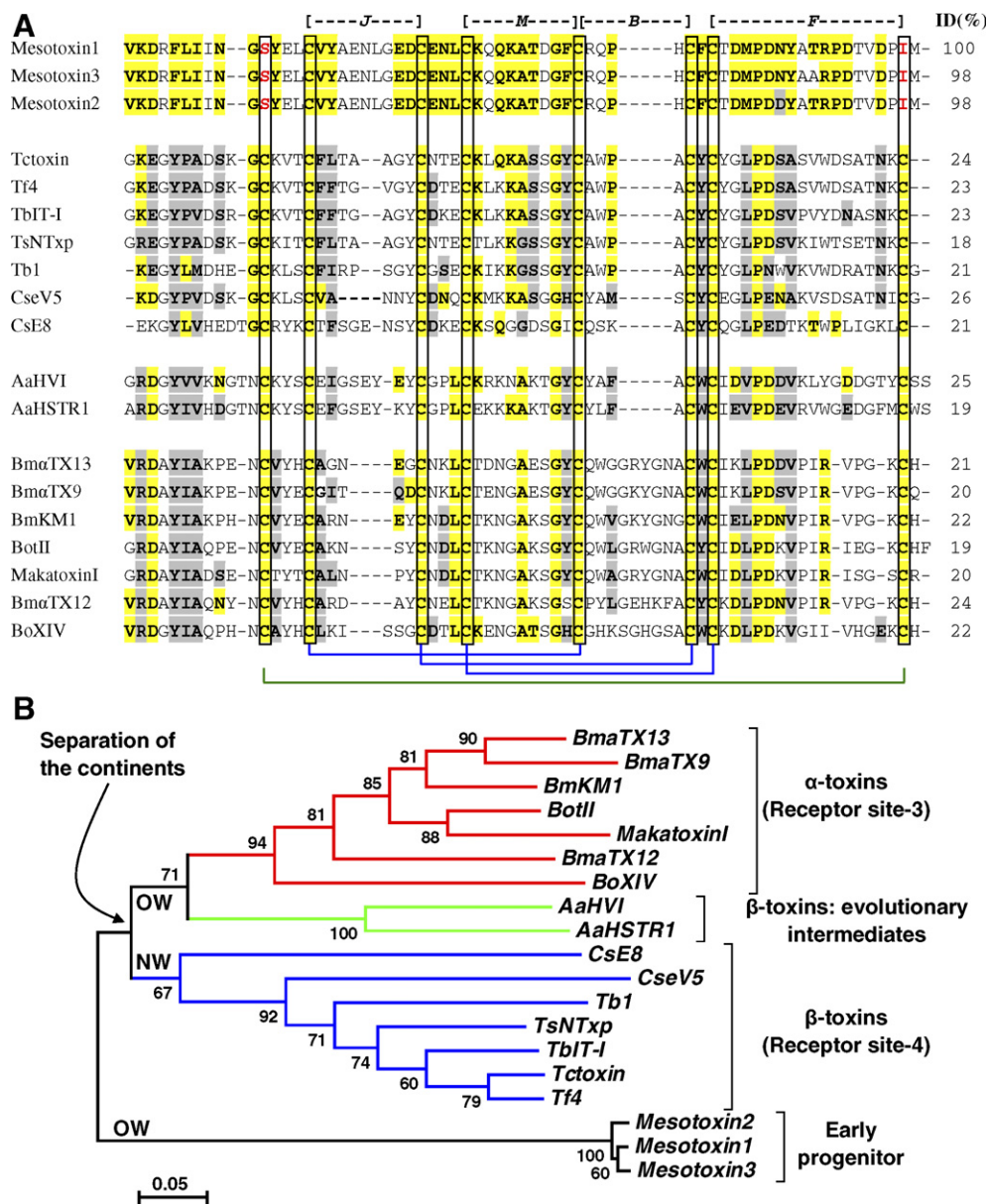
Fig. 3. Mesotoxins and related peptides. (A) Sequence alignment. Identical residues and conservation replacement to Mesotoxin1 are respectively shadowed in yellow and grey. Mesotoxin3 was derived from genomic amplification using two primers for constructing a recombinant expression vector (Table 1). Cysteines forming disulfides are boxed and residues in an equivalent position to Cys1 and Cys8 of the typical NaScTxs are highlighted in red. Three buried disulfides and the WDB are respectively indicated in blue and green lines. (B) Evolutionary analysis of scorpion α- and β-toxins. A bootstrap consensus tree is based upon 1000 replication of the neighbor-joining algorithm with p-distance. The tree is rooted with Mesotoxins. The scale bar shows total amino acid divergence. Two similar trees were also obtained using UPGMA and minimum evolution methods (results not shown). OW, Old World; NW, New World.

Despite independent evolution for typical α- and β-toxins [11–13], it appear that the WDB only originated once in these toxins with the WDB linking their N- and C-termini which possibly occurred before the separation of the continents due to the existence of evolutionary intermediates (AaHVI and AaHSTR1) in Old World scorpions which are insect-specific. To these toxins, a possible divergence route can be elucidated: adding a WDB in the Mesotoxin-like peptide scaffold led to the evolutionary intermediates which subsequently gave birth to the α-toxins acting on both mammal and insects by the extension of the B-loop in Old World scorpions. Extension of the B-loop could be a key event responsible for functional

loss of β-effect because a critical lysine side chain important for the β-toxins is shielded [25]. In New World scorpions, the ancestor with the WDB added produces typical β-toxins. Comparison of the functional surfaces of the α-toxin LqhαIT and the β-toxin Css4 revealed the existence of group-specific functional surface respectively located at the NC domain and the α-helical region [43,44]. However, a conserved region common to α- and β-toxins was found in the loop preceding the α-helix and the B loop (Fig. S2). We presumed that these two loops might represent the earliest exploited functional region responsible for the channel binding in the three-disulfide ancestor, such as Mesotoxins.

A

Fig. 4. Structural analysis of Mesotoxin1. (A) Structure-based sequence alignment. The key to the structural environment of each amino acid residue is based on formatting of the alignment by JOY [46]. The structural features of Mesotoxin1 are taken from the 3D model. Triangles indicate functionally important sites for the β-toxin Css4 [44]. (B) Ribbon diagram of the 3D model of Mesotoxin1 (right) and structure of CsEv2a (left). Two amino acids equivalent to Cys1 and Cys8 are highlighted in cyan sticks. (C) The exposed three-amino acid cluster in the structure of Mesotoxin1 is displayed in stick representation.

From an evolutionary perspective, the first NaScTx should affect insect rather than mammal sodium channels given rapid radiation of insects in the early history of scorpion evolution is the most selective factors driving the emergence of the Na-ScTx, whereas toxicity to mammals was later evolved for de-fense of these animals when they appeared on the earth. This explains the existing cause of the β-toxins specific for insect so-dium channels in Old World scorpions. Moreover, Mesotoxin possesses β-toxin features, again suggesting that the ancestor of the NaScTxs should be the β-toxin rather than the α-toxin

[19]. Also, our proposal presented here is not contradictory to the hypothesis that Drosomycin might be the ancestor of the NaScTxs [22] if we consider the adding of a NC region in the Drosomycin scaffold as the first evolutionary step in which the WDB did not form yet. Finally, it is also worth mentioning that more similarity in gene organization (location and size) between Mesotoxin and Martentoxin-1 strengthens their evolutionary link and also hints a possible ancestor role of Mesotoxins involved in the origin of short-chain scorpion toxins.

Regardless of the origin, the discovery of Mesotoxins provides a start point for experimental study of the evolutionary biology of the NaScTx family. Several immediate experiments can be followed, which include: (1) confirming Mesotoxin's β-effect acting on insect sodium channels and determining whether its functional surface is situated in the two loops; (2) by adding the WDB in the Mesotoxin scaffold to check the functional consequence of this modification.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.febslet.2006.09.071.

References

[1] Polis, G.A., Ed., (1990). The Biology of Scorpions, Stanford University Press, Stanford, CA.
[2] Possani, L.D., Becerril, B., Delepierre, M. and Tytgat, J. (1999) Scorpion toxins specific for Na$^+$-channels. Eur. J. Biochem. 264, 287–300.
[3] Possani, L.D., Merino, E., Corona, M., Bolivar, F. and Becerril, B. (2000) Peptides and genes coding for scorpion toxins that affect ion-channels. Biochimie 82, 861–868.
[4] Tytgat, J., Chandy, K.G., Garcia, M.L., Gutman, G.A., Martin-Eauclaire, M.F., van der Walt, J.J. and Possani, L.D. (1999) A unified nomenclature for short-chain peptides isolated from scorpion venoms: alpha-KTx molecular subfamilies. Trends Pharmacol. Sci. 20, 444–447.
[5] Gurevitz, M., Gordon, D., Ben-Natan, S., Turkov, M. and Froy, O. (2001) Diversification of neurotoxins by C-tail 'wiggling': a scorpion recipe for survival. FASEB J. 15, 1201–1205.
[6] Liang, S. (2004) An overview of peptide toxins from the venom of the Chinese bird spider *Selenocosmia huwena* Wang *Ornithoctonus huwena* (Wang). Toxicon 43, 575–585.
[7] Terlau, H. and Olivera, B.M. (2004) Conus venoms: a rich source of novel ion channel-targeted peptides. Physiol. Rev. 84, 41–68.
[8] Oguiura, N., Boni-Mitake, M. and Radis-Baptista, G. (2005) New view on crotamine, a small basic polypeptide myotoxin from South American rattlesnake venom. Toxicon 46, 363–370.
[9] Catterall, W.A. (2000) From ionic currents to molecular mechanisms: The structure and function of voltage-gated sodium channels. Neuron 26, 13–25.
[10] Cestéle, S. and Catterall, W.A. (2000) Molecular mechanisms of neurotoxin action on voltage-gated sodium channels. Biochimie 82, 883–892.
[11] Zhu, S., Bosmans, F. and Tytgat, J. (2004) Adaptive evolution of scorpion sodium channel toxins. J. Mol. Evol. 58, 145–153.
[12] Froy, O., Sagiv, T., Poreh, M., Urbach, D., Zilberberg, N. and Gurevitz, M. (1999) Dynamic diversification from a putative common ancestor of scorpion toxins affecting sodium, potassium, and chloride channels. J. Mol. Evol. 48, 187–196.
[13] Rodriguez de la Vega, R.C. and Possani, L.D. (2005) Overview of scorpion toxins specific for Na$^+$ channels and related peptides: biodiversity, structure-function relationships and evolution. Toxicon 46, 831–844.
[14] Gordon, D., Gilles, N., Bertrand, D., Molgo, J., Nicholson, G.M., Sauviat, M.P., Benoit, E., Shichor, I., Lotan, I., Gurevitz, M., Kallen, R.G. and Heinemann, S.H. (2002) in: Perspectives in Molecular Toxinology (Ménez, A., Ed.), pp. 215–238, Wiley, West Sussex.
[15] Inceoglu, B., Lango, J., Wu, J., Hawkins, P., Southern, J. and Hammock, B.D. (2001) Isolation and characterization of a novel type of neurotoxic peptide from the venom of the South African scorpion *Parabuthus transvaalicus* (Buthidae). Eur. J. Biochem. 268, 5407–5413.
[16] Inceoglu, B., Hayashida, Y., Lango, J., Ishida, A.T. and Hammock, B.D. (2002) A single charged surface residue modifies the activity of ikitoxin, a beta-type Na$^+$ channel toxin from *Parabuthus transvaalicus*. Eur. J. Biochem. 269, 5369–5376.
[17] Inceoglu, B., Lango, J., Pessah, I.N. and Hammock, B.D. (2005) Three structurally related, highly potent, peptides from the venom of *Parabuthus transvaalicus* posses divergent biological activity. Toxicon 45, 727–733.
[18] Srairi-Abid, N., Guijarro, J.I., Benkhalifa, R., Mantegazza, M., Cheikh, A., Ben Aissa, M., Haumont, P.Y., Delepierre, M. and El Ayeb, M. (2005) A new type of scorpion Na$^+$-channel-toxin-like polypeptide active on K$^+$ channels. Biochem. J. 388, 455–464.
[19] Froy, O. and Gurevitz, M. (2003) New insight on scorpion divergence inferred from comparative analysis of toxin structure, pharmacology and distribution. Toxicon 42, 549–555.
[20] Zeng, X., Luo, F. and Li, W. (2006) Molecular dissection of venom from Chinese scorpion *Mesobuthus martensii*: Identification and characterization of four novel disulfide-bridged venom peptides. Peptides 27, 1745–1754.
[21] Soudani, N., Gharbi-Chihi, J., Srairi-Abid, N., Yazidi, C.M., Planells, R., Margotat, A., Torresani, J. and Ayeb, M.E. (2005) Isolation and molecular characterization of LVP1 lipolysis activating peptide from scorpion *Buthus occitanus tunetanus*. Biochim. Biophys. Acta 1747, 47–56.
[22] Zhu, S., Gao, B. and Tytgat, J. (2005) Phylogenetic distribution, functional epitopes and evolution of the CSαβ superfamily. Cell. Mol. Life Sci. 62, 2257–2269.
[23] Zhu, S., Darbon, H., Dyason, K., Verdonck, F. and Tytgat, J. (2003) Evolutionary origin of inhibitor cystine knot peptides. FASEB J. 17, 1765–1767.
[24] Jones, D.T. (1999) GenTHREADER: an efficient and reliable protein fold recognition method for genomic sequences. J. Mol. Biol. 287, 797–815.
[25] Cook, W.J., Zell, A., Watt, D.D. and Ealick, S.E. (2002) Structure of variant 2 scorpion toxin from *Centruroides sculpturatus* Ewing. Protein Sci. 11, 479–486.
[26] Eisenberg, D., Luthy, R. and Bowie, J.U. (1997) VERIFY3D: assessment of protein models with three-dimensional profiles. Methods Enzymol. 277, 396–404.
[27] Sippl, M.J. (1993) Recognition of errors in three-dimensional structures of proteins. Proteins 17, 355–362.
[28] Guex, N. and Peitsch, M.C. (1996) Swiss-Pdb Viewer: a fast and easy-to-use PDB viewer for Macintosh and PC. Prot. Data Bank Q. Newslett. 77, 7.
[29] Koradi, R., Billeter, M. and Wuthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. J. Mol. Graph. 14, 51–55.
[30] Diego-Garcia, E., Batista, C.V., Garcia-Gomez, B.I., Lucas, S., Candido, D.M., Gomez-Lagunas, F. and Possani, L.D. (2005) The Brazilian scorpion *Tityus costatus* Karsch: genes, peptides and function. Toxicon 45, 273–283.
[31] Wagner, S., Castro, M.S., Barbosa, J.A., Fontes, W., Schwartz, E.N., Sebben, A., Rodrigues Pires Jr., O., Sousa, M.V. and Schwartz, C.A. (2003) Purification and primary structure determination of Tf4, the first bioactive peptide isolated from the venom of the Brazilian scorpion *Tityus fasciolatus*. Toxicon 41, 737–745.
[32] Lodish, H., Berk, A., Zipursky, S.L., Matsudaira, P., Baltimore, D. and Darnell, J. (2000) Molecular Cell Biology, W. H. Freeman and Company, New York.

[33] Zhu, S., Li, W., Zeng, X., Jiang, D., Mao, X. and Liu, H. (1999) Molecular cloning and sequencing of two 'short chain' and two 'long chain' K$^+$ channel-blocking peptides from the Chinese scorpion *Buthus martensii* Karsch. FEBS Lett. 457, 509–514.

[34] Zhu, S., Li, W., Zeng, X., Liu, H., Jiang, D. and Mao, X. (2000) Nine novel precursors of *Buthus martensii* scorpion alpha-toxin homologues. Toxicon 38, 1653–1661.

[35] Meijer, H.A. and Thomas, A.A.M. (2002) Control of eukaryotic protein synthesis by upstream open reading frames in the 5′-untranslated region of an mRNA. Biochem. J. 367, 1–11.

[36] Huang, H., Chien, C., Jen, K. and Huang, H. (2006) RegRNA: an integrated web server for identifying regulatory RNA motifs and elements. Nucleic Acids Res. 34, W429–W434.

[37] Hess, M.A. and Duncan, R.F. (1996) Sequence and structure determinants of *Drosophila* Hsp70 mRNA translation: 54-UTR secondary structure specifically inhibits heat shock protein mRNA translation. Nucleic Acids Res. 24, 2441–2449.

[38] Zhu, S., Cao, Z. and Li, W. (2001) Analysis of splicing signals of scorpion venom peptide introns. Acta Biophys. Sin. 17, 695–702.

[39] Sharon, I., Birkland, A., Chang, K., EL-Yaniv, R. and Yona, G. (2005) Correcting BLAST e-values for low-complexity segments. J. Comput. Biol. 12, 978–1001.

[40] Ye, X., Bosmans, F., Li, C., Zhang, Y., Wang, D.C. and Tytgat, J. (2005) Structural basis for the voltage-gated Na$^+$ channel selectivity of the scorpion alpha-like toxin BmKM1. J. Mol. Biol. 353, 788–803.

[41] Sun, Y.M., Liu, W., Zhu, R.H., Goudet, C., Tytgat, J. and Wang, D.C. (2002) Roles of disulfide bridges in scorpion toxin BmKM1 analyzed by mutagenesis. J. Pept. Res. 60, 247–256.

[42] Hogg, P.J. (2003) Disulfide bonds as switches for protein function. Trends Biochem. Sci. 28, 210–214.

[43] Karbat, I., Frolow, F., Froy, O., Gilles, N., Cohen, L., Turkov, M., Gordon, D. and Gurevitz, M. (2004) Molecular basis of the high insecticidal potency of scorpion α-toxins. J. Biol. Chem. 279, 31679–31686.

[44] Cohen, L., Karbat, I., Gilles, N., Ilan, N., Benveniste, M., Gordon, D. and Gurevitz, M. (2005) Common features in the functional surface of scorpion beta-toxins and elements that confer specificity for insect and mammalian voltage-gated sodium channels. J. Biol. Chem. 280, 5045–5053.

[45] Ji, Y.H., Wang, W.X., Ye, J.G., He, L.L., Li, Y.J., Yan, Y.P. and Zhou, Z. (2003) Martentoxin, a novel K$^+$-channel-blocking peptide: purification, cDNA and genomic cloning, and electrophysiological and pharmacological characterization. J. Neurochem. 84, 325–335.

[46] Mizuguchi, K., Deane, C.M., Blundell, T.L., Johnson, M.S. and Overington, J.P. (1998) JOY: protein sequence-structure representation and analysis. Bioinformatics 14, 617–623.