5624–5629 *Nucleic Acids Research, 1998, Vol. 26, No. 24* *© 1998 Oxford University Press*

Genetic characterization of the mRNAs encoding α-bungarotoxin: isoforms and RNA editing in *Bungarus multicinctus* gland cells

Li-Feng Liu, Chun-Chang Chang, Ming-Yi Liau[1] and Kou-Wha Kuo*

Department of Biochemistry, Kaohsiung Medical College, Kaohsiung 807, Taiwan and [1]Department of Health, National Institute of Preventive Medicine, Taipei 115, Taiwan

Received August 31, 1998; Revised and Accepted November 2, 1998

DDBJ/EMBL/GenBank accession nos AF056400–AF056417

ABSTRACT

The mRNA encoding α-bungarotoxin (α-Butx) was prepared from the venom glands of *Bungarus multicinctus* by Cap-finder cDNA synthesis. The sequences of the 3′- and 5′-flanking regions including a signal peptide of α-Butx were almost identical with those of Elapidae and Hydrophiidae toxins, suggesting that they may have the same origin. Sixteen polymorphic mRNA sequences of α-Butx were detected from *B.multicinctus* gland cells. Analysis of the genomic DNA of α-Butx indicated that the polymorphic mRNA originated from one DNA sequence. Most of the mutations in α-Butx mRNA were silent and the hot-spot variations occurred at 78, 107, 129, 198 and 201 nt in α-Butx mRNA. Ten distinct protein sequences of α-Butx could be deduced from the polymorphic mRNA and one of the isoforms has already been isolated. Since α-Butx DNA is a single copy in the genome, the RNA polymorphism might result from post-transcriptional editing. These results indicate that the authentic α-Butx is in fact derived from edited mRNAs. RNA editing may contribute a common mechanism toward the diversity of α-neurotoxins in snake glands.

INTRODUCTION

The postsynaptic α-neurotoxins derived from the venoms of Elapidae and Hydrophiidae snakes bind selectively with high affinity to the nicotinic acetylcholine receptor (nAChR) to block neuromuscular transmission (1,2). So far, >80 primary structures of α-neurotoxins have been sequenced (3) and can be categorized into short neurotoxins (60–62 residues and four disulfide bridges) and long neurotoxins (66–74 residues and five disulfide bridges). Both classes of neurotoxins share a significant similarity of amino acid sequence and overall topology, which is characterized as a triple-stranded anti-parallel β-sheet and three finger-like loops protruding from a globular core (4). The sequence comparisons among the known α-neurotoxins have identified 12 residues that are highly conserved throughout the long and short neurotoxins. These residues were postulated to be involved either directly or indirectly in binding to nAChR (5,6).

α-Bungarotoxin (α-Butx), a long neurotoxin comprised of 74 residues and five disulfide bonds, was isolated from the venom of the Taiwan banded krait (*Bungarus multicinctus*). Because of its extremely high affinity for nAChR from vertebrate muscles, electric ray (*Torpedo*) and electric eel (*Electrophorus*) organs, α-Butx has provided a critically important tool for studies directed at understanding neuromuscular transmission and structure–function of nAChR (7,8). The X-ray crystal structure of α-Butx has previously been reported (9). Refinements to the structure and the complex structure of a library-derived peptide have been solved by two-dimensional NMR (10–12). It has been shown that the Asp[30] and Arg[36] residues in α-Butx form an ion pair near the tip of loop II that mimics the structure of acetylcholine, thereby contributing to toxin specificity (13). Despite the fact that recombinant α-Butx can be prepared by functional expression of a constructed gene (14) or a cDNA from the cellular RNA of venom glands (15), the complete mRNA sequence of α-Butx and the post-transcriptional process in gland cells are still unclear. RNA editing, defined as an alteration in the coding capacity of mRNA other than splicing or 3′-end processing, was observed in α-Butx mRNA. The variations in the mRNA sequence of α-Butx were determined and the polymorphism of α-Butx mRNA might result from post-transcriptional editing. These results will reveal the genetic information of α-Butx and may provide insight into the regulation of the diversity of α-neurotoxins in snake glands.

MATERIALS AND METHODS

Molecular cloning and sequence analysis of full-length α-Butx cDNA

The venom glands of *B.multicinctus* were provided by the Department of Health, National Institute of Preventive Medicine of Taiwan. The cellular RNA was isolated from the glands, which had been stored in liquid nitrogen immediately after death. Two deep frozen glands from one snake were homogenized to extract the total RNA by a guanidinium isothiocyanate/phenol chloroform isolation kit (Stratagene, La Jolla, CA). α-Butx cDNA was generated by the Cap-finder cDNA synthesis method (Clontech, Palo Alto, CA), in which a modified oligo(dT) primer (CDS/3′PCR primer) was used to generate the first-strand cDNA. In this method, the CapSwitch oligonucleotide serves as a short and extended template at the 5′-end for the reverse transcription.

*To whom correspondence should be addressed. Tel: +886 7 312 1101; Fax: +886 7 321 8309; Email: kuokw@cc.kmc.edu.tw

The MMLV enzyme switches template and continues replicating to the end of the CapSwitch oligonucleotide when reverse transcription reaches the 5′-end of the mRNA. This switching occurs at the 7-methylguanosine cap structure, which is present on the 5′-end of all eukaryotic mRNAs (16). The full-length α-Butx cDNA was then obtained by long distance PCR with 5′PCR, CDS/3′PCR and gene-specific primers (BXA and BXB) of α-Butx. The 5′- and 3′-cDNA fragments were amplified with a high fidelity *Pfu* DNA polymerase (Stratagene) with the primer combinations 5′PCR and BXB, and CDS/3′PCR and BXA, respectively. The primer sequences of BXA and BXB were 5′-CGGGATCCATCGTGTGCCATACAAC-3′ (addition of a *Bam*HI site at the 5′-end) and 5′-GGAATTCCTATCCAGGTCT-TTGTTTAG-3′ (addition of an *Eco*RI site and one stop codon at the 5′-end). The PCR products were cloned into pCR-Script™ Amp SK(+) vectors according to the manufacturer's protocol (Stratagene). The inserts were sequenced using fluorescence-labeled dideoxyribonucleotides (ABI Prism Dye-Terminator; Perkin-Elmer, CA) on an ABI 373A DNA autosequencer (Applied Biosystems).

Analysis of the genomic DNA sequence of the α-Butx coding region

The genomic DNA was extracted from the liver of *B.multicinctus*. The organ was gently homogenized in a Potter B homogenizer at 800 r.p.m. in PBS supplemented with 2 mM EDTA, pH 7.4. The centrifugation pellet (0.2 ml) was dissolved in 1.8 ml of lysis buffer (10 mM Tris–HCl, 400 mM NaCl, 2 mM EDTA, pH 8.0, 0.5% SDS) and incubated in the presence of 0.2 mg/ml proteinase K (Sigma, St Louis, MO) at 37°C for 1 h. The lysate was continuously incubated in the presence of 0.02 mg/ml heat-treated RNase (Gibco BRL, Grand Island, NY) for an additional 1 h. The DNA was gently extracted (1:1 v/v) with phenol:chloroform: isoamyl alcohol (25:24:1) and then precipitated by adding 0.5 vol 7.5 M ammonium acetate and 2 vol ethanol. The DNA was monitored on a 0.3% agarose gel. The fragment size should be >50 kb. The PCR was conducted using the genomic DNA, α-Butx flanking primers (BXA and BXB) and *Pfu* DNA polymerase to amplify the protein coding region of α-Butx DNA. The products were cloned into pCR-Script™ vectors and the inserts were sequenced by autosequencer as described above.

Two-dimensional polyacrylamide gel electrophoresis and western blotting

HPLC-purified α-Butx from the venom of *B.multicinctus* was purchased from Sigma. α-Butx (50 µg) was diluted 3-fold with an IPG sample buffer (9 M urea, 65 mM DTE, 1% v/v carrier ampholytes 3–10, 65 mM CHAPS, 35 mM Tris, 5% 2-mercapto-ethanol and 0.1% bromophenol blue). IPG strips were rehydrated in the sample buffer for 10 min. The first dimension separation was performed in a 4.5% acrylamide gel in a glass tube of 2.2 × 150 mm. The cathode electrode buffer was 0.1 M ethylene diamine and the anode electrode buffer was 0.01 M iminodiacetic acid. The gels were pre-run at 300 V for 30 min and then at 400 V for 30 min. After loading the sample solution, the gel was run at 500 V for 20 h and then at 800 V for 1 h. The gel was removed from the tube and equilibrated in a solution containing 10% glycerol, 5% 2-mercaptoethanol, 2.3% SDS and 0.0625 M Tris–HCl (pH 6.8) for <5 min. The second dimension was performed as a vertical 2-dimensional PAGE in a uniform gel (1 × 140 × 130 mm) composed of 3.4% polyacrylamide in the stacking gel and 12.5% in the separation gel to which 1% SDS was added. The running buffer contained 0.025 M Tris base, 0.192 M glycine and 0.1% SDS. Electrophoresis was performed at a constant current of 10 mA/gel for the separation phase. All the chemicals for electrofocusing and PAGE were of the highest grade available.

Proteins were separated by SDS–PAGE on a 12.5% (w/v) gel and transferred onto a nitrocellulose membrane (Hybond C; Amersham, Arlington Heights, IL) according to established procedures. The molecular weight of recombinant protein was estimated by comparing the standard markers (Bio-Rad, Hercules, CA). Detection of α-Butx was performed with anti-α-Butx serum (1:300 dilution). Goat anti-rabbit IgG conjugated with alkaline phosphatase (Bio-Rad) was added to detect α-Butx antibody according to the standard procedure.

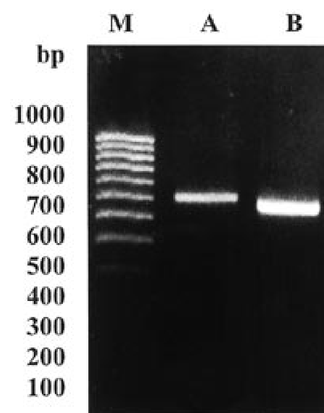

Figure 1. Cap-finder PCR cDNA synthesis of the 5′- and 3′-regions of α-Butx cDNA. Procedures for the generation of α-Butx 5′- and 3′-cDNAs are described in Materials and Methods. The PCR products were separated on a 2% agarose gel with ethidium bromide staining. M, 100 bp DNA ladder marker; A, 5′-cDNA, and B, 3′-cDNA of α-Butx.

RESULTS

Molecular cloning and sequencing of the full-length α-Butx cDNA

To determine the complete mRNA sequence of α-Butx from the venom gland cells of *B.multicinctus*, Cap-finder cDNA synthesis, which is advantageous for the analysis of the 5′-ends of eukaryotic genes, was employed. The mRNA was reverse transcribed into cDNA by MMLV reverse transcriptase according to the manufacturer's protocol. As shown in Figure 1, the 5′- and 3′-cDNA segments of α-Butx were amplified by a high fidelity *Pfu* DNA polymerase using commercial CDS/3′PCR and 5′PCR CapSwitch primers and α-Butx-specific primers (BXA and BXB). After cloning of the cDNA segments into pCR-Script vectors, the full-length cDNA sequence of α-Butx was analyzed. As shown in Figure 2, it is ~530 bp consisting of a 5′-untranslated region (UTR) (33 bp), signal peptide (63 bp), α-Butx coding region (222 bp) and 3′-UTR (195 bp) with a TGA termination codon and an AATAAA signal polyadenylation sequence located 13 nt upstream of the site of cleavage. The complete nucleotide sequence of α-Butx was submitted to a computer-aided sequence similarity search in databases. Astonishingly, the sequences of the

```
GAC ACT TTG CAG GCT CCC AGA GAA GAT CGC AAG ATG AAA ACT CTG   45
                                             M   K   T   L     4
TTG CTG ACC TTG GTG GTG GTG ACA ATC GTG TGC CTG GAC TTA GGA   90
 L   L   T   L   V   V   V   T   I   V   C   L   D   L   G    19
TAT ACC ATC GTA TGC CAC ACA ACA GCT ACT TCG CCT ATT AGC GCT  135
 Y   T   I   V   C   H   T   T   A   T   S   P   I   S   A    34
GTG ACT TGT CCA CCT GGG GAG AAC CTA TGC TAT AGA AAG ATG TGG  180
 V   T   C   P   P   G   E   N   L   C   Y   R   K   M   W    49
TGT GAT GCA TTC TGT TCC AGC AGA GGA AAG GTA GTC GAA TTG GGG  225
 C   D   A   F   C   S   S   R   G   K   V   V   E   L   G    64
TGT GCT GCT ACT TGC CCT TCA AAG AAG CCC TAT GAG GAA GTT ACC  270
 C   A   A   T   C   P   S   K   K   P   Y   E   E   V   T    79
TGT TGC TCA ACA GAC AAG TGC AAC CCA CAT CCG AAA CAG AGA CCT  315
 C   C   S   T   D   K   C   N   P   H   P   K   Q   R   P    94
GGT TGA GTT TTG CTC TCA TCC ATC ATG GAC CAT CCT TGA AAA TTT  360
 G   *
ATG CTT GTG GCC TTT ACC ACC AGA TGG TCC ATC ATT CCC CTC TCC  405
CCT GCT GTC TTT GAC ATC TCA ACA TCT TTC CCT TTT TCT CTC ATT  450
CTG TAA GTT TCC TTC TGC TAG TTC TGT AGT TTT GAG AAT CAA ATA  495
AAC CTC AGC ATT CAC AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA  540
```

Figure 2. Nucleotide sequence of α-Butx cDNA. The amino acid sequence of the signal peptide is underlined. A polyadenylation signal (AATAAA) at the 3′-end is marked with bold letters. The complete cDNA sequence of α-Butx has been deposited in the EMBL/GenBank database under accession no. AF056400.

5′- and 3′-UTRs and the signal sequences of α-Butx were almost identical with those of toxins from Elapidae and Hydrophiidae snakes (Fig. 3). The results imply that these snakes may share a common evolutionary origin.

RNA polymorphism of the protein coding region of α-Butx

To characterize the presence of isoformic cDNA sequences of α-Butx in the cellular RNA, >100 positive clones of the cDNA containing 5′- or 3′-segments of α-Butx were analyzed in a DNA autosequencer. The sequences of the 5′- and 3′-UTRs and the signal peptide segments in α-Butx were highly consistent, suggesting that the regions might have been important during the evolution of these snakes. Nevertheless, a total of 16 polymorphic mRNA sequences of the protein coding region of α-Butx were identified (Fig. 4). Most mutations in the α-Butx sequences were silent. The 'hot-spot' variations occurred at 78, 107, 129, 198 and 201 nt and a nonsense mutation occurred at 163 nt of the α-Butx mRNA (R_7) in which GAG (Glu^{55}) was changed to TAG (stop codon). In addition, 10 distinct amino acid sequences of α-Butx could be deduced from the polymorphic mRNA sequences. Two major protein sequences of α-Butx in which the R_8, R_9, R_{10} and R_{11} mRNA sequences translated into authentic α-Butx and R_1, R_2 and R_3 into a major isoform with the variation G→A at nt 107, resulting in substitution of Arg^{36} by Lys^{36}, were detected. Nevertheless, the polymorphic α-Butx mRNAs could be categorized into two major groups according to the amino acid at position 36 in the sequence, i.e. Lys36 α-Butx (α-ButxK^{36}) and Arg36 α-Butx (α-ButxR^{36}). R_1 α-Butx mRNA seemed to be the original transcript from the genomic DNA. After transcription, all the α-ButxK^{36} transcripts were accompanied by silent mutations at nt 198 (AA<u>C</u>→AA<u>T</u>) and nt 201 (CC<u>A</u>→CC<u>T</u>) encoding Asn^{66} and Pro^{67} in the α-Butx sequence, respectively. In contrast, those positions in

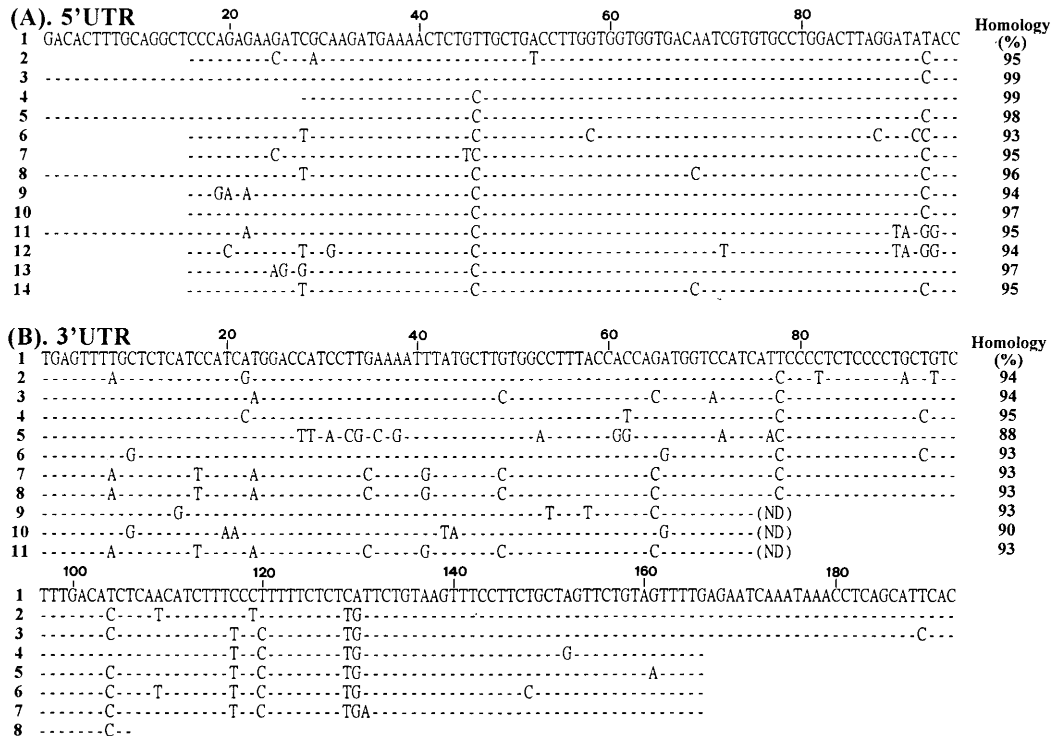

Figure 3. Sequence similarity of the 5′- and 3′-flanking regions of α-Butx and other toxins. From *Bungarus multicinctus*: 1, α-Butx; 2, neurotoxin-like peptide 1 (accession no. X64593); 3, neurotoxin homolog (X51414); 4, κ-neurotoxin (CB1) (X51412). From *Dendroaspis angusticeps*: 5, synergistic-like toxin (X51467); 6, angusticeps-type protein (DaF8) (X53409); 7, muscarinique 1 (X52292). From *Aipysurus laevis*: 8, toxin B (X13373); 9, toxin D (X13372). From *Laticauda semifasciata*: 10, erabutoxin a (X02533); 11, erabutoxin b (X16950); 12, erabutoxin c (X51410). From *Naja naja atra*: 13, cardiotoxin 4 (Y12493); 14, cobrotoxin (Y12492). ND at the 3′-end indicates no sequence determination after this point.

```
                        20                  40                  60                  80                  100                 120
#    ATCGTATGCCACACAACAGCTACTTCGCCTATTAGCGCTGTGACTTGTCCACCTGGGGAGAACCTATGCTATAGAAAAATGTGGTGTGATGCATTCTGTTCCAGCAAAGGAAAGGTAGTCGAATTGGGATG
R1   ...................................................................................................................................
R2   ...................................................................................................................................
R3   ...................................................................................................................................
R4   ..............................................................................G....................................................
R5   .............................................................................G.........................A...........................
R6   ...................................................................................................................................
R7   ...................................................................................................................................
R8   .............................................................................G............................G.....................G..
R9   .............................................................................G............................G.....................G..
R10  .............................................................................G............................G.....................G..
R11  ..........................................................................................................G.....................G..
R12  .............................................................................G............................G.....................G..
R13  ...........................................................................................T..C...........G.....................G..
R14  .............................................................................G.............T..............G.....................G..
R15  ..........................................................................................................G.....................G..
R16  ..............................................................................................C...........G.....................G..

            140                 160                 180                 200                 220
#    TGCTGCTACTTGCCCTTCAAAGAAGCCCTATGAGGAAGTTACCTGTTGCTCAACAGACAAGTGCAACCCACATCCGAAACAGAGACCTGGT
R1   ...........................................................................................
R2   ..................................................................T..T.....................
R3   ......C...........................................................T..T.....................
R4   ..................................................................T..T.....................
R5   ..................................................................T..T.....................
R6   .............T....................................................T..T.....................
R7   .......................G.......T..................................T..T.....................
R8   .........................................G.................................................
R9   ...........................................................................................
R10  ..................................................................T..T.....................
R11  ...........................................................................................
R12  ..........................................................C.....C..........................
R13  ...........................................................................................
R14  ...........................................................................................
R15  ...................G....................T..................................................
R16  .......C...........G....................T..................................................
```

#	Position of the mutation	pI	%
R1		7.93	3.33
R2	66, 67	7.93	65.00
R3	46, 66, 67	7.93	1.67
R4	$27^{Met\text{-}Val}$, 66, 67	7.93	1.67
R5	26, $35^{Ser\text{-}Asn}$, 66, 67	7.93	1.67
R6	$49^{Pro\text{-}Ser}$, 66, 67	7.93	1.67
R7	$52^{Lys\text{-}Arg}$, $55^{Glu\text{-}STOP}$, 66, 67	7.93	1.67
R8	26, $36^{Lys\text{-}Arg}$, 43, $58^{Thr\text{-}Ser}$	7.93	1.67
R9	$36^{Lys\text{-}Arg}$, 26, 43	7.93	5.00
R10	$36^{Lys\text{-}Arg}$, 26, 43, 66, 67	7.93	5.00
R11	$36^{Lys\text{-}Arg}$, 43	7.93	3.33
R12	26, $36^{Lys\text{-}Arg}$, $64^{Lys\text{-}Gln}$, $66^{Asn\text{-}His}$, 43	7.93	1.67
R13	$31^{Ala\text{-}Val}$, $32^{Phe\text{-}Ser}$, $36^{Lys\text{-}Arg}$, 43	7.93	1.67
R14	26, $31^{Ala\text{-}Val}$, $36^{Lys\text{-}Arg}$, 43	7.93	1.67
R15	$36^{Lys\text{-}Arg}$, 43, $51^{Lys\text{-}Glu}$, $58^{Thr\text{-}Ser}$	7.38	1.67
R16	$32^{Phe\text{-}Ser}$, $36^{Lys\text{-}Arg}$, 43, $47^{Thr\text{-}Pro}$, $51^{Lys\text{-}Glu}$, $58^{Thr\text{-}Ser}$	7.38	1.67

Figure 4. The polymorphic mRNA sequences of the mature protein coding region of α-Butx in gland cells. # indicates the DNA sequence of α-Butx in the genome. R1–R16 are the polymorphic sequences of α-Butx cDNAs in the cells. The genomic DNA and polymorphic mRNA sequences of α-Butx have been deposited in the EMBL/GenBank database under accession nos AF056401–AF056417.

α-ButxR^{36} mRNAs were frequently invariant, suggesting that the nt 198 and 201 variations might be related to the editing of G to A at nt 107 of α-ButxK^{36} mRNA. However, all the α-ButxR^{36} transcripts were accompanied by a silent mutation at nt 129 (GG<u>A</u>→GG<u>G</u>) encoding Gly^{43}. The isoelectric points (pI) of α-Butx and isoforms translated from the polymorphic mRNA sequences were predicted by a DNA and protein sequence analysis software system (DNASIS Pro™; Hitachi, San Bruno, CA). The pI values were 7.93 for R_1–R_{14} and 7.38 for R_{15} and R_{16} of α-Butx mRNA. To verify the existence of isoforms, 2-dimensional PAGE followed by immunoblotting was applied for the identification of isoforms of α-Butx. As shown in Figure 5, although the native α-Butx demonstrated a single band on SDS–PAGE and reverse-phase HPLC, the blot displayed two spots after incubation with α-Butx antibody. The pI values of the spots were estimated to be essentially identical to those predicted for α-Butx and isoforms.

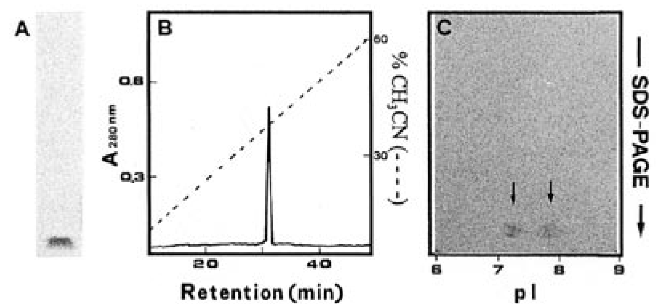

Figure 5. Determination of the presence of isoforms in HPLC-purified α-Butx by 2-dimensional PAGE followed with western blotting. (**A**) SDS–PAGE analysis of the purified α-Butx (10 µg). Molecular weight of the single protein band was 8 kDa; (**B**) 100 µg of α-Butx was loaded onto a C-18 reverse-phase column and eluted with acetonitrile in reverse-phase HPLC; (**C**) 50 µg of α-Butx was electrophoresed on an IEF gel (pI 3–10). The gel strip was utilized for SDS–PAGE according to procedures described in Materials and Methods. After transfer onto a nitrocellulose membrane, the α-Butx was detected by polyclonal α-Butx antibody according to the standard protocol of western blotting.

Genomic DNA sequence of the protein coding region of α-Butx

Because the gene structure of α-Butx is still unclear, the genomic DNA was extracted from the liver of *B.multicinctus*. To determine the protein coding sequence of α-Butx in genomic DNA, the DNA was amplified by PCR using *Pfu* DNA polymerase and the BXA and BXB primers. As shown in Figure 6, the size of the PCR product generated with the genomic DNA template was the same as that from the cellular RNA in gland cells using the same specific primers. The bands were further verified by Southern hybridization using [^{32}P]α-Butx cDNA as probe. It was found that the protein coding region of α-Butx contains no intron in the gene structure. The PCR products were then cloned into PCR-Script vectors and more than 50 positive clones containing the protein coding region of α-Butx genomic DNA were sequenced. All analyses resulted in only one sequence which was completely identical with that of R_1 cDNA. To investigate the copy number of α-Butx DNA in the genome, the genomic DNA was digested with *Eco*RI, *Sau*3A and the combination of both restriction enzymes. *Sau*3A is frequently

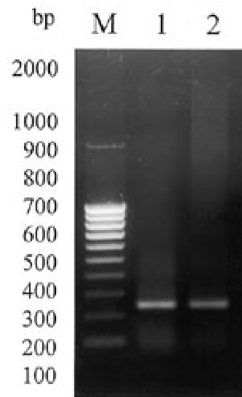

Figure 6. PCR and RT–PCR of the mature protein coding region of α-Butx. The mature protein coding regions of α-Butx in the genomic DNA and cellular RNA in gland cells were amplified with flanking primers by PCR and RT–PCR. The products were separated on a 2% agarose gel with ethidium bromide staining. The RNA was carefully prepared and treated with DNase to avoid DNA contamination. Lane M, DNA markers; lane 1, PCR of genomic DNA; lane 2, RT–PCR of cellular RNA.

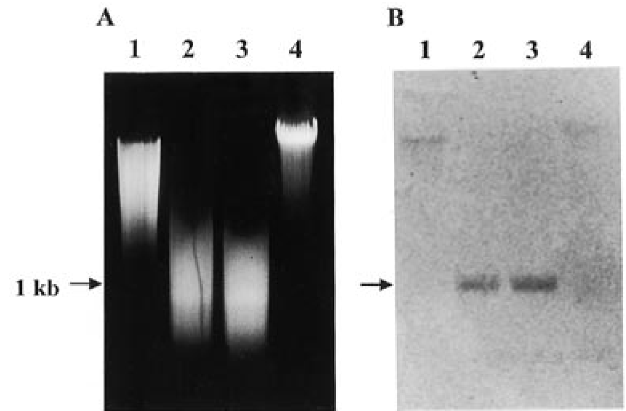

Figure 7. Analysis of the genomic DNA from *B.multicinctus* cells by digestion with restriction enzymes. The genomic DNA (60 µg) was digested with *Eco*RI, *Sau*3A and the combination of both enzymes. (**A**) The digests were electrophoresed on a 1% agarose gel with ethidium bromide staining. (**B**) Southern hybridization of the digested DNA with the mature protein coding region of α-Butx cDNA probe labeled with [^{32}P]dCTP. Lane 1, *Eco*RI; lane 2, *Sau*3A; lane 3, *Eco*RI and *Sau*3A; lane 4, control. The arrows indicate the sizes of the products.

used in the preparation of genomic libraries of cells. The complete cDNA sequence of α-Butx contains no cleavage sites for *Sau*3A or *Eco*RI. As shown in Figure 7, all the digests showed a single band on Southern hybridization. The DNA fragments generated by *Eco*RI digestion were larger than those due to *Sau*3A. However, the sizes of positive DNA bands from *Sau*3A and the combination of *Sau*3A and *Eco*RI were ~1 kb, suggesting that α-Butx DNA is a single copy in the genome. Thus, the R_1 mRNA is directly transcribed from the genomic template of α-Butx, while the other polymorphic mRNA sequences may have resulted from post-transcriptional RNA editing. These results indicate that post-transcriptional editing of α-Butx may occur in the venom gland cells. Therefore, the authentic α-Butx is in fact translated from the edited mRNA.

DISCUSSION

This is the first paper describing the complete mRNA sequence and the genetic information about a snake long neurotoxin. Originally, the α-Butx mRNA sequence was determined by the Cap-finder cDNA synthesis and Marathon cDNA amplification (Clontech, Palo Alto, CA) methods. Both the cDNA sequences analyzed by the two methods above were essentially identical. However, the Cap-finder PCR cDNA synthesis generated a complete sequence at the 5′-end of α-Butx cDNA. Comparing the nucleotide sequence of α-Butx with known toxins in the GenBank database, the 3′- and 5′-UTRs and signal sequences were almost identical to those from various genuses of Elapidae and Hydrophiidae snakes, suggesting that these toxins might share a common evolutionary origin. The sequences of signal peptides among the snake toxins appeared to be conserved, implying that they might be secreted by a common pathway from gland cells.

On the other hand, an RNA polymorphism has been verified in the mRNA of α-Butx. The 5′- and 3′-flanking sequences of the mRNAs of α-Butx and its isoforms were highly conserved, but the nucleotide sequence of the mature protein coding region of α-Butx varied. A similar result was observed with the cDNA sequences of phospholipase A_2 isolated from Crotalidae and Viperidae snakes, in which the nucleotide sequences flanking the signal peptide and stop codon were shown to be highly conserved (17,18). In addition, the cDNAs encoding short neurotoxins in the venom glands of the sea snakes *Aipysurus laevis* and *Latticauda semifasciata* have a common organization and present high similarity in nucleotide sequences, not only in the coding region but also the 3′- and 5′-UTRs (19). In this experiment, 10 distinct amino acid sequences of α-Butx could be deduced from the 16 polymorphic RNA sequences. The sequence with the substitution of Arg^{36} by Lys^{36} (R_2) was the major isoform of α-Butx mRNA. This substitution by Lys^{36} was the only variation among the conserved residues in α-Butx, but the positive charge of Lys^{36} might be able to maintain the ionic bond, like Arg^{36} with Asp^{30}. All the other hot-spot variations did not alter codons which encode the conserved residues in α-Butx. Since long and short neurotoxins may be from the same origin, enormous variations in the protein coding region during evolution may be responsible for the diverse sequences of α-neurotoxins. This finding may explain why the same conserved residues and a common secondary structure are found in all long and short neurotoxins derived from various species of snakes. In addition, all the conserved residues were gathered around the cluster of three loops and the tip of loop II. This illustration was consistent with the fact that the loop II structure of α-neurotoxin appeared to be an important binding domain for nAchR (20,21).

The process of post-transcriptional modification of RNA to alter the coding specificity of the sequence is termed RNA editing. The process may be caused by a variety of mechanisms in different organisms (22) and appears to occur independently of transcription, splicing and translation (23–25). Since the genome of *B.multicinctus* might contain a single copy of α-Butx DNA, the polymorphic mRNA sequences of α-Butx might result from post-transcriptional editing. Meanwhile, almost all the α-neurotoxins from Elapidae and Hydrophiidae snakes possess more than one homologous isoform and their sequences have been verified

(3). Erabutoxins, the short neurotoxins from *L.semifasciata*, are composed of a, b and c isoforms, but only the DNA sequence of erabutoxin c was identified from the genome (26). We are interested in understanding why snakes make use of such extensive RNA editing. The most reasonable explanation seems to be that snakes contain a primitive genetic system. It has been shown that RNA editing is regulated to produce different mRNAs under different conditions, so it can be viewed as a primitive way to change the expression of genes (27).

The mechanism of RNA editing in snake cells has not been identified. Deamination of cytosine and adenine to form uracil and inosine are performed by the APOBEC-1 and DRADA deaminases in mammalian cells. Inosine pairs preferentially with cytosine and it interacts with ribosomes in the same way as guanine. In both reverse transcription and translation, edited mRNAs that have undergone an A→I transition appear to have exchanged an A for a G. To date, four classes of mRNA editing have been observed in cells, i.e. C→U, A→G, U→C and U→A (28). As shown in Figure 4, most the edited nucleotides in α-Butx polymorphic sequences are A→G and C→U. Thus, deamination of C and A may be a common process in RNA editing of *B.multicinctus* cells.

α-ButxK^{36} (R_1–R_6) and α-ButxR^{36} (R_7–R_{16}) comprised six and 10 out of 16 polymorphic mRNAs. It appears that at least two major α-Butx proteins with either Arg or Lys at position 36 in the sequence are present in the venom. To verify the existence of α-Butx isoforms, the α-Butx was separated by 2-dimensional PAGE followed by western blotting. Two positive spots were identified and the spot locations in the blot matched the pI predictions for α-Butx and isoforms. In fact, one of the α-Butx isoforms (R_{14}) with a replacement of Ala^{31} by Val^{31} has been isolated from purified α-Butx by Mono-S cationic chromatography (11). These genetic characteristics of α-Butx may contribute to our understanding of the mechanism of gene regulation for the translation of α-neurotoxins in venom glands and provide an approach to the further verification of the α-Butx sequences.

ACKNOWLEDGEMENTS

The authors are grateful to Dr J. C. Bartimus for reading this manuscript. This work was supported by grants from the National Science Council of Taiwan (NSC 86-2314-B037-011 and NSC 87-2314-B037-063).

REFERENCES

1. Chang,C.C. (1979) In Lee,C.Y. (ed.), *Handbook of Experimental Pharmacology*: *Snake Venoms*. Springer-Verlag, Berlin, Germany, Vol. 52, pp. 309–376.
2. Changeux,J.-P. (1990) *Trends Pharmacol. Sci*., **11**, 485–492.
3. Endo,T. and Tamiya,N. (1991) In Harvey,A.L. (ed.), *Snake Toxins*. Pergamon Press, New York, NY, pp. 165–222.
4. Chiappinelli,V.A. (1993) In Harvey,A.L. (ed.), *Nature and Synthetic Neurotoxins*. Academic Press, New York, NY, pp. 65–128.
5. Pillet,L., Tremeau,O., Ducancel,F., Drevet,P., Zinn-Justin,S., Pinkasfeld,S., Boulain,J.-C. and Menez,A. (1993) *J. Biol. Chem*., **268**, 909–916.
6. Tremeau,O., Lemaire,C., Drevet,P., Pinkasfeld,S., Ducancel,F., Boulain,J.-C. and Menez,A. (1995) *J. Biol. Chem*., **270**, 9362–9369.
7. Balass,M., Katchalski-Katzir,E. and Fuchs,S. (1997) *Proc. Natl Acad. Sci. USA*, **94**, 6054–6058.
8. Changeux,J.-P., Devillers-Thiery,A., Galzi,J.-L. and Bertrand,D. (1992) *Trends Pharmacol. Sci*., **13**, 299–301.
9. Love,R.A. and Stroud,R.M. (1986) *Protein Engng*., **1**, 37–46.
10. Basus,V.J., Billeter,M., Love,R.A., Stroud,R.M. and Kuntz,I.D. (1988) *Biochemistry*, **27**, 22763–22771.
11. Kosen,P.A., Janet,F.-M., McCarthy,M.P. and Basus,V.J. (1988) *Biochemistry*, **27**, 2775–2781.
12. Scherf,T., Balass,M., Fuchs,S., Katchalski-Katzir,E. and Anglister,J. (1997) *Proc. Natl Acad. Sci. USA*, **94**, 6059–6064.
13. Tsernoglou,D., Petsko,G.A. and Hudson,R.A. (1978) *Mol. Pharmacol*., **14**, 710–716.
14. Rosenthal,J.A., Hsu,S.H., Schneider,D., Gentile,L.N., Messier,N.J., Vaslet,C.A. and Hawrot,E. (1994) *J. Biol. Chem*., **269**, 11178–11185.
15. Kuo,K.W., Chen,Y.C. and Chang,C.C. (1995) *Biochem. Biophys. Res. Commun*., **216**, 1088–1094.
16. Furuichi,Y. and Miura,K. (1975) *Nature*, **253**, 374–375.
17. Nakashima,K., Nobuhisa,I., Deshimaru,M., Ogawa,T., Shimohigashi,Y., Fukumaki,Y., Hattori,S. and Ohno,M. (1995) *Proc. Natl Acad. Sci. USA*, **92**, 5605–5609.
18. Wang,Y.M., Wang,J.H. and Tsai,I.H. (1996) *Toxicon*, **34**, 1191–1196.
19. Ducancel,F., Guignery-Frelat,G., Boulain,J.C. and Menez,A. (1990) *Toxicon*, **28**, 119–123.
20. Kuo,K.W., Chang,L.S. and Chang,C.C. (1995) *J. Biochem*., **117**, 438–442.
21. Peng,S.S., Kumar,T.K., Jayaraman,G., Chang,C.C. and Yu,C. (1997) *J. Biol. Chem*., **272**, 7817–7832.
22. Cattaneo,R. (1991) *Annu. Rev. Genet*., **25**, 71–88.
23. Gualberto,J.M., Bonnard,G., Lamattina,L. and Grienenberger,J.M. (1991) *Plant Cell*, **3**, 1109–1120.
24. Sutton,C.A., Conklin,P.L., Pruitt,K.D. and Hanson,M.R. (1991) *Mol. Cell. Biol*., **11**, 4274–4277.
25. Zeltz,P., Hess,W.R., Neckermann,K., Borner,T. and Kossel,H. (1993) *EMBO J*., **12**, 4291–4296.
26. Shields,D.C. and Wolfe,K.H. (1997) *Mol. Biol. Evol*., **14**, 344–349.
27. Alberts,B., Bray,D., Lewis,J., Raff,M., Robert,K. and Watson,J.D. (1994) *Molecular Biology of the Cell*, 3rd Edn. Garland Publishing, New York, NY.
28. Ashkenas,J. (1997) *Am. J. Hum. Genet*., **60**, 278–283.