# The Nucleotide Sequence of the Translated and Untranslated Regions of a cDNA for Myotoxinafrom the Venom of Prairie Rattlesnake (Crotalus viridis viridis)

**ARTICLE** *in* BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS · FEBRUARY 1997

Impact Factor: 2.3 · DOI: 10.1006/bbrc.1996.6021 · Source: PubMed



**3 AUTHORS**, INCLUDING:

Anthony T. Tu
Colorado State University
**236** PUBLICATIONS **4,218** CITATIONS

SEE PROFILE

BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS **230,** 607–610 (1997)
ARTICLE NO. RC966021

# The Nucleotide Sequence of the Translated and Untranslated Regions of a cDNA for Myotoxin *a* from the Venom of Prairie Rattlesnake (*Crotalus viridis viridis*)

Jeffrey W. Norris, Rachel M. Fry, and Anthony T. Tu[1]
*Department of Biochemistry and Molecular Biology, Colorado State University, Fort Collins, Colorado 80523*

Received December 10, 1996

**Myotoxin *a*, isolated from *C. viridis viridis* venom, is a myonecrotizing agent present in many snake venoms. A cDNA library was prepared from mRNA obtained from the venom glands of a *C. viridis viridis*. The complete base sequence of a cDNA corresponding to an mRNA encoding myotoxin *a* was determined. The 5′ untranslated region has 15 nucleotides, while the 3′ untranslated region has 109 nucleotides. The translated portion of the myotoxin *a* cDNA encodes a start methionine, a signal peptide, the myotoxin *a* peptide sequence, and an additional lysine residue. It is likely that myotoxin *a* is secreted as the cDNA encodes a signal peptide immediately 5′ to the myotoxin *a* peptide code.** 

Myotoxin *a* is a major component (20% of the dry weight) of the venom from *Crotalus viridis viridis*, and causes rapid degradation of muscle tissue (1, 2). Thus, myotoxin *a* contributes significantly to the local tissue damage caused by rattlesnake bites. Further, a number of structurally and functionally homologous peptides exist in other members of the genus *Crotalus* (2-6). Together, these novel toxins bear little sequence homology to postsynaptic neurotoxins and cardio-toxins (6). Myotoxin *a* inhibits $Ca^{2+}$ uptake into the sarcoplasmic reticulum by binding the $Ca^{2+}$ ATPase (7, 8). Thus, myotoxin *a* may cause its pathology by disrupting intracellular calcium flow.

Despite the important biological function of myotoxin *a*, it is unclear whether or not myotoxin *a* is released as the product of posttranslational process. This problem was addressed by determining the complete nucleotide sequence of cDNA for myotoxin *a*.

[1] To whom correspondence should be addressed. Fax: (970) 491-6313.

## MATERIALS AND METHODS

*Venom glands.* A mature, individual *C. viridis viridis* was collected near Ault, Colorado. Its venom glands were removed four days after venom extraction, as the highest mRNA concentration in the venom glands of the snakes *Echis carinatus* and *Vipera palestinae* occurs after this period of time (9, 10).

*Construction of cDNA library.* 368 $\mu$g of whole cell RNA was isolated from the venom glands (RNA Isolation Kit, Stratagene, La Jolla, CA), from which 7.74 $\mu$g of poly $(A)^+$ RNA was purified (Poly $(A)^+$ Quik mRNA Purification Kit, Stratagene). These quantities were determined by measuring the optical density of the RNA solutions at 260 nm.

The cDNA library was constructed from 5 $\mu$g of poly $(A)^+$ RNA (ZAP-cDNA) synthesis kit, Stratagene). The resulting lambda library was plated on *Escherichia coli* SURE cells, thereby amplifying $1.0 \times 10^5$ independent colonies once. Conversion of the bacteriophage library to a pBluescript plasmid library was accomplished by *in vivo* excision. Plasmids isolated from a liquid culture of cells from the excision were used in the following Polymerase Chain Reactions (PCR).

*PCR - degenerate primers.* Two pools of degenerate primers (Primer Pool N and Primer Pool C, Macromolecular Resources, Ft. Collins, CO) (Figure 1) were used to amplify a DNA fragment from the cDNA library unique to myotoxin *a* (Figure 2A). The one hundred micro-liter reactions contained the following reagents: 50 pmol of each primer pool, 2 ng of plasmid DNA, 3 units of AmpliTaq DNA polymerase (Perkin Elmer, Foster City, CA), 10 $\mu$L 10× PCR Buffer II (Perkin Elmer), 3.0 mM $MgCl_2$, and 100 $\mu$M dTP's. Thirty rounds of amplification were carried out, each of which had the following parameters: denaturation at 95°C for 1 min (5 min on the first round), primer annealing at 43°C for 1 min, and extension of product at 72°C for 1 min.

The resulting 91 bp PCR product (1.2 ng) was cloned into the pCR II vector (Invitrogen, San Diego, CA), and transformed into *E. coli* XL1 Blue MRF′ cells. Sequencing of the insert was carried out by the amplicycle cycle sequencing method (Perkin Elmer), and by the Amplicycle cycle sequencing method (Perkin Elmer), and directed from the M13 Forward (−20) and M13 Reverse primer sites in the vector. Twenty-five femtomols of target DNA were subjected to twenty-five rounds of amplification (Coy thermocycler): denaturation at 95°C for 1 min (5 min on the first round), primer annealing at 48°C for 1 min, and extension of product at 72°C for 1 min.

*Determination of the full myotoxin a cDNA sequence.* The sequence of the 91 bp product provided information from which nondegenerate primers could be designed. Primer C2 (Figure 1) (Macromo-



0006-291X/97 $25.00

I. Degenerate Primer Pools

Primer Pool N

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Amino Acid Sequence | H | C | F | P | K | E | K |
| Oligo Sequence | CAT | TGT | TTT | CCA | AAA | GAA | A |
| | C | C | C | C | G | G | |
| | | | | G | | | |
| | | | | T | | | |

Primer Pool C

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Amino Acid Sequence | G | K | K | C | C | W | K |
| Oligo Sequence | CCT | TTT | TTA | CAA | CAT | TCC | A |
| | C | C | G | G | C | | |

II. Nondegenerate Primers

Primer C2

TTTCCATCGACAGTCCA

Primer N2

AAATATGTATTCCTCCATC

**FIG. 1.** Oligonucleotide primers.

lecular Resources) was coupled with the M13 Reverse primer of pBluescript to direct the amplification of the 5′ untranslated region of the myotoxin *a* cDNA (Figure 2B). The PCR consisted of those reagents described in the preceding section with the following modifications: 30 pmol of each primer and 2.5 mM $MgCl_2$. Twenty-five rounds of amplification were conducted as described in that section, except the annealing temperature was 48°C for 30 sec. The product was purified by gel electrophoresis and sequenced in both directions (Primer C2 and Primer M13 Reverse), as described in the preceding section.

The sequence of the 3′ untranslated region of myotoxin *a* was determined by coupling Primer N2 (Macromolecular Resources, Figure 1) to the M13 Forward (−20) primer of pBluescript (Figure 2C). This procedure was analogous to that used to amplify the 5′ untranslated region. Alignment of these sequences produces the full sequence of the myotoxin *a* cDNA.

*Sequence analysis.* The nucleotide sequence of the myotoxin *a* cDNA was compared to those in Genbank, European Molecular Biology Laboratories (EMBL), DNA Database of Japan (DDBJ), and Brookhaven Protein Data Bank (PDB) with the BLAST email server (blast@ncbi.nlm.nil.gov).

## RESULTS AND DISCUSSION

The coding regions of mRNA's encode a protein and are bracketted by untranslated regions. The mRNA encoding myotoxin *a* (Figure 3) has 5′ and 3′ untranslated regions which are 15 and 109 nucleotides, respectively. Nucleotides 16 - 81 encode the start methionine and a 22 amino acid signal peptide characteristic of those which target peptides to the endoplasmic reticulum. Nucleotides 82 - 210 encode the myotoxin *a* peptide sequence, while nucleotides 211 - 213 encode a lysine residue which is not present in myotoxin *a*. Apparently, this lysine is removed by a posttranslational process. The lysine is followed by a stop codon at nucleotides 214 - 216. The mRNA's of several snake venom proteases encode propeptides which presumably inactivate

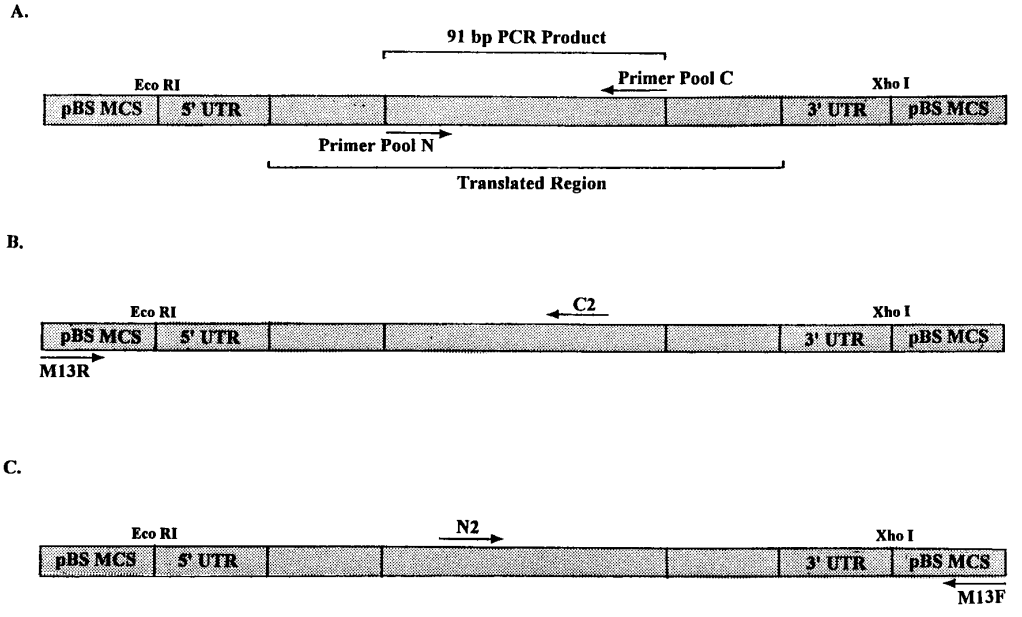


**FIG. 2.** Cloning strategy. This is a diagram of the pBluescript multicloning site (pBS MCS) with the myotoxin *a* cDNA cloned between the Eco RI and Xho I restriction sites. (A) Primer Pool C and Primer Pool N were used to amplify a 91 bp fragment of the myotoxin *a* translated region. (B) Primer C2 and Primer M13 Reverse (M13R) were used to amplify a portion of the myotoxin *a* cDNA which includes the 5′ UTR. (C) Primer N2 and Primer M13 Forward (−20) (M13F) were used to amplify a portion of the myotoxin *a* cDNA which includes the 3′ UTR.

these proteins prior to secretion (11); however, the myotoxin *a* mRNA doesn't encode such a sequence. Comparison of the myotoxin *a* nucleotide sequence with those in the GenBank, EMBL, DDBJ, and PDB databases indicated it has no significant similarity to reported sequences.

Aird et al. have purified proteins, from the venom of an individual *C. viridis viridis*, with slight variations of the myotoxin *a* peptide sequence (12, 13). Our data do not provide information about the heterogeneity of the myotoxin *a* mRNA population. Direct sequencing by PCR is an averaging process which cannot detect minor sequence variations of the target DNA (14). Instead, our data represent the sequence of an mRNA for the major myotoxin *a* component.

Although the secondary structures of proteins are commonly predicted, those of nucleic acids have not been extensively explored. Ogawa (15) used method of

```
                              10              20              30
PCR 1       : CGGCAGGAGCTCAGC ATG AAG ATC CTT TAT CTG CTG TTC
PCR 2       :
AMINO ACID:                       Met  Lys  Ile   Leu Tyr  Leu  Leu  Phe

                 40              50              60              70
PCR 1       : GCA TTT CTT TTC CTT GCA TTC CTG TCT GAA CCA GGG AAT
PCR 2       :
AMINO ACID: Ala  Phe  Leu  Phe Leu Ala  Phe  Leu Ser  Glu  Pro  Gly  Asn

                  80              90              100             110
PCR 1       : GCC TAT AAA CAG TGT CAG AAG AAA GGA GGA CAC TGC
PCR 2       :
AMINO ACID: Ala  Tyr  Lys  Gln  Cys  His  Lys  Lys  Gly  Gly  His  Cys

                        120             130             140             150
PCR 1       : TTT CCC AAG GAG AAA ATA TGT ATT CCT CCA TCT TCT GAC
PCR 2       :                       AAA ATA TGT ATT CCT CCA TCT TCT GAC
AMINO ACID: Phe  Pro  Lys  Glu  Lys  Ile  Cys Ile  Pro  Pro  Ser  Ser  Asp

                           160             170             180
PCR 1       : CTT GGG AAG ATG GAC TGT CGA TGG AAA
PCR 2       : CTT GGG AAG ATG GAC TGT CGA TGG AAA TGG AAA TGC
AMINO ACID: Phe Gly  Lys  Met Asp  Cys  Arg  Trp  Lys  Trp  Lys  Cys

                  190             200             210             220
PCR 1       :
PCR 2       : TGT AAA AAG GGA AGT GGA AAA TAA TGCCATCTCCATCTA
AMINO ACID: Cys Lys  Lys   Gly  Ser  Gly  Lys  Stop

                    230             240             250             260
PCR 1       :
PCR 2       : GGACCATGGATATCTTCAAGATATGGCCAAGGACCTGAGAGT
AMINO ACID:

                  271             280             290             300             310
PCR 1       :
PCR 2       : GCCGGTGCTATTGCCTTTATCTTTCTTTATCTAAATAAAATTG
AMINO ACID:

                        320
PCR 1       :
PCR 2       : CTACCTATC poly(A)
AMINO ACID:
```

**FIG. 3.** Sequence of Overlapping PCR Products. PCR 1 is the sequence of the product generated with Primer C2 and Primer M13 Reverse (Lane 7 of Figure 2). PCR 2 is the sequence of the product generated with Primer N2 and Primer M13 Forward (−20) (Lane 5 of Figure 2). The pBluescript, Primer N2, and Primer C2 sequences have been removed from PCR 1 and PCR 2.

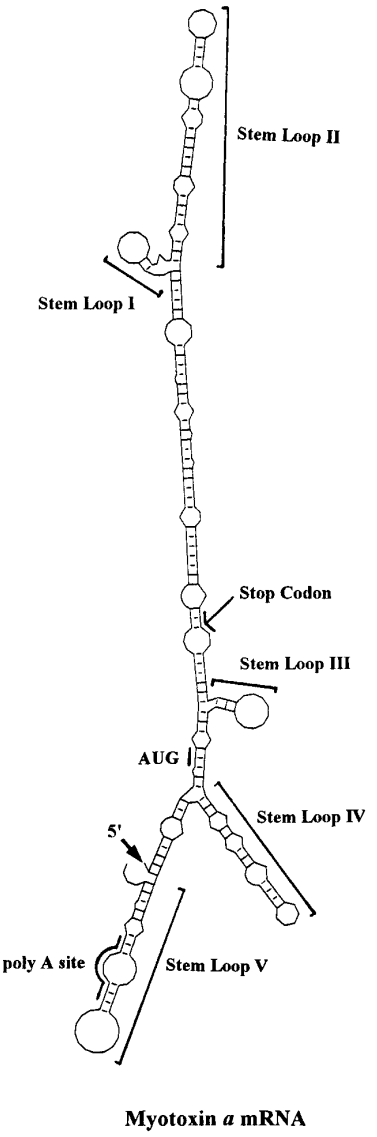


**FIG. 4.** Predicted Secondary Structures of the myotoxin *a* mRNA.

Zukker and Steigler (16) to determine the secondary structure of mRNA's encoding phospholipase $A_2$ toxins from *Trimeresurus flavovoridis* (Habu snake). Further, they suggested that stem loops in the untranslated regions of these toxins may alter the translation rate or stability of snake toxin mRNA's. We also used the method of Zukker and Steigler to determine a secondary structure for the myotoxin *a* mRNA (Figure 4). There are two major differences between the myotoxin *a* mRNA secondary structure and the phospholipase $A_2$ mRNA secondary structures. First, the coding region of the phospholipase $A_2$ encoding mRNA's does not participate in secondary structure formation. However, the secondary structure predicted for the myotoxin *a* mRNA is very different as the coding region does par-

ticipate as part of the structure. Second, there are no stem loop structures shared by the mRNA's encoding myotoxin *a* and the phospholipase $A_2$ toxins.

## ACKNOWLEDGMENTS

The work is supported by NIH Grant 5R37GM15591. The authors thank Brenda Baker, Lynn Taylor, and Nancy Tress for assistance with the experiments.

## REFERENCES

1. Cameron, D. L., and Tu, A. T. (1977) *Biochemistry* **16,** 2546–2553.
2. Ownby, C. L., Cameron, D. L., and Tu, A. T. (1976) *Am. J. Pathol.* **85,** 149–166.
3. Laure, C. J. (1975) *Hoppe-Seyler's Z. Physiol. Chem.* **356,** 213–215.
4. Maeda, N., Tamiya, N., Pattabhiraman, T. R., and Russell, F. E. (1978) *Toxicon* **16,** 431–441.
5. Engle, C. M., Becker, R. R., Bailey, T., and Bieber, A. L. (1983) *J. Toxicol. - Tox. Rev.* **2,** 267–283.
6. Fox, J. W., Elzinga, M., and Tu, A. T. (1979) *Biochemistry* **18,** 678–684.
7. Volpe, P., Damiani, E., Maurer, A., and Tu, A. T. (1986) *Arch. Biochem. Biophys.* **246,** 90–97.
8. Utaisincharoen, P., Baker, B., and Tu, A. T. (1991) *Biochemistry* **30,** 8211–8216.
9. Paine, M., Desmond, H., Theakston, R., and Crampton, J. (1992) *Toxicon* **30,** 379–386.
10. Rotenberg, D., Bamberger, E., and Kochva, E. (1971) *Biochem. J.* **121,** 609–612.
11. Siigur, E., Aaspõllo, A., Tu, A. T., and Siigur, J. (1996) *Biochem. Biophys. Res. Commun.* **224,** 229–236.
12. Griffin, P. R., and Aird, S. D. (1990) *FEBS Lett.* **274,** 43–47.
13. Aird, S., Kruggel, W., and Kaiser, I. (1991) *Toxicon* **29,** 265–268.
14. Gyllensten, U. B. (1989) *Biotechniques* **7,** 700–708.
15. Ogawa, T., Oda, N., Nakashima, K., Sasaki, H., Hattori, M., Sakaki, Y., Kihara, H., and Ohno, M. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 8557–8561.
16. Zukker, M., and Steigler, P. (1981) *Nucleic Acids Res.* **9,** 133–148.