Toxicon 58 (2011) 18–27

Contents lists available at ScienceDirect

Toxicon

journal homepage: www.elsevier.com/locate/toxicon

Biochemical and molecular characterization of the venom from the Cuban scorpion *Rhopalurus junceus*

B.I. García-Gómez [a], F.I.V. Coronas [a], R. Restano-Cassulini [a], R.R. Rodríguez [b], L.D. Possani [a,*]

[a] Department of Molecular Medicine and Bioprocesses, Instituto de Biotecnologia, Universidad Nacional Autonoma de Mexico, Avenida Universidad, 2001, Apartado Postal 510-3, Cuernavaca MOR. 62210, Mexico
[b] Centro de Desarrollo de la Montaña, Ministerio de Ciencia, Tecnología y Medio Ambiente, Limonar de Monte Roux, El Salvador Guantanamo, Cuba

ARTICLE INFO

Article history:
Received 22 December 2010
Received in revised form 17 February 2011
Accepted 12 April 2011
Available online 13 May 2011

Keywords:
Cuba
Gene cloning
Ion-channel
Scorpion toxin
Rhopalurus junceus

ABSTRACT

This communication describes the first general biochemical, molecular and functional characterization of the venom from the Cuban blue scorpion *Rhopalurus junceus*, which is often used as a natural product for anti-cancer therapy in Cuba. The soluble venom of this arachnid is not toxic to mice, injected intraperitoneally at doses up to 200 μg/20 g body weight, but it is deadly to insects at doses of 10 μg per animal. The venom causes typical alpha and beta-effects on Na^+ channels, when assayed using patch-clamp techniques in neuroblastoma cells *in vitro*. It also affects K^+ currents conducted by ERG (*ether-a-go-go related gene)* channels. The soluble venom was shown to display phospholipase, hyaluronidase and anti-microbial activities. High performance liquid chromatography of the soluble venom can separate at least 50 components, among which are peptides lethal to crickets. Four such peptides were isolated to homogeneity and their molecular masses and N-terminal amino acid sequence were determined. The major component (RjAa12f) was fully sequenced by Edman degradation. It contains 64 amino acid residues and four disulfide bridges, similar to other known scorpion toxins. A cDNA library prepared from the venomous glands of one scorpion allowed cloning 18 genes that code for peptides of the venom, including RjA12f and eleven other closely related genes. Sequence analyses and phylogenetic reconstruction of the amino acid sequences deduced from the cloned genes showed that this scorpion contains sodium channel like toxin sequences clearly segregated into two monophyletic clusters. Considering the complex set of effects on Na^+ currents verified here, this venom certainly warrant further investigation.

1. Introduction

Research on scorpion venoms has been largely restricted to a few species within the medically important Buthidae family, most of the studies deal with scorpions from the North Africa area, continental America, China, India and the Caucasian regions. Buthid scorpions are cosmopolitan (Polis, 1990), but far less is known about scorpion venoms from many other regions, including the Caribbean Islands. Striking among these scorpions is the Cuban *Rhopalurus junceus*, whose venom has been used by doctors to treat humans suffering from cancer (http://kinastchile.cls/ccc43.htm, http://www.escozul-cancer.com/es/escozul-investigacion.html). There are several communications dealing with various aspects of the venom from *R. junceus*, such as toxicity and pharmacology (Cao et al., 1997; Pérez et al., 2004; Hernández-Betancourt et al., 2009a), electrophoretic and chromatographic separation (Hernández-Betancourt et al., 2009b) and anti-microbial activity (Rodríguez et al., 2004), however, to the best of our knowledge, the venom from this species, and others from the same region, have not been studied using more in depth biochemical and molecular biological approaches. One of

* Corresponding author. Tel.: +52 777 3171209; fax: +52 777 3172388.
E-mail address: possani@ibt.unam.mx (L.D. Possani).

0041-0101/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.toxicon.2011.04.011

the few publications dealing with venom from the genus *Rhopalurus* is that performed with *Rhopalurus agamemnon*, by Nishikawa et al. (1994). It is worth recalling that scorpion venoms are known to contain many different peptides and proteins with a variety of pharmacological activities. The best documented are peptides that block or modify ion-channel functions in excitable and non excitable cells. The short-chain scorpion peptides, which contain from 20 to 43 amino acid residues are usually effective blocker of K^+-channels, whereas the long-chain scorpion toxins, containing from 58 to 76 amino acid residues are modifiers of the Na^+-channels gating mechanism (reviewed in: Possani and Rodríguez de la Vega, 2006). A substantial amount of literature is available on this subject and a general view of the biodiversity of components found in scorpion venoms has recently been published by our group (Rodríguez de la Vega et al., 2010).

This communication describes a general biochemical and molecular characterization of the venom from the Cuban species *Rhopalurus junceus*, from here on abbreviated *R. junceus*. Several peptides where isolated to homogeneity and a full amino acid sequence of a novel insect toxin was obtained. From a cDNA library several genes were cloned and are reported here. Additionally it was shown that this venom is not dangerous to mammals (mice) assayed at doses 40 fold higher than scorpions of the genus *Centruroides* of Mexico. However, the venom certainly affects the Na^+ currents of neuroblastoma cells, showing both an alpha and beta-effects, which are completely reversible. It also affects ERG potassium channels. Presently, a complete proteomic analysis of venom from *R. junceus* collected in different areas of Cuba is under investigation (Rodríguez R.R. et al., in preparation).

2. Material and methods

2.1. Venom, chemical and reagents

Venom of scorpions of the species *R. junceus* was collected in Guantanamo area of Cuba, and they were milked for venom by electrical stimulation. The venom was freeze-dried. Upon arrival in the laboratory in Mexico it was dissolved in distilled water and centrifuged at 15,000 *g* for 15 min, lyophilized immediately and stored at −20 °C until use. All chemicals were analytical grade reagents. The enzymes trypsin and protease V8 from *Staphylococcus aureus* were from Roche Diagnostics GmbH (Mannheim, Germany). Double distilled water over quartz was used all through the procedures. All reagents used were analytical grade as mentioned in earlier publications by our group (Barona et al., 2006; Batista el al., 2004, 2007).

2.2. Lethality and other general characteristics

The soluble venom was assayed using mice and crickets. Mice of the strain (CD1) were injected intraperitoneally with samples containing 50, 100 and 200 μg protein per 20 g body weight. The content of protein was estimated by reading the venom at 280 nm and assuming that one unit of absorbance is equal to 1 mg/ml solution in a 1 cm pathway cuvette. Crickets (strain *Acheta domestica*) were injected intrathoracically between the second and third pair of legs, with venom and fractions at concentrations from 10 to 30 μg. Non toxic venom or fractions means the animal injected showed similar symptoms of those injected with phosphate buffer saline (PBS), pH 7.4, used as control. Toxic means the insect lost equilibrium (problems to stand-up), partial paralysis, salivation, but usually recover after a few hours. Lethal means the cricket apart from showing the symptoms described became paralyzed and died within 24 h after injection.

Phospholipase and hyaluronidase activities were verified by the method described by Habermann, 1972, and Tolksdorf et al.,1949, respectively. The anti-microbial activity was identified as described in Rodriguez et al., (2004).

2.3. Chromatographic separations

The soluble venom was initially separated by high performance liquid chromatography (HPLC) using a C18 reverse-phase column (Vydac, Hisperia, CA) eluted with a linear gradient of solution A [0.12% trifluoroacetic acid (TFA) in water] to 60% solution B (0.10% TFA in acetonitrile), run for 60 min, using equipment described earlier by our group (Batista et al., 2004). Fractions were collected according to the absorbance at 230 nm, freeze-dried and when required further separated using different gradient conditions as mentioned in the figure legends.

2.4. Amino acid sequence and mass spectrometry determination

Amino acid sequence determination of pure peptides was performed by automatic Edman degradation in a Beckman LF 3000 Protein Sequencer (Palo Alto, CA, USA) using the chemicals and procedures previously described (Batista et al., 2004, 2007). Mass spectrometry determination was obtained with protein samples of 0.1–0.5 μg/μl dissolved in 50% acetonitrile with 1% acetic acid and directly applied into a Finnigan LCQ ion trap mass spectrometer (San Jose, CA, USA), using a Surveyor MS syringe pump delivery system, as initially described in Batista et al. (2007). The amino acid sequence of the peptide RjAa12f, fully sequenced, is deposited into UniProtKB accession number P86685.

2.5. Cloning and DNA sequencing

RNA was isolated from the venomous glands (telson) of one scorpion of the species *R. junceus*, according to the Promega Total RNA isolation system specifications.

For the 3′RACE reaction (Rapid Amplification of Complementary DNA Ends), the total RNA isolated was used for cDNA synthesis by means of a 3′ outer oligonucleotide from Ambion First Choice RLM-RACE Kit (Applied Biosystems) following all the specifications of the providers. For the polymerase chain reaction (PCR), a sample of the cDNA first strand reaction (1 μl) was added to the solution containing: 1X Vent DNA polymerase buffer (in mM concentration: 10 KCl, 10 $(NH_4)SO_4$, 2 $MgSO_4$, 20 Tris–HCl at pH 8.8, 0.1% Triton X-100, at 25 °C) plus 200 μM dNTPs, 0.25 μM 15-

mer forward degenerated primer (5′-AAR GAR GGN TAY CCN-3′),0.25 μM 3′RACE outer primer and two units of Vent DNA polymerase (New England Biolabs, Beverly, MS, USA) in a final volume of 50 μl. The reaction was performed using a Perking Elmer 9600 thermo cycler with the following protocol: incubation of the mixture during 3 min at 94 °C, plus 7 min at 42 °C before addition of the enzyme, followed by 30 s at 72 °C for one cycle. After this initial cycle, the mixture was incubated at 94 °C for 30 s followed by 42 °C for 40 s and 72 °C for 30 s per cycle, and repeated 35 times before a 5 min final step at 72 °C. The main product of this amplification reaction had a size of approximately 350 base-pairs (bp). The final cloning and sequencing were obtained using the PCR products purified in a Centricon 100 Column (Amicon, Beverly, MA, USA) following the manufacturer instructions and then ligated into the EcoRV of pKS- phagemid (Stratagene, La Jolla, CA, USA). The construct was used to transform *E. coli* DH5-α cells. Selection of the clones bearing an insert was performed by plating transformants on a Petri dish containing LB/agar in the presence of X-gal/IPTG. White colonies were grown for plasmid preparation. Plasmid DNA was sequenced from both strands, using fluorescent nucleotides in an automatic Perkin Elmer Applied Biosystems apparatus (Forster City, CA. USA), as described by the manufacturer. For the 5′ RACE reaction the total RNA isolated was used for cDNA synthesis with the outer oligonucleotide from Ambion First Choice RLM-RACE Kit (Applied Biosystems) following all the specifications of the providers.

The 5′ nucleotide sequence for the most representative toxin-clone (RjAa2) was obtained and used to design two specific reverse primers (Rjrev1 and Rjrev2). The polymerase chain reaction (PCR) conditions were the following: a sample of the first strand reaction cDNA (1 μl), was added to 1X Vent DNA polymerase buffer (same as described before for the 3′RACE), plus 200 μM dNTPs, 0.25 μM 5′ RACE Outer primer from the kit and a specific Rjrev1 primer (5′-ATT GGA TTA AAT GTC CGA GG-3′) and two units of Vent DNA polymerase (New England Biolabs, Beverly, MS, USA) in a final volume of 50 μl. The product previously obtained made use of a template in a nested PCR reaction, using the 5′ RACE-Inner primer provided in the RLM-RACE Kit and the specific internal Rjrev2 primer (5′-AAT ATC CAT CCG ATT GTC TCC A-3′), using the same PCR reaction conditions.

Following amplification, 5′ RACE products were cloned into the EcoRV site of pKS- phagemid (Stratagene) with blue-white selection. The ligation reaction was used to transform competent *E. coli* DH5-α cells. Positive clones were sequenced from both strands, using fluorescent nucleotides in an automatic Perkin–Elmer Applied Biosystems apparatus (Foster City, CA), as recommended by the manufacturer. The nucleotide sequences determined are deposited into GenBank accession numbers HM233939 to HM233956, and JF309048.

2.6. Complete nucleotide sequence of RjAa12f clone

The complete nucleotide sequence for the RjAa12f toxin was obtained using as a template a cDNA synthesized from total RNA. For the PCR reaction a specific external RjAa12f direct primer (5′-GAA GTC AAA ATG AAG ATT TTG ATA T-3′) and a reverse 3′RACE outer primer from the kit were used using the same PCR conditions described above.

2.7. Electrophysiological measurements

F11 cells (a fusion product of mouse neuroblastoma cells with embryonic rat dorsal-root ganglion – DRG- neurons) were routinely cultured in DMEM (Dulbeco's Modified Eagle Medium, Gibco) supplemented with 10% fetal bovine serum (Gibco), at 37 °C in 95% humidity and 5% CO_2 atmosphere. Standard external solution contained (in mM): 130 NaCl, 5 KCl, 2 $CaCl_2$, 2 $MgCl_2$, 10 HEPES, 5 glucose, pH 7.3 adjusted with NaOH. Standard intracellular solution contained (in mM): 10 NaCl, 2 $MgCl_2$, 130 potassium aspartate, 10 EGTA, 10 HEPES buffer adjusted with KOH at pH 7.3. A high potassium external solution ($[K]_o = 40$ mM) was used during ERG potassium currents records, where NaCl was replaced by an equimolar amount of KCl. From a holding potential at −80 mV, cells were stepped at +60 mV for 500 ms and the ERG tail current was recorded during the subsequent step at −120 mV (Rodríguez de la Vega et al., 2009). Sodium currents from F11 cells were elicited by a depolarization steps protocol from −80 to +20 mV, from a holding potential of −80 mV. Where indicated, a 5 ms pre-pulse at 50 mV preceded the depolarization step. Patch-clamp experiments were performed by using a Multiclamp 700B amplifier with a DigiData 1440A. pClamp 10 (Molecular Device, U.S.A.) and Origin 7 (Microcal Inc, USA) software were routinely used during data acquisition and analysis.

2.8. Sequence analyses

Amino acid sequences were aligned with MAFFT (Katoh and Toh, 2008) and the alignments submitted to hhm search (Eddy, 1998) against UniProt database. The phylogenetic relationships of *R. junceus* sequences were reconstructed by Bayesian inference with MrBayes 3.1 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003) as described elsewhere (Rodríguez de la Vega and Possani, 2005). Briefly, 219 sodium channel scorpion toxins (NaScTx) sequences were aligned with CLUSTAL_X and manually refined in gap-rich regions. The alignment was then used to reconstruct the phylogeny under the Wag substitution model, with a mixed rate model with invariant and gamma distribution sites. The last 250 sampled trees of two parallel runs, 500,000 generations each sampled every 500 iterations, were merged and used to calculate the 50% consensus tree.

3. Results and discussion

3.1. General venom characterization

The soluble venom of this scorpion when applied intraperitoneally into mice at doses up to 200 μg per 20 g mouse weight causes a slight discomfort, visible by increments of bowel movement and abdominal stretching and distention, but does not show typical symptoms of intoxication such as the one seen when injecting venom of Buthidae scorpions of the genera *Centruroides* or *Tityus*.

These symptoms typically include important excitability, respiratory problems, intense nose and eye secretions, convulsions, paralysis of legs and eventually death (Possani et al., 1977). Medical doctors in Cuba have not reported intoxication symptoms from persons stung by this species. Actually, one of us (RRR) when collecting scorpion in the field was stung in different occasions, and only felt intense pain at the moment of the accident, followed by a mild paresthesia that disappeared after a few hours. Thus, it seems that the venom causes a mild anesthesia in the local site of the sting. No respiratory or circulatory problems were felt. Concerning the cricket experiments, soluble venom of this scorpion is lethal (see Material and Methods) at doses of 10 μg per animal. This is also comparable to what we have observed with venom from the *Centruroides* species of Mexico (Selisko et al., 1996).The venom also displays some enzymatic activities such as hyaluronidase and phospholipase. While hyaluronidase activity appears ubiquitous, all venom samples obtained from scorpions collected in distinct geographical areas showed this activity, phospholipase activity was observed with scorpion collected in the humid area of Baracoa, Guantanamo Province (Rodríguez, R.R. et al., in preparation). The presence of anti-microbial activity was verified only with freshly collected venom, as described by Rodriguez et al., (2004). In our hand, the lyophilized and stored samples of venom did not show this activity. Because the soluble venom was not toxic to mammals at doses 40 times higher than other classical toxins from Buthidae scorpions, such as *Centruroides noxius*, whose LD50 is 5 μg/20 g body weight (Licea et al., 1996), and due to the possibility of a cancer inhibition activity of some venom fractions, it was important to verify if the venom contained components capable of recognizing mammalian Na^+- and K^+-ion channels, as most scorpions dangerous to human do. In order to test this possibility, the venom activity was assayed on ion channels endogenously expressed in F11 cell line (Fig. 1 and Fig. 2). These cells were previously successfully used in venom (Corona et al., 2002) and toxin screening (Gurrola et al., 1999; Barona et al., 2006). In vivo, 200 μg/20 gr mouse of *R. junceus* venom has no toxic effect, so we decided to test a higher concentration (500 μg/ml) to look for possible effects on mammalian channels. In our previous experiments (Corona et al., 2002), using venom from the Mexican buthid scorpions of the genus *Centruroides* the application of 180 μg/ml venom produced an almost complete blockade of potassium channels in the same cellular model. Whole-cell sodium currents were elicited with the pulse stimulations shown in panels *d* and *h* of Fig. 1 (see methods section for details). Fig. 1a–c shows sodium currents in control, after 0.5 mg/ml venom application and after recovery; for clarity only the traces corresponding to −80, −60, −40 and −10 mV are reported. Application of 0.5 mg of soluble venom resulted in a modification of both activation and inactivation kinetics. In control condition at −40 mV almost all channels are in a closed state and no current are visible (Fig. 1a, gray trace), but in presence of the venom, channels are able to open at more negative potential, resulting in a consistent current also at −40 mV (Fig. 1b, graytrace). This effect is described as beta-effect (Rodríguez de la Vega and Possani, 2007) and is dependent from the brief pre-pulse used in the protocol in Fig. 1d (Cestèle et al., 1998). In addition, the venom produces an inactivation kinetic delay (Fig. 1b and f), that is, the channel takes more time to inactivate. When the same type of experiment is performed by using a protocol without pre-pulse (Fig. 1h), no beta-effect are visible, but the inactivation kinetic delay is still present (Fig. 1f). In control experiments, at −10 mV, after 7.5 ms, almost all channels are inactivated (Fig. 1e, arrow), but in presence of the venom, in the same condition, part of the channels are still open resulting in about 20% of the maximal current (Fig. 1f, arrow). This is known as alpha-effect (Bosmans and Tytgat, 2007). Both alpha-and beta-effect produced by the venom resulted reversible (Fig. 1c and g). Based on these results it is clear that the soluble venom of *R. junceus* must contain peptides that belong to both the alpha and beta-scorpion toxin types.

Using the same cellular line F11, the effect of the soluble venom was assayed at similar concentration (0.5 mg/ml) on ERG potassium channels. Fig. 2a shows an example of the traces obtained. It shows the ERG currents in control (black line), after venom application (gray line) and after recovery (black dotted line). Currents were elicited by using the protocol shown in the inset *b* (see methods section for details).Venom application produced a rapid and reversible block of ERG channels as shows in Fig. 2a and in Fig. 2c, where the fractional currents were plotted against time to obtain the blockade (black bar corresponds to the venom application), followed by recovery after washing without toxin.

3.2. Venom fractionation and peptide sequencing

The soluble venom when separated by HPLC shows to contain more than 50 chromatographic components as it can be seen in Fig. 3 and confirmed by mass spectrometry determination (Rodríguez, R.R. et al, in preparation). Most of the components are eluting from the C18 reverse-column around minutes 20 to 40 of the gradient, which is usually the place where the K^+- and Na^+-channel specific toxins elute in other scorpion venoms (Batista et al., 2007). Several components are very well represented into the venom, making difficult to visualize components present in lower relative concentrations. Five major components, collectively accounting for about 64% of the total venom, were further used for purification and assay (see Fig. 3). After additional chromatographic separations of these fractions, at least 5 peptides were obtained in homogeneous form. Insets *a* and *b* of Fig. 3 shows the profile of two examples: one from the fraction that elutes at 34.34 and the other at 37.99 min. Components labeled with asterisk were homogeneous and were further characterized.

The major component of the fraction eluting at 30.06 min had a molecular mass of 8099 atomic mass units (amu), with the N-terminal amino acid sequence: KEGYPKNSEGCK ITCLFNDPYCKGLCINLSTQADY… This peptide injected into crickets (20 μg) caused symptoms of intoxication, accompanied by loss of equilibrium (could not stand-up) and died within 12 h after injection.

The peptide eluting at 30.50 min had a molecular mass of 8073 amu, with the N-terminal sequence:

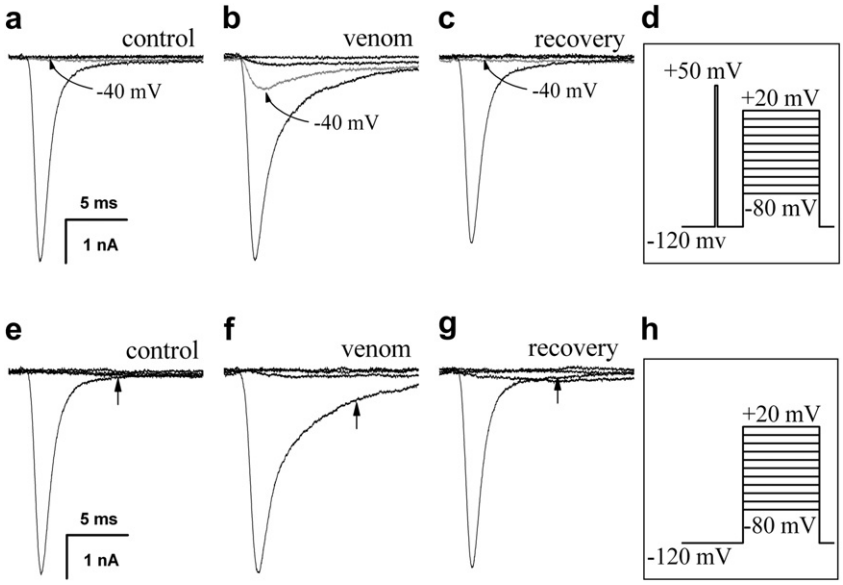

Fig. 1. Soluble whole venom effect on sodium channels in F11 cell line The soluble venom of *R. junceus* was applied at concentration of 0.5 mg/ml. Sodium currents were elicited with protocols shown in panels *d* and *h*, where steps depolarization were preceded or not by a strong brief depolarization pre-pulse respectively (see methods for details). In panels a–c and e–g are depicted representative currents and for clarity only the traces corresponded at −80, −60, −40 and −10 mV are reported. When the pre-pulse was applied, the venom produced a left shift of the voltage dependent activation (beta-effect): in control at −40 mV all channels are still closed (panel a, gray trace) but after venom application at the same potential some of the channels are able to open generating a visible current (panel b, grey trace). This effect is reversible as shown in panel *c*. Venom also produces a delay in the inactivation kinetic (alpha-effect) that is well appreciable using a protocol without pre-pulse: in control at −10 mV and after 7.5 s all channels reach the inactivation conformation (panel e, arrow); in presence of venom at the same potential and after the same time, many channels are still open (panel f, arrow). Also this effect is reversible as shown in panel *h*.

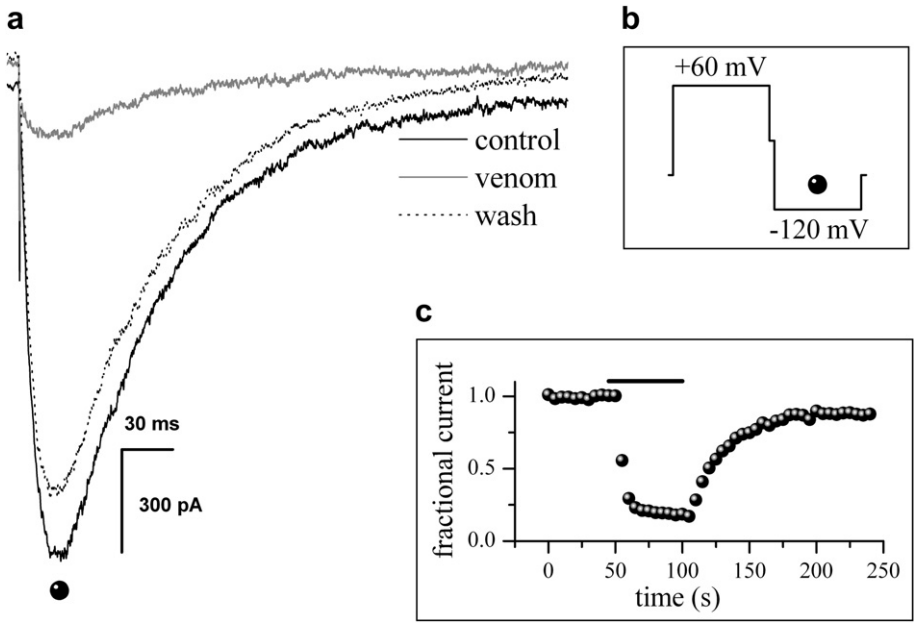

Fig. 2. Soluble whole venom effect on ERG potassium channels in F11 cell line The soluble venom of *R. junceus* was applied at concentration of 0.5 mg/ml. ERG potassium currents were elicited with protocols shown in panel *b*, where tail currents were elicited at −120 mV (panel *b*, sphere), after a depolarization at 60 mV (see methods for details). Panel a show traces in control (black) after venom application (gray) and after washing (black dot). Venom produces a rapid and reversible block of ERG current as also shown in panel c, where fractional currents were plotted against time to produce a time-course graphic (black bar correspond to the venom application).

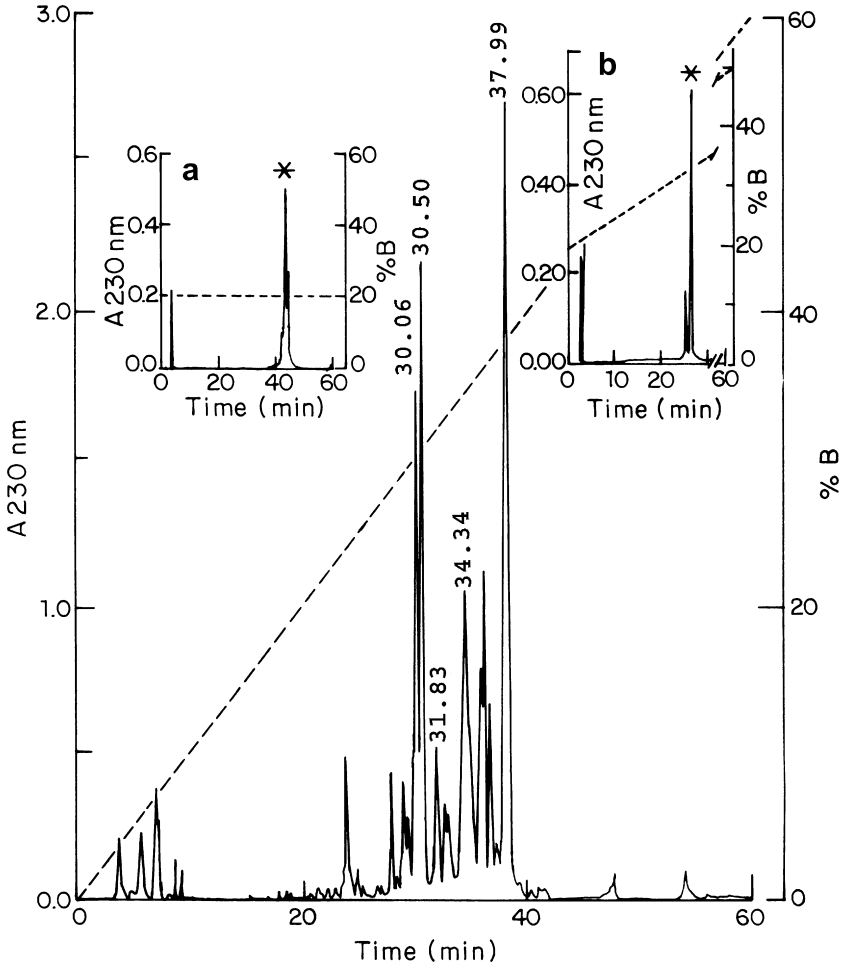

Fig. 3. HPLC separation of soluble venom Soluble venom (1 mg) was applied to a C18 reverse-phase column and resolved using a gradient from solution A to 60% solution B, during 60 min (see material and methods). Fractions eluting at 30.06, 30.50, 31.83, 34.34 and 37.99 min corresponded respectively to 8.1, 14.5, 3.5, 10.5 and 27.3% of the material recovered. Insert *a* shows the profile of further separation of fraction eluting at 34.34 min, using a C18 reverse-phase column run for 60 min with constant concentration of 20% solution A. From a sample containing 180 μg applied to the column, the peptide indicated by the asterisk corresponded to 54% of the material recovered. Insert *b* shows HPLC profile of further separation of 80 μg of component that elutes at 37.99 min using a gradient from 20% to 50% solution B, during 60 min. Component labeled with asterisk was the main component, and corresponded to 84% of the material recovered.

KEGYPTNSEGCKITXLFNDPYCKGXCINLSTQAD.., where X stands for a non identified amino acid. This component injected into crickets at doses of 20 μg did not show any symptom of intoxication.

Peptide eluting at 31.83 min when sequenced gave the N-terminal sequence: KEGYPDGQNGKKIPCAINDNISKTXE QA… This peptide injected into crickets (20 μg) showed scarce symptoms of intoxication, such as some difficulties to stand-up and mild paralysis, but recover overnight.

The major component of the fraction that elutes at 34.34 min (see inset a, labeled with asterisk) had a molecular mass of 7402 amu, with the N-terminal amino acid sequence:
KEGYMGSDGCKMSCVINDQFCDTECQAKLKGSTGYCYFXGLA-CYXXG… This component injected into crickets at doses of 20 μg/animal showed paralysis of rear limbs and some contracture of the body, but recovered overnight.

The available material allowed to fully determining the primary structure of the principal component, corresponding to 25% of the soluble venom (inset *b* of Fig. 3). This component, named RjAa12f, when subjected to direct Edman degradation procedure gave the first 37 amino acid residues of the N-terminal region. The peptide was digested with two hydrolytic enzymes: trypsin and endoprotease V8 from *S. aureus*, using conditions earlier described by our group (Batista et al., 2004, 2007). Fig. 4A shows the HPLC profile of the separation of this material after tryptic digestion. The sub-peptides indicated by numbers were further applied to the amino acid sequencer and the overlapping sequence shown in Fig. 4B was obtained. For details see legend of Figs. 3 and 4. The sequence of the peptides obtained after endoprotease V8 cleavage were confirmatory of the sequence found and are not shown here. The purified peptide has 64 amino acid residues, stabilized by 4

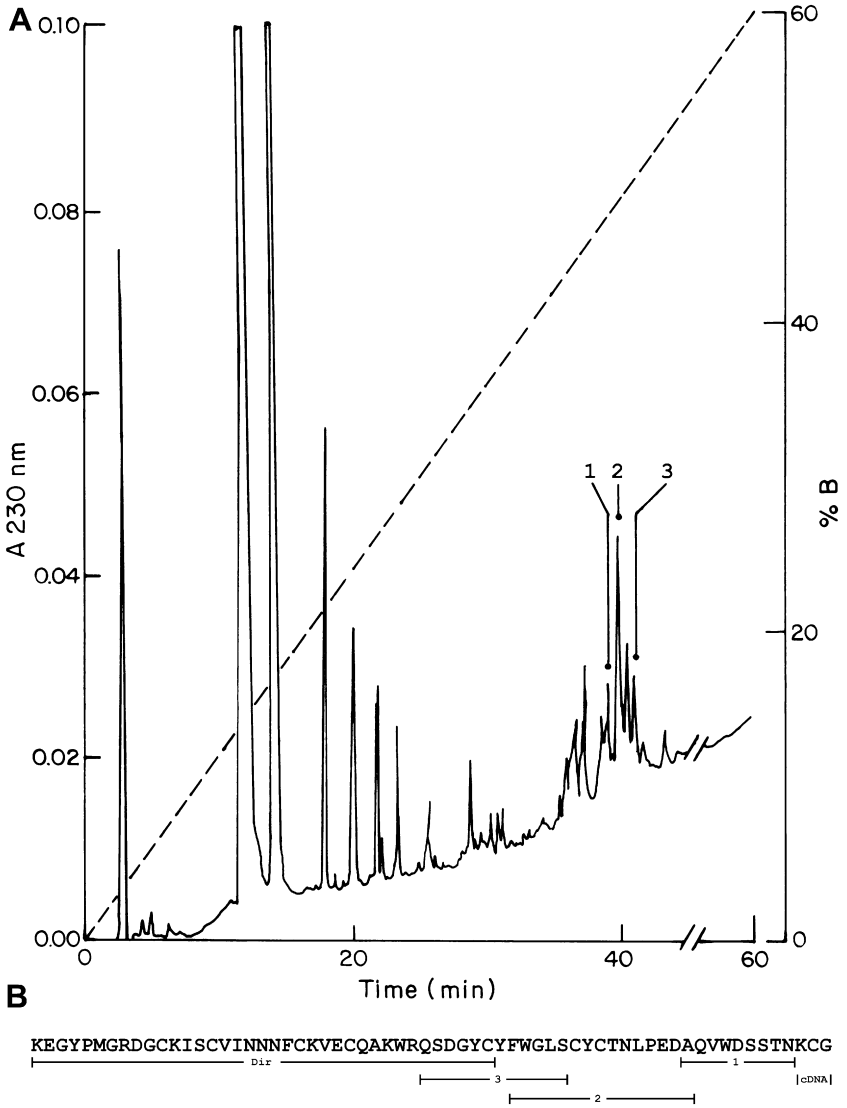

Fig. 4. HPLC separation of digested peptide and amino acid sequence of peptide RjAa12_fp. A., Pure peptide from Fig. 3, insert b (40 μg) was digested with trypsin (1:20 ratio enzyme:peptide), using a 100 mM Tris–HCl buffer, pH 8.5, incubated overnight at 37 °C. The digested peptides were separated into a C18 reverse-phase column using a linear gradient from solution A to 60% solution B, during 60 min. Peptides numbered 1 to 3 were necessary to complete the overlapping sequence, however many other peptides were sequenced, both from the tryptic digestion as well from the protease V8 digestion but were not used, because they were confirmatory of the final sequence found. The last three residues KCG were confirmed by the cDNA clone of gene RjAa12f. B. The first 37 amino acids at the N-terminal region were directly determined by Edman degration (see labeled Dir). Several peptides purified as indicated in Fig. 4A were sequenced by Edman degradation. The sequences used for overlapping segments were obtained from peptide 1 (amino acids A53 to N61), peptide 2 (amino acids F39 to A54 and peptide 3 from Q32 to S43. The three last residues were identified from the cDNA sequence of clone RjAa12f.

disulfide bridges, similar to other scorpion toxins. As discussed in the next section, a gene containing identical sequence was isolated and cloned. The theoretical expected mass value for the peptide sequenced was 7294.13 amu and the molecular weight experimentally determined was 7293.5 amu, within an acceptable error of the spectrometer used for this determination. As it will be discussed below the cloned gene had two glycine residues at the C-terminal region, before the stop codon, which suggests that the mature peptide is processed after transcription, loosing one of the last glycines.

Thus, RjAa12f is the first peptide fully sequenced from the venom of the scorpion *R. junceus*. It is not toxic to mammals, but it is lethal to crickets, which when injected with the pure peptide showed the classical symptoms of problems to stand-up, paralysis and died within 24 h after injection. This peptide was also tested for its function on F11 sodium channels. Experiments were performed as previously described for the total venom assay, using up to 10 μM of the peptide. At this concentration no significant effect was produced by the RjAa12f (data not shown). Since this peptide is capable of killing crickets at doses of 20 μg, but is not toxic to mammal, and in consequence of the lack of efficacy on mammal sodium channel, it is reasonable to assume that RjAa12f is an insect sodium toxin.

3.3. Gene cloning

Fig. 5 shows the nucleotide sequence of the gene RjAa12f and its corresponding amino acid sequence. The signal peptide is underlined, the 5′ and 3′ non-translated region are written in lower-case letters and the putative polyadenilation site is in bold. The mature peptide has 65 amino acid residues, whereas the signal peptide codes for a peptide 18 amino acid residues long. As mentioned in the previous section the cloning of this gene helped confirming the primary structure found by direct Edman degradation of the major component that elutes at 37.99 min, which is actually the most abundant component of the soluble venom. The last two residues at the C-terminal side of the gene are both glycine. However, the last one is probably processed during the maturation period. We have not seen this amino acid during the direct sequencing determination of the pure peptide. We assume that it is eliminated by post-translational modification, as it is the case for many scorpion toxins (Becerril et al., 1993). The molecular weight found experimentally is in agreement with the sequence determined by Edman degradation, as mentioned before. Thus, either the last residue is eliminated after transcription by a carboxypeptidase or there are two genes with identical sequence, except that one has two glycines at the C-terminal region.

Several additional genes were cloned and sequenced as indicated by their deduced amino acid sequences in Fig. 6. This figure shows 18 amino acid sequences deduced from *R. junceus* cloned genes (all abbreviated by the letters RjAa; 4 in Fig. 6A and 14 in Fig. 6B). Six genes were cloned including the region encoding signal peptides). For information regarding the signal sequence consult GeneBank, accession numbers HM233939 to HM233956, and JF309048. Comparing these sequences with other known scorpion toxins it is clear that these genes code for peptides with same or very similar length and identical number of disulfides of the known Na^+-channel specific peptides described in scorpion venom (Rodríguez de la Vega and Possani, 2005).

3.4. Sequence analyses

Multiple sequence alignment based on translated sequences revealed the cloned genes conforms two groups of sequences, one comprising RjAa9, RjAa13, RjAa16 and RjAa44 (Fig. 6A), while the other contains the remaining sequences, including RjAa12f (Fig. 6B). The star (*) after RjAa12f showed that this was the leading sequence used for comparison with the other ones. Database searches were carried out in order to identify the closest homologs amongst known scorpion toxins. Fast HMMer against UniProt retrieved NaScTx sequences from *Centruroides*, *Tityus* and *Parabuthus* species when the alignment of the first group was used as query, whereas the second group retrieved NaScTx sequences from *Centruroides*, *Tityus* and several Old World scorpions (see Fig. 6).

Both database searches and phylogenetic reconstruction (Fig. 6A) supports a monophyletic RjAa9, RjAa13, RjAa16 and RjAa44 plus CsE5 (P46066). The closest relatives of these groups are alpha NaScTx from South American *Tityus* scorpions (v.gr. Ts3 and Tb3) and promiscuous toxins from South African *Parabuthus* species (v.gr. Kurtoxin, which modulates both Na^+ and Ca^{2+} channels (Chuang et al., 1998)). On the other hand, sequences related to RjAa12f clustered as a single group within beta NaScTx. Two alternative groupings were retrieved in the final tree set, one supports monophyletic RjAa12f-related sequences plus CsEv5 (P58779) group as a sister clade of *Tityus* gamma-like sequences, while the other places them as the sister group of Old World insect-selective NaScTx. Despite the phylogenetic reconstructions implemented here (available upon request) it did not resolve the high order relationships amongst beta NaScTx. It is interesting to note that RjAa12f-related sequences always appeared closer to insect-selective toxins, in agreement with the lethality test conducted here with purified RjAa12f.

Although the sequence diversity of NaScTx preclude homology-based classification of RjAa sequences, it is clear from this analysis that *R. junceus* has followed a completely different route of NaScTx diversification as compared with the two most prevalent buthids in the New World. Venoms from *Centruroides* species are known to contain almost exclusively classic beta type NaScTx, whereas *Tityus* species are rich in gamma-like NaScTx with only a few NaScTx with alpha type activity (Rodríguez de la Vega and Possani, 2005, 2007). *R. junceus* hyperdiversified Ts3/CsE5/Kurtoxin-like sequences are quite striking, providing, except for CsE5, no-other close homologs of this group are known outside austral scorpions (*Tityus* and *Parabuthus* genera).

3.5. Final commentaries

In the introductory section of this communication we mentioned the use of *R. junceus* venom for treatment of

```
gaaqtcaaaATGAAGATTTTGATATTCATCATCGCTTCTTTCATGCTTATTGGCGTAGAG      60
          M  K  I  L  I  F  I  I  A  S  F  M  L  I  G  V  E        17
TGCAAAGAGGGATATCCTATGGGTAGAGATGGTTGCAAAATCTCCTGTGTAATAAATAAT     120
C  K  E  G  Y  P  M  G  R  D  G  C  K  I  S  C  V  I  N  N       37
AACTTTTGCAAAGTTGAATGCCAAGCGAAATGGAGACAATCGGATGGATATTGCTACTTT     180
N  F  C  K  V  E  C  Q  A  K  W  R  Q  S  D  G  Y  C  Y  F       57
TGGGGACTGTCGTGCTATTGTACAAATCTACCAGAAGACGCCCAGGTTTGGGATTCTAGC     240
W  G  L  S  C  Y  C  T  N  L  P  E  D  A  Q  V  W  D  S  S       77
ACCAATAAATGTGGAGGATAAtgtaaccgtcactcaatgcctcggacatttaatccaatg     300
T  N  K  C  G  G  *                                              83
taatgtattctcactgataaca**aaaaataaaa**gcatataatagttaaaagaaaaaaaaaaaa  362
```

Fig. 5. Nucleotide sequence encoding RjAa12f The deduced amino acid sequence is given below the nucleotide sequence. The 24-mer oligonucleotide used for the PCR amplification is underlined (5′-gaa gtc aaa ATG AAG ATT TTG ATA T-3′). The 3′non-traduced region is written in lower-case letters, and the polyadenylation signal is in bold. The number of the nucleotides (total 362) and position of the amino acid residues (83) are indicated on the right side of the figure. Protein sequence deposited at UniProtKB accession number P86685.

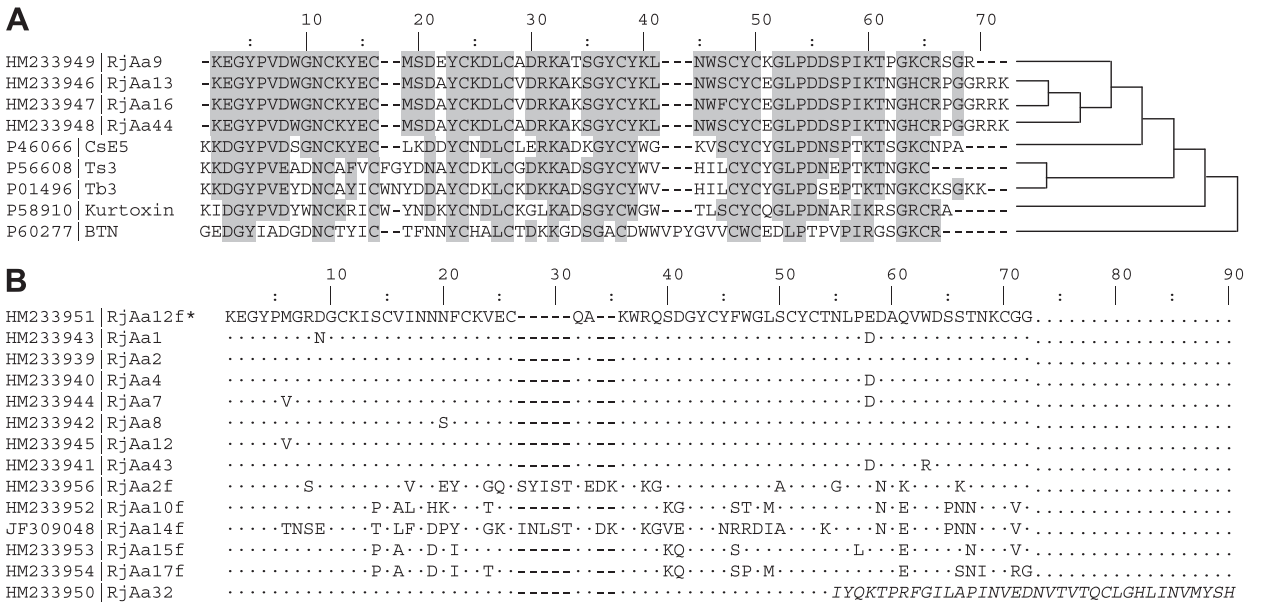

Fig. 6. Multiple sequence alignments of *R. junceus* NaScTx A) Deduced amino acid sequences of RjAa9, RjAa13, RjAa16 and RjAa44 aligned with CsE5, Ts3, Tb3, Kurtoxin and BTN (sister clade). The simplified cladogram on the right depicts the phylogenetic relationships as inferred by Bayesian reconstruction, all partitions occurred in at least 55% of the final tree set (details available upon request). UniProt accession codes are provided for nont *R. junceus* sequences. B) Multiple sequence alignment of RjAa12f-related sequences based on nucleotide sequences of cDNA clones. Dots (·) indicate the same amino acid as in RjAa12f (asterisk). Dashes (-) indicate insertions in RjAa2f and RjAa14f. An single nucleotide insertion in RjAa32 (residues in italics) leads to a frameshift of the coding sequence, which otherwise remains identical at the nucleotide level to RjAa12f (see GenBank entries HM233950|RjAa32 and HM233951|RjAa12f).

cancer. This scorpion is known by the vulgar name of "Blue Scorpion" and a diluted solution of its venom known as "Escozul" (a trade mark) is an alternative drug used for cancer treatment, mainly in Cuba. We want to make it clear that this report does not advocate in favor or against this product. Based on our experiments here reported, this venom does not seem to be toxic to mammals, although it contains toxins to insects. It does contain components that recognize Na^+- and K^+-channels of excitable cells. It contains phospholipase and hyaluronidase enzymes, as well as an anti-microbial component similar to many others scorpion venom. The most abundant peptide in its venom is an insect toxin with 64 amino acid residues, compacted by 4 disulfide bridges, similar to many other insect toxins described in scorpion venoms (see review by Possani and Rodríguez, 2006). A partial genomic analysis suggests that it contains genes that code for some unusual peptides, for which the actual function is totally unknown at this moment. In conclusion, more studies should be performed with the venom of this interesting "Blue scorpion" of Cuba.

Acknowledgments

The authors are obliged with Prof. Misael Bodier Chivaz from the *Facultad de Ciencias Médicas de Guantanamo*, Cuba and their collaborators, for providing us with the first sample of venom used for this article. This publication is dedicated to his memory. Prof. Bordier died before this article was written and without his consent, we could not include his name as co-author, however his contribution to this communication was important. The authors are deeply indebted to Dr. Ricardo Rodríguez de la Vega from the European Molecular Biology Laboratory of Heidelberg, Germany, for conducting the phylogenetic analysis presented here. The authors want also to thank Dr. Fernando Z. Zamudio for helping with the amino acid sequencer and Dr. Cesar Ferreira Batista of the Unit of Proteomic of the Biotechnology Institute for the molecular mass determination of the pure peptides. This work was partially supported by a grant from Instituto Bioclon S.A de C.V. (Mexico) and Dirección General del Personal Académico of UNAM (DGAPA No. 204110 to LDP). Rodolfo RodríguezRavelo (abbreviated Rodríguez, R.R.) is a Ph.D. student, recipient of the scholarship UNAM - University of Habana, Cuba.

References

Barona, J., Batista, C.V., Zamudio, F.Z., Gomez Lagunas, F., Wanke, E., Otero, R., Possani, L.D., 2006. Proteomic analysis of the venom and characterization of toxins specific for Na^+ and K^+ channels from the Colombian scorpion *Tityus pachyurus*. Biochim. Biophys. Acta 1764 (1), 76–84.

Batista, C.V., del Pozo, L., Zamudio, F.Z., Contreras, S., Becerril, B., Wanke, E., Possani, L.D., 2004. Proteomics of the venom from the Amazonian scorpion *Tityus cambridgei* and the role of prolines on mass spectrometry analysis of toxins. J. Chromatogr. B Analyt. Technol. Biomed. Life Sci. 803 (1), 55–66.

Batista, C.V., Román-González, S.A., Salas-Castillo, S.P., Zamudio, F.Z., Gómez-Lagunas, F., Possani, L.D., 2007. Proteomic analysis of the venom from the scorpion *Tityus stigmurus*: biochemical and physiological comparison with other *Tityus* species. Comp. Biochem. Physiol. C. Toxicol. Pharmacol. 146 (1–2), 147–157.

Becerril, B., Corona, M., Mejía, M.C., Martin, B.M., Lucas, S., Bolivar, F., Possani, L.D., 1993. The genomic region encoding toxin gamma from the scorpion *Tityus serrulatus* contains an intron. FEBS Lett. 335 (1), 6–8.

Bosmans, F., Tytgat, J., 2007. Voltage-gated sodium channel modulation by scorpion alpha-toxins. Toxicon 49, 149–158.

Cao, J., Rivera, F., Bello, F., 1997. Algunos aspectos bioecologicos y farmacologicos del veneno crudo procedente de dos especies de escorpiones cubanos. Resumenes IV Simposio de Zoologia 70 La Habana.

Cestèle, S., Qu, Y., Rogers, J.C., Rochat, H., Scheuer, T., Catterall, W.A., 1998. Voltage sensor-trapping: enhanced activation of sodium channels by beta-scorpion toxin bound to the S3–S4 loop in domain II. Neuron 21 (4), 919–931.

Chuang, R.S., Jaffe, H., Cribbs, L., Perez-Reyes, E., Swartz, K.J., 1998. Inhibition of T-type voltage-gated calcium channels by a new scorpion toxin. Nat. Neurosci. 1, 668–674.

Corona, M., Gurrola, G.B., Merino, E., Restano-Cassulini, R., Valdez-Cruz, N. A., García, B., Ramírez-Domínguez, M.E., Coronas, F.I.V., Zamudio, F.Z., Wanke, E., Possani, L.D., 2002. A large number of novel Ergtoxin-like genes and ERG K^+-channels blocking peptides from scorpions of the genus. Centruroides. FEBS Lett. 532, 121–126.

Eddy, S., 1998. Profile hidden Markov models. Bioinformatics 14 (9), 755–763.

Gurrola, G.B., Rosati, B., Rocchetti, M., Pimenta, G., Zaza, A., Arcangeli, A., Olivotto, M., Possano, L.D., Wanke, E., 1999. A toxin to nervous, cardiac, and endocrine ERG K^+ channels isolated from Centruroides noxius scorpion venom. FASEB J. 13 (8), 953–962.

Habermann, E., Hardt, K.L., 1972. A sensitive and specific plate test for the quantitation of phospholipases. Anal. Biochem. 50 (1), 163–173.

Hernández-Betancourt, O., Casado-Hernández, I., Iglesias-Huerta, E., Ramírez-Labrada, A., del Risco-Ramos, J., Rodríguez-Pargas, A., 2009a. Evaluación de la toxicidad in vitro del veneno del alacrán *Rophalurus junceus* a través de un ensayo celular. Rev. Cubana Invest. Biomed. 28, 1–11.

Hernández-Betancourt, O., Compte, A.O., Quesada-Leiva, L., Rodríguez-Pargas, A., 2009b. Caracterización electroforética y cromatográfica del veneno del alacrán *Rhopalurus junceus*. Acta. Med. Colomb. 13 online.

Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17 (8), 754–755.

Katoh, K., Toh, H., 2008. Recent developments in the MAFFT multiple sequence alignment program. Brief Bioinform. 9 (4), 286–298.

Licea, A.F., Becerril, B., Possani, L.D., 1996. Fab fragments of the monoclonal antibody BCF2 are capable of neutralizing the whole soluble venom from the scorpion Centruroides noxius Hoffmann. Toxicon 34, 843–847.

Nishikawa, A.K., Caricati, C.P., Lima, M.L.S.R., Dos Santos, M.C., Kipnis, T.L., Eickstedt, V.R.D., Knysak, I., Da Silva, M.H., Higashi, H.G., Dias Da Silva, W., 1994. Antigenic cross-reactivity among the venoms from several species of Brazilian scorpions. Toxicon 32, 989–998.

Pérez Capote, M., Rodríguez, C., Guevara, I., Romeaux, R., 2004. Estudio de toxicidad aguda de las toxinas de los escorpiones *Rophalurus junceus* y Centruroides Gracilis. Rev. Cub. Farm., 38. Supl. Esp.

Polis, G.A., 1990. The Biology of Scorpions. Stanford University Press, Stanford, CA, USA.

Possani, L.D., Alagon, A., Fletcher Jr., P., Erickson, B.W., 1977. Purification and properties of mammalian toxins from the venom of the Brazilian scorpion *Tityus serrulatus* Lutz & Mello. Arch. Biochem. Bioph. 180, 394–403.

Possani, L.D., Rodríguez de la Vega, R.C., 2006. Scorpion venom peptides. In: Kastin, A.J. (Ed.), The Handook of Biologically Active Peptides. Academic Press, San Diego, CA, USA, pp. 339–354.

Rodríguez, R., Guerra, O., Baro, E., Silva, J., Magdelaine, R., 2004. Acción antimicrobiana Del veneno del escorpion *Rhopalurus junceus* (Scorpionida: buthidae). Revista Cubana de Farmacología 38, 134.

Rodríguez de la Vega, R.C., 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.

Rodríguez de la Vega, R.C., Possani, L.D., 2007. Novel paradigms on scorpion toxins that affects the activating mechanism of sodium channels. Toxicon 49 (2), 171–180.

Rodríguez de la Vega, R.C., Barraza, R., Restano, R., Possani, L.D., 2009. Toxins active on Herg channel: structure and function. In: de Lima, M. E. (Ed.), Animal Toxins: State of the Art. Perspectives in Health and Biotechnology. Editora UFMG. Belo Horizonte, Brazil, pp. 193–204.

Rodríguez de la Vega, R.C., Schwartz, E.F., Possani, L.D., 2010. Mining on scorpion venom biodiversity. Toxicon 56, 1155–1161.

Rodríguez, R.R., Coronas, F.I.V., Zamudio, F.Z., Batista, C.V.F., (in preparation). Proteomic analysis of venom from the scorpion Rhopalurus junceus collected in different geographical areas of Cuba.

Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19 (12), 1572–1574.

Selisko, B., Garcia, C., Becerril, B., Delepierre, M., Possani, L.D., 1996. An insect-specific toxin from *Centruroides noxius* Hoffmann cDNA, primary structure, three-dimensional model and electrostatic surface potentials in comparison with other toxin variants. Eur. J. Biochem. 242, 235–242.

Tolksdorf, S., McCready, M., McCullagh, D., Schwenk, E., 1949. The turbidimetric assay of hyaluronidase. J. Lab. Clin. Med. 34, 74–89.