ELSEVIER

Toxicon 42 (2003) 747–752

TOXICON

www.elsevier.com/locate/toxicon

Structure and chromosomal localization of the gene for crotamine, a toxin from the South American rattlesnake, *Crotalus durissus terrificus*

Gandhi Rádis-Baptista[a], Tai Kubo[b], Nancy Oguiura[c,*], Marta Svartman[d], Therezinha M.B. Almeida[e], Radenka F. Batistic[c], Eduardo B. Oliveira[f], Ângela M. Vianna-Morgante[d], Tetsuo Yamane[a]

[a]*Laboratório de Toxinologia Molecular, Instituto Butantan, Av. Vital Brazil, 1500, São Paulo 05503-900, Brazil*
[b]*Molecular Neurophisiology Laboratory, National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1 Higashi, Tsukuba Central 6, Tsukuba 305-8566, Japan*
[c]*Laboratório de Herpetologia, Instituto Butantan, Av. Vital Brazil, 1500, São Paulo 05503-900, Brazil*
[d]*Departamento de Biologia, Instituto de Biociências, Universidade de São Paulo, CP 11461, São Paulo 05422-970, Brazil*
[e]*Laboratório de Genética, Instituto Butantan, Av. Vital Brazil, 1500, São Paulo 05503-900, Brazil*
[f]*Departamento de Bioquímica, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto 14049-900, Brazil*

Received 8 August 2003; revised 26 August 2003; accepted 3 October 2003

Abstract

Crotamine is a 42 amino acid-long basic polypeptide, one of the major components of the South American rattlesnake, *Crotalus durissus terrificus*, venom. The mRNA has about 340 nucleotides and codifies a pre-crotamine, including the signal peptide, the mature crotamine, and a final lysine. In this report, we describe the crotamine gene with 1.8 kb organized into three exons separated by a long phase-1 (900 bp) and a short phase-2 (140 bp) introns. Exon 1 includes the 5′-untranslated region and codifies the first 19 amino acids of the signal peptide. Exon 2 codifies 42 amino acids, three belonging to the signal peptide and 39 to the mature crotamine. Exon 3 codifies the last three amino acids of the mature toxin and the terminal lysine. The crotamine gene was mapped by in situ hybridization to the end of the long arm of chromosome 2, the intensity of signals differing between the two homologues. This may reflect a difference in gene copy numbers between chromosomes, a possible explanation for the variable amounts of crotamine found in the venom.
© 2003 Elsevier Ltd. All rights reserved.

Keywords: Snake toxins; Gene mapping; Gene structure; Myotoxin

1. Introduction

Crotalus durissus is a rattlesnake found from Southern Mexico to the pampas of Argentina (Campbell and Lamar, 1989). Vellard (1939) reported the dependence of the physiopathological effect of *C. durissus* venom on its geographical origin, reflecting the variability of the venom composition.

Crotamine was first isolated from the venom of the South American rattlesnake *C. durissus terrificus* (Gonçalves, 1956). It belongs to a group of closely related small basic polypeptide myotoxins (SBPM), commonly found in *Crotalus* venoms, and its contents vary according to the subspecies or the geographical location (Schenberg, 1959; Bober et al., 1988; Ownby, 1998), reaching an average of 17% (w/w) of the dry crude venom (Gonçalves and Arantes, 1956; Oguiura et al., 2000). The amino acid sequences of

* Corresponding author. Tel.: +55-11-3726-7222x2267; fax: +55-11-3726-1505.
E-mail address: nancyoguiura@butantan.gov.br (N. Oguiura).

0041-0101/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/j.toxicon.2003.10.019

SBPMs from various *Crotalus* species have been determined and they showed a high degree of similarity ranging from 83 to 98% (Samejima et al., 1991).

The crotamine cDNAs from venom glands of *C. durissus terrificus* was found to be 340–360 bp in length, encompassing an open reading frame of 198 nucleotides (nt), with 5′ and 3′ untranslated regions (UTR) of variable sizes (Rádis-Baptista et al., 1999), and showing a high degree of sequence similarity with the myotoxin a mRNA from *Crotalus viridis viridis* (Norris et al., 1997).

According to their structure, snake venom toxin genes can be grouped into: (a) genes with five exons and four introns, as in batroxobin, a thrombin-like venom enzyme from *Bothrops atrox* (Itoh et al., 1988); (b) genes with four exons and three introns as in phospholipases A_2 (PLA_2) from *Crotalus scutulatus scutulatus* (John et al., 1994), *Trimeresurus gramineus*, *Trimeresurus flavoridis* (Nakashima et al., 1995), *Trimeresurus okinavensis* (Nobuhisa et al., 1996) and *Naja sputatrix* (Jeyaseelam et al., 2000); (c) genes with three exons and two introns, as in erabutoxin c from *Laticauda semifasciata* (Fuse et al., 1990), cardiotoxin, cobrotoxin, cobrotoxin b from *Naja naja atra* (Chang et al., 1997a,b), cardiotoxin, α-neurotoxin from *N. n. sputatrix* (Lauchumanan et al., 1998; Afifiyan et al., 1999) and α-neurotoxin short-chain from *Pseudonaja textilis* (Gong et al., 2000). An unusual characteristic of these genes is the high conservation of intron sequences, in contrast to a high frequency of nucleotide substitutions observed in coding regions (accelerated evolution), as it has been clearly shown for PLA_2 genes (Nakashima et al., 1995).

In Viperidae, the family to which the *Crotalus* genus belongs, the karyotype includes 36 chromosomes: eight macrochromosome pairs and 10 microchromosome pairs. The fourth pair of macrochromosomes is the sex pair, which is heteromorphic in females, characterizing a ZZ/ZW sex determining system (Beçak, 1964). Chromosomal studies of snakes are relatively scarce and gene assignment has never been described.

Herein, we report the genomic sequence of the crotamine gene of *C. d. terrificus*, and its chromosomal localization.

2. Material and methods

2.1. Snake specimens

Three specimens of crotamine-positive *C. d. terrificus* (*Cdt* 9705, *Cdt* 9706 and *Cdt* 6724), captured in different locations in the State of São Paulo, Brazil, were obtained from Instituto Butantan.

2.2. Isolation and sequencing of the crotamine gene (Crt-p1)

DNA was isolated from liver according to Ausubel et al. (1995). Isolation of the crotamine gene from the specimen *Cdt* 9706 was performed by genome walking using the Universal Genome Walker Kit (Clontech, Palo Alto, CA). The five libraries obtained were amplified using ExLONGase enzyme Mix (Life Technologies, Grand Island, NY), with AP1 sense primer supplied with the kit, and the specific antisense primer GSP2 (5′-ACTGTCGATGGAGATGGAAAATGCTGTA-3′). The PCR products were re-amplified with nested primers, forward AP2 and reverse specific 3′crot163 (5′-CATCTCCATCGACAGTCCAT-3′). The specific primers GSP2 and 3′crot163 were designed based on the 3′end sequence of crotamine cDNA precursors (Rádis-Baptista et al., 1999). The PCR products were cloned into pCR2.1-TOPO vector (InVitroGen Life Technologies, San Diego). The gene was sequenced with ABI Prism Big Dye Terminator (Perkin Elmer, Foster City, CA) in an automated sequencer (ABI Prism 373 or 377, Perkin Elmer) using synthetic oligonucleotides designed from the crotamine cDNA precursor sequences.

The *Crt-p1* gene from the specimen *Cdt* 9705 was isolated from its genomic library constructed with Lambda-FIXII/*Xho* I partial Fill-in Kit (Stratagene, LaJolla, CA). Screening was performed by the PCR-based protocol of Israel (1995), with some modifications. The library was subdivided into 32 aliquots of 1000 plaque-forming units (pfus) per well in a microplate. After propagation, 1 μl of phage suspension from each well was used in an arbitrary PCR (AR-PCR) with the forward primer H010 (5'-AAGCAGTCTCAGCATGAAGATC-3'), which corresponds to the signal peptide sequence of the crotamine cDNA, and the reverse primer 3′crot163 yielding a PCR product of ~ 1.1 kbp. Phages from positive wells were diluted to 25 pfus per well, re-amplified and re-screened. Phages from positive wells were plated, and the isolated clone MR-20, containing a 15 kb genomic insert harboring the crotamine gene, was prepared as described by Ausubel et al. (1995). The excised *Crt-p1* was subcloned into pGEM 3 Zf (Promega, Madison, WI) and transprimers were randomly inserted into the clone pKT206 with the Genome Priming System GPS-1 (New England Biolabs, Beverly, MA). Sequencing was performed as described above, using the primers supplied in the GPS-1 kit.

2.3. Sequence analysis

The contigs were assembled using the Biocomputing Software Lasergene (DNAStar, Inc., Madison, Wiscosin). The intron/exon junctions of the *C. d. terrificus* crotamine gene were identified by comparing genomic DNA sequences with crotamine cDNA sequences in the GenBank, using the BLAST program. To find the theoretical binding sites of proteins that regulate transcription, the *Crt-p1* sequence of *Cdt* 9705 was analyzed using TFSEARCH 1.3, 1995 from Yutaka Akiyama (Kyoto University).

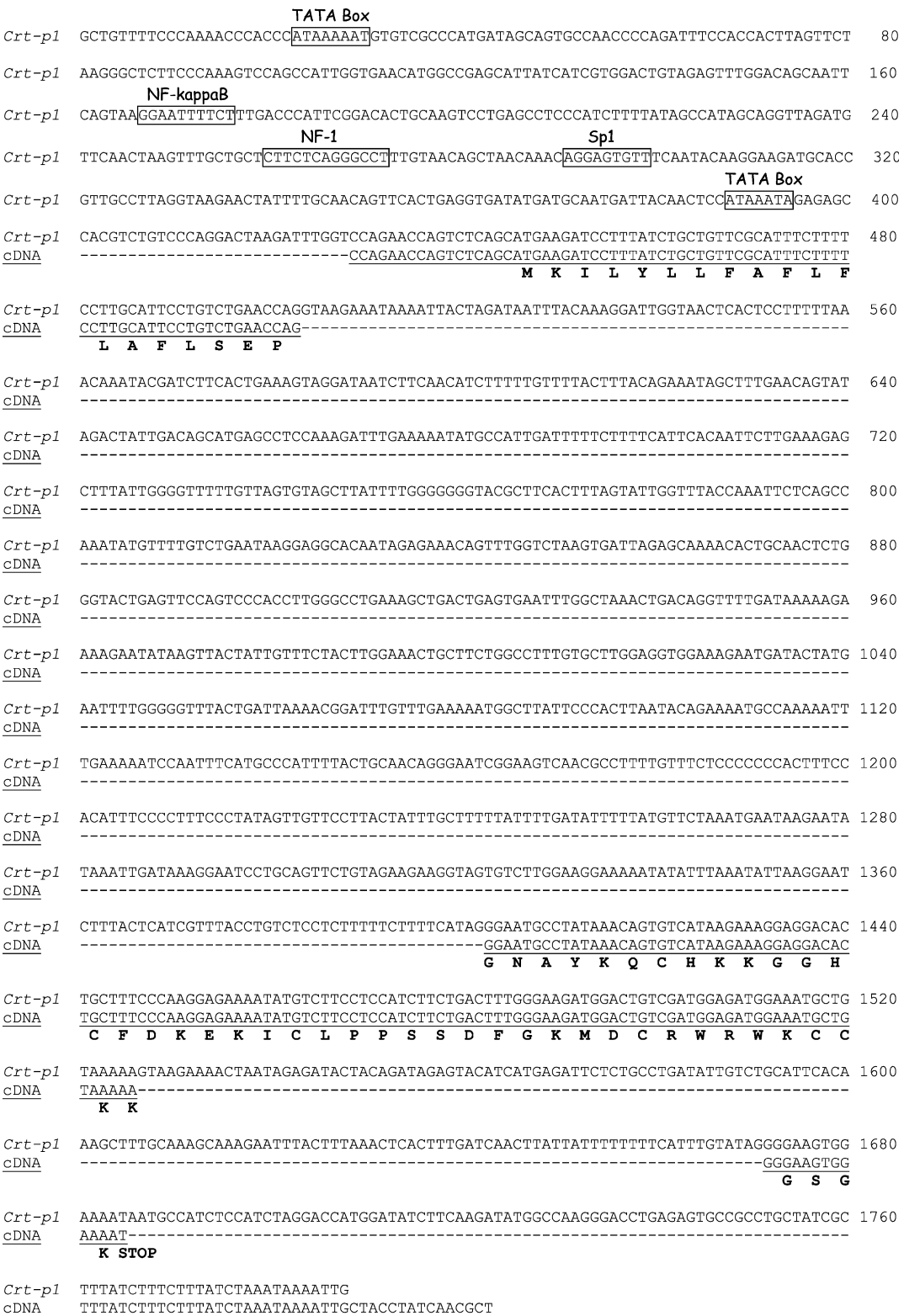

Fig. 1. Complete nucleotide sequence of the crotamine gene aligned to its cDNA (underlined sequence) and amino acid sequence (in bold). Promoter consensus sequences are shown inside boxes and the third line represents amino acid sequences.

2.4. Crotamine gene mapping

Chromosome preparations were obtained from rib bone marrow of the specimen *Cdt* 6724. The probe, MR20 phage DNA containing the *Crt-p1* from *Cdt* 9705 was labeled with biotin by nick translation (BioNick, InVitroGen Life Technologies, San Diego). In situ hybridization was performed as previously described by Viegas-Pequinot (1992) with minor modifications, using avidin conjugated to FITC (Vector, Burlingame, CA) for probe detection. Analyses and documentation were performed in a Carl Zeiss Axiophot epifluorescent microscope equipped with a CCD camera and the ISIS software (Metasystems, Altlussheim, Germany).

3. Results

The *Crt-p1* gene sequences from *Cdt* 9705 and *Cdt* 9706 differed just for some base changes in intron 1 and in the 3′-UTR (GenBank AF223946 and AF223947, respectively). The crotamine gene consists of 1785 nt and, by comparing it with the published crotamine cDNA nucleotide sequences (Rádis-Baptista et al., 1999), the exon/intron boundaries could be identified. It contains three exons, with coding regions of 58, 124 and 13 nt, totaling 195 nt, intervened by two introns of 898 and 145 nt. The 5′-splice donor and 3′-splice acceptor sequences conform to the GT-AG rule, gAGGT/AGGa (Senapathy et al., 1990). This gene structure is similar to that of cardiotoxin (Lauchumanan et al., 1998), a myotoxin presenting a different pharmacological activity and similar molecular properties, being both highly basic proteins (Ownby, 1998).

The first exon encompasses 58 nt coding for 19 amino acid residues out of 22 that characterize the signal peptide of the crotamine precursor; the second exon (124 nt) codes for the last three amino acids of the signal peptide and for the most part of the mature toxin (39 out of 42 amino acids); the third exon encodes for the remaining four amino acids residues of crotamine and includes the downstream flanking region containing the 3′-UTR of the crotamine cDNA. The upstream flanking region contains the 5′-UTR, two TATA boxes and theoretical binding sites for NF-κB, NF-1, SP1 (Fig. 1).

In situ hybridization allowed the mapping of the crotamine gene to the end of the long arm of chromosome 2 in metaphases of the specimen *Cdt* 6724. The intensity of the signal differed between the two homologues in all cells analyzed (Fig. 2).

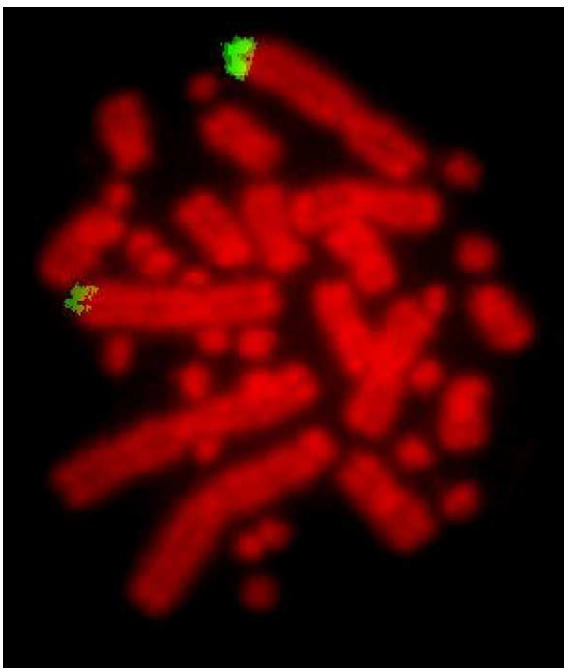

Fig. 2. Chromosomal localization of *Crt-p1* gene by fluorescence in situ hybridization: hybridization signal visible as green spots on the tip of the long arm of chromosome 2. Note the difference in signal intensities. Chromosomes were counterstained with propidium iodide.

4. Discussion

The structural organization of the crotamine gene herein described is very similar to that of the majority of toxin genes: the first exon contains most of the leader sequence followed by a relatively long intron; the second exon codes for most part of the mature toxin and is followed by the relatively short second intron; and the third exon encompasses the toxin C-terminal region and the 3′-UTR. One peculiarity of the *Crt-p1* gene is the phase of introns which differs from other toxins: the 3′-ends of exons 1 and 3 are interrupted after the first nucleotide of the last codon (phase 1-intron), and the 3′-end of exon 2 has its 3′-end codon interrupted after the second nucleotide (phase 2-intron). The exons of other toxins, like post-synaptic α-neurotoxin and cardiotoxin genes, are interrupted by introns of the same phase (phase 1). Although several amino acid sequences of SBPM have been described, their variability proved to be quite small, the two most diverging SBPM having 86% similarity suggesting that these toxins evolved rather recently and have not accumulated mutations related to the accelerated evolution observed in other snake toxins. When the crotamine and the myotoxin a cDNAs were compared, the coding sequences showed a higher similarity (100 and 98% for the signal peptide and mature myotoxin, respectively) than the UTR (60 and 80% for 5′ and 3′ regions). This characteristic differentiates these members of SBPM family from other toxins whose genes

have higher similarity in non-coding sequences (Norris et al., 1997; Rádis-Baptista et al., 1999).

Our previous study (Rádis-Baptista et al., 1999) had shown six different crotamine cDNAs, four coding for the crotamine-ile 19 (Santos et al., 1993) and two for crotamine (Laure, 1975), obtained from the venom gland of the specimen *Cdt* 9706 herein used. Therefore, it was quite surprising when we found only one gene coding for crotamine, although a search for others was made. The possibility remains that the gene although present in the gland was not detected. Alternatively, the C → A change leading to the leucine for isoleucine substitution at position 19 may occur as a post-transcriptional event (Maas and Rich, 2000).

The crotamine gene is the first gene to be mapped on snake chromosomes. The consistent difference in the intensity of the hybridization signals between homologues may well result from a variable number of copies of the gene on each chromosome. This in turn would be a possible explanation for the variable amounts of crotamine found in the venom of crotamine-positive *C. d. terrificus*.

Acknowledgements

We thank K.F. Grego and Nathália Fernandes Azevedo for technical assistance. DNA sequencing was performed in the Laboratório de Bioquímica e Biofísica, Instituto Butantan. This work was supported by FAPESP and carried out in compliance with the current guidelines governing genetic experimentation in Brazil.

References

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K. (Eds.), 1995. Current Protocols in Molecular Biology, Wiley, New York.

Afifiyan, F., Armugam, A., Tan, C.H., Gopalakrishnakone, P., Jeyaseelan, K., 1999. Postsynaptic α-neurotoxin gene of the spitting cobra, *Naja naja sputatrix*: structure, organization, and phylogenetic analysis. Genome Res. 9, 259–266.

Beçak, W., 1964. Karyotypes, Sex Chromosomes, and Chromosomal Evolution in Snakes, Venomous Animals and their Venoms, vol. 1. Academic Press, New York, pp. 53–95.

Bober, M.A., Glenn, J.L., Straight, R.C., Ownby, C.L., 1988. Detection of myotoxin a-like proteins in various snake venoms. Toxicon 26, 665–673.

Campbell, J.A., Lamar, W.W., 1989. The venomous reptiles of Latin America. Comstock Publishing Associates, New York, pp. 330–346.

Chang, L.-S., Chou, Y.-C., Lin, S.-R., Wu, B.-N., Lin, J., Hong, E., Sun, Y.-J., Hsiao, C.-D., 1997a. A novel neurotoxin, cobrotoxin b, from *Naja naja atra* (Taiwan Cobra) venom: purification, characterization, and gene organization. J. Biochem. (Tokyo) 122, 1252–1259.

Chang, L.-S., Lin, J., Chou, Y.-C., Hong, E., 1997b. Genomic structures of cardiotoxin 4 and cobrotoxin from *Naja naja atra* (Taiwan Cobra). Bioch. Biophys. Res. Commun. 239, 756–762.

Fuse, N., Tsuchiya, T., Nonomura, Y., Menez, A., Tamiya, T., 1990. Structure of the snake short-chain neurotoxin, erabutoxin c, precursor gene. Eur. J. Biochem. 193, 629–633.

Gonçalves, J.M., 1956. Purification and properties of crotamine. In: Bucley, E.E., Porges, N. (Eds.), Venoms, American Association for the Advancement of Science, Washington, DC, pp. 261–274.

Gonçalves, J.M., Arantes, E.G., 1956. Estudos sobre Venenos de Serpentes Brasileiras. III—Determinação Quantitativa de Crotamina no Veneno de Cascavel Brasileira. An. Acad. Bras. Cienc. 28, 369–371.

Gong, N.-L., Armugam, A., Jeyaseelan, K., 2000. Molecular cloning, characterization and evolution of the gene encoding a new group of short-chain α-neurotoxins in an Australian elapid, *Pseudonaja textilis*. FEBS Lett. 473, 303–310.

Israel, D.I., 1995. A PCR-based method for screening DNA libraries. In: Dieffenbach, C.W., Dveksler, G.S. (Eds.), PCR Primer—a Laboratory Manual, Cold Spring Harbor Laboratory Press, New York, pp. 453–461.

Itoh, N., Tanaka, N., Funakoshi, I., Kawasaki, T., Mihashi, S., Yamashina, I., 1988. Organization of the gene for batroxobin, a thrombin-like snake venom enzyme. J. Biol. Chem. 263, 7628–7631.

Jeyaseelam, K., Armugam, A., Donghui, M., Tan, N.-H., 2000. Structure and phylogeny of the venom Group I phospholipase A_2 gene. Mol. Biol. Evol. 17, 1010–1021.

John, T.T., Smith, L.A., Kaiser, I.I., 1994. Genomic sequences encoding the acidic and basic subunits of mojave toxin: unusually high sequence identity of non-coding regions. Gene 139, 229–234.

Lauchumanan, R., Armugam, A., Tan, C.H., Jeyaseelan, K., 1998. Structure and organization of the cardiotoxin genes in *Naja naja sputatrix*. FEBS Lett. 433, 119–124.

Laure, C.J., 1975. Die primarstruktur des crotamins. Hoppe Seylers Z. Physiol. Chem. 356, 213–215.

Maas, S., Rich, A., 2000. Changing genetic information through RNA editing. Bioessays 22, 790–802.

Nakashima, K.-I., Nobuhisa, I., Deshimaru, M., Nakai, M., Ogawa, T., Shimogashi, Y., Fukumaki, Y., Hattori, M., Sakaki, Y., Hattori, S., Ohno, M., 1995. Accelerated evolution in the protein-coding regions is universal in crotalinae snake venom gland phospholipase A_2 isozyme genes. Proc. Natl Acad. Sci. USA 92, 5605–5609.

Nobuhisa, I., Nakashima, K.-I., Deshimaru, M., Ogawa, T., Shimogashi, Y., Fukumaki, Y., Sakaki, Y., Hattori, S., Kihara, H., Ohno, M., 1996. Accelerated evolution of *Trimeresurus okinavensis* venom gland phospholipase A_2 isozyme-encoding genes. Gene 172, 267–272.

Norris, J.W., Fry, R.M., Tu, A.T., 1997. The nucleotide sequence of the translated and untranslated regions of a cDNA for myotoxin a from the venom of Prairie Rattlesnake (*Crotalus viridis viridis*). Biochem. Biophys. Res. Commun. 230, 607–610.

Oguiura, N., Camargo, M.E., da Silva, A.R.P., Horton, D.S.P.Q., 2000. Quantification of crotamine, a small basic myotoxin, in South American rattlesnake (*Crotalus durissus terrificus*) venom by enzyme-linked immunosorbent assay with parallel-lines analysis. Toxicon 38, 443–448.

Ownby, C., 1998. Structure, function and biophysical aspects of the myotoxins from snake venoms. J. Toxicol. Toxin. Rev. 17, 213–238.

Rádis-Baptista, G., Oguiura, N., Hayashi, M.A.F., Camargo, M.E., Grego, K.F., Oliveira, E.B., Yamane, T., 1999. Nucleotide sequence of crotamine isoform precursors from a single South American rattlesnake (*Crotalus durissus terrificus*). Toxicon 37, 973–984.

Samejima, Y., Aoki, Y., Mebs, D., 1991. Amino acid sequence of a myotoxin from venom of the Eastern Diamondback Rattlesnake (*Crotalus adamanteus*). Toxicon 29, 461–468.

Santos, M.C., dos Morhy, L., Ferreira, L.C.L., Oliveira, E.B., 1993. Purification and properties of a crotamine analog from *Crotalus durissus ruruima* venom. Toxicon 31, 166.

Schenberg, S., 1959. Geographical pattern of crotamine distribution in the same rattlesnake Subspecies. Science 129, 1361–1363.

Senapathy, P., Shapiro, M.B., Harris, N.L., 1990. Splice junctions, branch point sites, and exons: sequence statistics, identification, and application to genome project. Meth. Enzymol. 183, 252–278.

Vellard, J., 1939. Variations géographiques du venin de *Crotalus terrificus*. C. R. Seances Acad. Sci. 130, 463–464.

Viegas-Pequinot, E., 1992. In situ hybridization to chromosomes with biotinilated probes. In: Willkison, D.G., (Ed.), In situ Hybridization: a Practical Approach, IRL Press, Oxford, pp. 137–158.