03088
3
2ej

UNIVERSIDAD NACIONAL AUTONOMA
DE MEXICO

UNIDAD ACADEMICA DE LOS CICLOS PROFESIONAL Y DE
POSGRADO DEL COLEGIO DE CIENCIAS Y HUMANIDADES
INSTITUTO DE BIOTECNOLOGIA

DETERMINACION DE LA ESTRUCTURA PRIMARIA Y CARACTERIZACION DE cDNAs QUE CODIFICAN *PARA* TOXINAS DEL VENENO DEL ALACRAN *Centruroides noxius* HOFFMANN

T E S I S

QUE PARA OBTENER EL GRADO DE

DOCTOR EN BIOTECNOLOGIA

P R E S E N T A :

M. en C. ALEJANDRA VAZQUEZ RAMOS

CUERNAVACA, MORELOS 1993

TESIS CON
FALLA DE ORIGEN

TESIS CON
FALLA DE ORIGEN

Universidad Nacional
Autónoma de México

UNAM – Dirección General de Bibliotecas Tesis Digitales Restricciones de uso

DERECHOS RESERVADOS © PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis está protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

ESTE TRABAJO SE REALIZO EN EL LABORATORIO DEL DR. BOLIVAR, EN EL DEPARTAMENTO DE BIOLOGIA MOLECULAR Y EN EL LABORATORIO DEL DR. POSSANI, EN EL DEPARTAMENTO DE BIOQUIMICA DEL INSTITUTO DE BIOTECNOLOGIA, UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO BAJO LA DIRECCION DE LOS DOCTORES: BALTAZAR BECERRIL LUJAN Y LOURIVAL DOMINGOS POSSANI POSTAY

DURANTE EL DESARROLLO DE ESTE TRABAJO LA AUTORA FUE BECARIA DE LA

DIRECCION GENERAL DE ASUNTOS DEL PERSONAL ACADEMICO

ESTA TESIS FUE FINANCIADA PARCIALMENTE POR LOS PROYECTOS

HOWARD HUGHES MEDICAL INSTITUTE 75191-527104

DGAPA- UNAM IN202689 E IN300991

CONACyT-0018-N9105

INDICE

CONTENIDO		PAGINA
	RESUMEN	1
	SUMMARY	2
1.	INTRODUCCION	
1.1.	PRESENTACION	3
1.2.	ANTECEDENTES	3
1.3.	OBJETIVO	10
2.	PUBLICACIONES	11
2.1.	Vázquez *et al.* 1993. FEBS Lett. 320:43-46	12
2.2.	Becerril *et al.* 1993. Gene. (en prensa)	16
3.	DATOS NO PUBLICADOS	
3.1.	OBTENCION Y COMPARACION DE LA SECUENCIA DE NUCLEOTIDOS DE cDNAs	23
3.2.	COMPARACION DEL PEPTIDO SEÑAL Y DE LA REGION 3´ QUE NO SE TRADUCE	25
3.3.	ANALISIS Y COMPARACION DE LA SECUENCIA DE AMINOACIDOS DEDUCIDA DE LOS cDNAs CON LA SECUENCIA DE AMINOACIDOS REPORTADAS	27
4.	PERSPECTIVAS FUTURAS	30
5.	REFERENCIAS	31

RESUMEN

En esta tesis se reporta la estructura primaria de la toxina Cn4 del veneno del alacrán *Centruroides noxius* Hoffmann, la cual se ha visto que bloquea canales de sodio dependientes de voltaje. También se reporta la clonación y secuenciación del cDNA que la codifica, así como la de otros cDNAs que codifican para toxinas del veneno del alacrán aún no reportadas por secuencia directa de los péptidos. A la vez, se discute el posible mecanismo de procesamiento del péptido hasta su forma madura y se lleva a cabo un análisis comparativo de las secuencias obtenidas con las reportadas en la literatura. Dicho análisis muestra algunos motivos estructurales que son comunes a todas las toxinas del veneno de alacranes, tanto del nuevo mundo, como del viejo mundo.

SUMMARY

In this thesis, we report the primary structure of toxin Cn4 from the venom of the scorpion *Centruroides noxius* Hoffmann, which has been shown to be a voltage dependent Na^{+}-channel blocker. In addition, we report the cloning and nucleotide sequence determination of the cDNA that codes for this toxin and for other cDNAs encoding toxins of the scorpion venom which have not been reported by direct sequence of the peptides. We also discuss the mechanism for processing the peptide to its mature form and performed a comparative analysis of the sequences obtained with those reported in the literature. This analysis shows some structural motifs that are common to all toxins of the scorpion venoms, from both Old World and New World.

1. INTRODUCCION

1.1. PRESENTACION

Esta tesis está enfocada principalmente al estudio de las características de los cDNAs que codifican para toxinas del veneno del alacrán mexicano *Centruroides noxius* Hoffmann, las cuales reconocen canales de sodio de membranas de células plasmáticas excitables.

La tesis se encuentra dividida en cuatro partes: en la primera se revisan brevemente los antecedentes que llevaron a la hipótesis de trabajo que se siguió durante el desarrollo de la tesis; la segunda parte contiene los artículos publicados con algunos de los datos experimentales obtenidos. En la tercera parte se mencionan algunos datos complementarios, aún no publicados; y en la última parte se presenta una visión global de las perspectivas futuras de este trabajo.

1.2. ANTECEDENTES

El efecto letal y tóxico del veneno de los alacranes se debe a la presencia de péptidos que actúan específicamente sobre canales iónicos dependientes de voltaje. De acuerdo a su especificidad de unión se han dividido en dos grupos: el primero

conformado por las toxinas que actúan específicamente sobre canales de potasio, descritas inicialmente por nuestro grupo de trabajo (Carbone *et al.*, 1982; Possani *et al.*, 1982) y posteriormente por otros grupos (Gimenez-Gallego *et al.*, 1988); y el otro grupo constituido por las toxinas que actúan específicamente sobre canales de sodio (Caterall, 1976; Couraud *et al.*, 1982).

Existen diferencias estructurales entre las toxinas que bloquean los canales de potasio o de sodio, sin encontrarse similitud importante en la estructura primaria de ambos grupos de toxinas. Las que actúan sobre canales de potasio son mas pequeñas, contienen de 38-39 resíduos de aminoácidos y su estructura se encuentra estabilizada por tres puentes disulfuro. Las que afectan a canales de sodio son mas grandes (60-70 resíduos) y presentan cuatro puentes disulfuro.

Durante los últimos años, se han realizado una serie de estudios enfocados hacia la purificación y caracterización de numerosas toxinas del veneno de alacranes (ver revisiones de Rochat *et al.*, 1979 y Possani, 1984). Las toxinas que actúan sobre canales de sodio se han dividido a su vez en tres grupos de acuerdo con su especificidad de acción en: toxinas contra mamíferos (Possani and Dent, 1981; Possani *et al.*, 1985), toxinas contra insectos (Babin *et al.*, 1974; Kopeyan *et al.*, 1990; Pelhate *et al.*, 1982) y toxinas contra crustáceos (Zlotkin *et al.*, 1975 y 1978).

Asimismo, se ha descrito que dentro de las toxinas que afectan canales de sodio existen dos grupos funcionales de acuerdo a su farmacología y sus propiedades de

unión: **toxinas tipo** α que se han encontrado en el veneno de alacranes de la subfamilia Buthinae (alacranes del viejo mundo) y afectan principalmente el mecanismo de inactivación del canal (Noner, 1979); y las **toxinas tipo ß** las cuales se encuentran en el veneno de alacranes de la subfamilia Centrurinae (alacranes del nuevo mundo) y que alteran el mecanismo de activación (Meves *et al.*, 1986). También se ha comprobado que ambos tipos de toxinas se unen a sitios distintos en el canal (Jover *et al.*, 1980; Couraud *et al.*, 1982). Sin embargo, se ha encontrado que existen venenos de alacranes que contienen toxinas con actividad α y toxinas con actividad β en el mismo veneno (Kirsch *et al.*, 1989).

Al comparar la estructura primaria de las toxinas que afectan a canales de sodio (Fig. 1), se observa que existe una gran similitud en la secuencia de aminoácidos en todas ellas. En particular cuando se alinean con respecto a los resíduos de cisteína, se ve claramente que estos 8 resíduos se encuentran en posiciones equivalentes en todas las secuencias comparadas (12,16,25,29,41,51,53 y 65 con respecto a la toxina Cn4). Es importante hacer notar que además tienen conservada la posición de los resíduos que son importantes para estabilizar la estructura secundaria, regiones de α-hélice (a) y ß plegada(b). La mayoría de los resíduos donde se permite variabilidad, deleciones y/o inserciones representadas por guiones en la Fig. 1, están en posiciones que corresponden a estructuras al azar o a asas dentro de la molécula. En la sección 3.3. se presenta una discusión mas amplia sobre esta figura.

```
TOXINA     SECUENCIA DE AMINOACIDOS                                                          REF.

                                      *****                 ***
Cn4       KEGYLVNSYTGCKYECFKLGDNDYCLRECKQQYGKGAGGYCYAF-----GCWCTHLYEQAVVWPL--KNKT-C-N          a
CngtV     KEGYLVNSYTGCKYECFKLGDNDYCLRECKQQYGKGAGGYCYAF-----GCWCTHLYEQAVVWPL--KNKT-C-NGK        a
Clt1      KEGYLVNHSTGCKYECFKLGDNDYCLRECRQQYGKGAGGYCYAF-----GCWCTHLYEQAVVWPL-P-NKT-CS           b
Cn3       KEGYLVELGTGCKYECFKLGDNDYCLRECKARYGKGAGGYCYAF-----GCWCTQLYEQAVVWPL--KNKT-CR           b
CssII     KEGYLVSKSTGCKYECLKLGDNDYCLRECKQQYGKSSGGYCYAF-----ACWCTHLYEQAVVWPL-P-NKT-CN           c
Cn2       KEGYLVDKNTGCKYECLKLGDNDYCLRECKQQGYKGAGGYCYAF-----ACWCTHLYEQAIVWPL-P-NKR-CS           b
CssIII    KEGYLVSKSTGCKYECLKLGDNDYCLRECKQQYGKSSGGYCYAF-----ACWCEALPDHTQVW--VP-NK--CT           c
CsEv2     KEGYLVNKSTGCKYGCLKLGENEGCDKECKAKNQGGSYGYCYAF-----ACWCEGLPESTPTYPL-P-NK--CSS          c
CngtII    KEGYLVNKSTGCKYGCLLLGKNEGCDKECKAKNQGGSYGYCYAF-----GCWCQGLPESTPTYPL-P-NKS-CSKK         d
CsEv1     KEGYLVKKSDGCKYDCFWLGKNEHCNTECKAKNQGGSYGYCYAF-----ACWCEGLPESTPTYPL-P-NK--CS           c
CngtIII   KEGYLVNKSTGCKYGCFWLGKNEACDKECKAKNQGGSYGYCYAF-----GCWCEVLPESTPTYPL-P-NKT-CSKK         d
CsEv3     KEGYLVKKSDGCKYGCLKLGENEGCDTECKAKNQGGSYGYCYAF-----ACWCEGLPESTPTYPL-P-NKS-C            c
Cn1       KDGYLVDA-KGCKKNCYKLGKNDYCNRECRMKHRGGSYGYCYGF-----GCYCEGLSDSTPTWPL--TNKT-C            b
CngtIV    KDGYLVDV-KGCKKNCYKLGENDYCNRECKMKHRGGSYGYCYGF-----GCYCEGLSDSTPTWPL-P-NKR-CGGK         d
CsEI      KDGYLVEK-TGCKKTCYKLGENDFCNRECKWKHIGGSYGYCYGF-----GCYCEGLPDSTQTWPL-P-NK--CT           c
Ts10      KEGYLMDHE-GCKLSCF-IRPSGYCGRECGIK-KGSS-GYCAWP-----ACYCYGLPNWVKVWDRA-TNK--C            b
Ts8       KEGYAMDHE-GCKFSCF-IRPAGFCDGYCKTHLKASS-GYCAWP-----ACYCYGVPDHIKVWDYA-TNK--C            b

AaHI     KRDGYIVYPNN-CVYHCVPP-----CDGLCKKN-GGSS-GSCSFLVPSGLACWCKDLPDNVPIKDT--SRK--CTRK         c
AaHIII   VRDGYIVNSKN-CVYHCVPP-----CDGLCKKN-GAKS-GSCGFLIPSGLACWCVALPDNVPIKDP--SYK--CHSR         c
LqhαIT   VRDAYIAKNYN-CVYECFR---DAYCNELCTKN-GASS-GYCQWAGKYGNACWCYALPDNVPIRVP---GK--CHRK         e
AaHII    VKDGYIVDDVN-CTYFCGR---NAYCNEECTKL-KGES-GYCQWASPYGNACYCYKLPDHVRTKGP---GR--CH           c
BotXI    LKDGYIVDDRN-CTFFCGR---NAYCNDECKKK-GGES-GYCQWASPYGNACWCYKLPDHVRTVQA---GR--CN           c
BotI    G-RDAYIAQPEN-CVYECAQ---NSYCNDLCTKQ-GATS-GYCDWLGKYGNACWCKDLPDNVPIKIP---GK--CHF         c
BotII   G-RDAYIADDKN-CVYTCGS---NSYCNTECTKN-GAES-GYCQWLGKYGNACWCIKLPDKVPIRIE---GK--CHF         c
LqqIV   GVRDAYIADDKN-CVYTCGS---NSYCNTECTKN-GAES-GYCQWLGKYGNACWCIKLPDKVPIRIP---GK--CR          f
LqqV     LKDGYIVDDKN-CTFFCGR---NAYCNDECKKK-GGES-GYCQWASPYGNACWCYKLPDRVSIKEK---GR--CN           f
LqhIT2     DGYIKRR-DGCKVACLI-G-NEGCDKECKAY-GG-SYGYC-WT--WGLACWCEGLPDDKTWKSE--TN--TCG          f
LqqIT2     DGYIRKR-DGCKLSCLF-G-NEGCNKECKSY-GG-SYGYC-WT--WGLACWCEGLPDEKTWKSE--TN--TCG          f
BjIT2      DGYIRKK-DGCKVSCII-G-NEGCNKECVAH-GG-SFGYC-WT--WGLACWCENLPDAVTWKSS--TN--TCG          f
AaIT4     EHGYLLNKYTGCKVWCVI--NNEECGYLCNKRRGG-YYGYCYF---WKLACYCQGARKSELWNY--KTNK--CDL          f
AaHIT    KKNGYAVD-SSGKAPECLL--SN-YCYNECTKVHYA-DKGYCCLLS-----CYCFGLNDDKKVLEISDTRKSYCDTTIIN     f
LqqIT1   KKNGYAVD-SSGKAPECLL--SN-YCYNECTKVHYA-DKGYCCLLS-----CYCFGLNDDKKVLEISDARKKYCDFVTIN     f
          bbbb                    aaaaaaaa       bbbbb        bbbb
                        -asaJ-                            -asaB--
```

Figura 1.

Comparación de la secuencia de aminoácidos de diferentes toxinas del veneno de alacrán que afectan canales de sodio. El alineamiento se realizó en base a los resíduos de cisteína de la toxina Cn4. Se introdujeron guiones (-) para maximizar la similitud. El primer bloque incluye toxinas de alacranes del nuevo mundo. El segundo bloque contiene toxinas de alacranes del viejo mundo. Se indican las regiones de α-hélice (a) y ß plegada (b). Los asteriscos (*) indican el motivo estructural propuesto por Kobayashi y colaboradores (1991).

Referencias: a) Vázquez *et al.*, 1993; b) Zamudio *et al.*, 1992; c) Meves *et al.*, 1984; d) Becerril *et al.*, 1993; e) Eitan *et al.*, 1990; f) Zlotkin *et al.*, 1991.

El estudio de la estructura secundaria y terciaria de las toxinas del veneno de los alacranes se ha llevado a cabo empleando técnicas como: difracción de rayos X (Almassy *et al.*, 1983; Fontecilla-Camps *et al.*, 1988), resonancia magnética nuclear (Darbon and Braun, 1991) y dicroismo circular (Loret *et al.*, 1990) principalmente. La combinación de estas técnicas ha permitido determinar los principales motivos estructurales de las toxinas, llegando a tener ya descrita la estructura terciaria de tres toxinas: dos correspondientes a un alacrán del viejo mundo, la toxina II y la toxina de insecto de *Androctonus australis* Hector (Fontecilla- Camps *et al.*, 1988; Darbon *et al.*, 1991), y la otra de un alacrán del viejo mundo, la variante 3 de *C. sculpturatus* Ewing (Almassy *et al.*, 1983). En todos los casos, se presenta un fragmento de α-hélice, tres láminas antiparalelas ß plegadas y un centro hidrofóbico estabilizado principalmente por tres de los cuatro puentes disulfuro (Fig. 2). En adición, este motivo estructural se encuentra presente en todas las toxinas de alacrán, independientemente de su tamaño, secuencia o función, así como en otro tipo de péptidos no relacionados (Bontems *et al.*, 1991; Ménez *et al.*, 1992; Bruix *et al.*, 1993). Se cree que todas las toxinas de alacrán constituyen una familia de péptidos similares que evolucionaron a partir de un ancestro común, por duplicación genética y evolución independiente de los genes duplicados.

Figura2.
Estructura tridimensional de la toxina Cn2 del veneno del alacrán *C. noxius* (datos no publicados del laboratorio de Dr. Possani).

La mayoría de los grupos de investigación en el área, se han abocado a la caracterización bioquímica de las toxinas del veneno de alacranes de diferentes géneros y especies. No obstante, el conocimiento que a la fecha existe con respecto a los aspectos de biología molecular es limitado, aún se desconoce el número y la organización de los genes que codifican para las toxinas, y tampoco se sabe como está organizada o como es procesada la información genética en la glándula venenosa. A la fecha existen solo 6 reportes en donde se da la secuencia de cDNAs que codifican para toxinas del veneno de alacranes, 3 de ellos pertenecientes a alacranes del viejo mundo (Bougis *et al.*, 1989; Zilberberg *et al.*, 1991; Gurevitz *et al.*, 1991) y otros 3 del nuevo mundo (Martín-Eauclaire *et al.*, 1992; este trabajo de tesis: Becerril *et al.*,1993 y Vázquez *et al.*, 1993).

El uso de las metodologiás de DNA recombinante, como mutagénesis dirigida para poder obtener variantes que pudieran diferir en su especificidad de acción (mamíferos, insectos o crustáceos), o en su farmacología (α o β), con base en la información que se tiene hasta el momento de los motivos estructurales fundamentales para la función de las toxinas, combinado con la expresión de genes o fragmentos de genes de las toxinas nativas y/o modificadas, permitiría poder contestar una serie de preguntas relevantes como : 1) ¿Cuales son las características conservadas de la estructura primaria, de las toxinas del veneno de los alacranes, que determinan su actividad sobre membranas excitables? 2) ¿Por qué las regiones variables posiblemente son las que definen el control fino de la acción de las toxinas, α o β? 3) ¿Que mutaciones puntuales generarían toxinas modificadas que permitan

estudiar la relación estructura-función? 4) ¿Se puede llegar a proponer la síntesis de un péptido inmunogénico que sirva como vacuna? Conociendo la respuesta a las preguntas anteriores; 5) ¿Se puede llegar a diseñar un nuevo fármaco?

Para poder llegar a contestar éstas y muchas otras preguntas, hace falta más información referente a la parte de biología molecular de las toxinas del veneno del alacrán. Por lo anterior, nuestro grupo de investigación está interesado en conocer la organización y expresión de los genes que codifican para las toxinas del veneno de alacranes mexicanos. El primer paso en la generación de conocimiento en este campo, es el aislamiento y la caracterización de diferentes clonas de cDNA del veneno del alacrán mexicano *C. noxius* Hoffmann.

1.3. OBJETIVO

El objetivo de esta tesis es: determinar la estructura primaria de la toxina Cn4, clonar y secuenciar el cDNA que la codifica, así como llevar a cabo la caracterización de diferentes cDNAs que codifiquen para toxinas del veneno del alacrán *Centruroides noxius* Hoffmann.

2. PUBLICACIONES

2.1.

Vázquez, A., Becerril, B., Martin, B. M., Zamudio, F., Bolivar, F. and Possani, L. D. (1993). Primary structure determination and cloning of the cDNA encoding toxin 4 of the scorpion *Centruroides noxius* Hoffmann. FEBS Lett. 320:43-46.

2.2.

Becerril, B., Vázquez, A., García, C., Corona, M., Bolivar, F. and Possani, L. D. (1993). Cloning and characterization of cDNAs that code for Na+-channel blocking toxins of the scorpio *Centruroides noxius* Hoffmann. Gene. (en prensa).

Volume 320, number 1, 43-46 FEBS 12270 March 1993
© 1993 Federation of European Biochemical Societies 00145793/93/$6.00

Primary structure determination and cloning of the cDNA encoding toxin 4 of the scorpion *Centruroides noxius* Hoffmann

A. Vázquez[a], B. Becerril[a], B.M. Martin[b], F. Zamudio[a], F. Bolivar[a] and L.D. Possani[a]

[a]*Instituto de Biotecnología, Universidad Nacional Autónoma de México, Av. Universidad 2001, Cuernavaca 62271, México* and [b]*National Institute of Mental Health, Molecular Neurogenetics Unit, Clinical Neuroscience Branch, Building 10 3N256 Bethesda, MD 20892, USA*

Received 15 February 1993

A peptide (toxin II-10), shown to be a Na^+ channel blocker, was purified from the venom of the scorpion *Centruroides noxius* Hoffmann and sequenced by Edman degradation. It has 66 amino acid residues with the C-terminal residue (asparagine) amidated, as demonstrated by mass spectrometry. In addition, we report the cloning and the nucleotide sequence of the cDNA (CngtV) that codes for this toxin. We discuss the mechanism for processing the precursor peptide to its final form and compare the primary structure to that of other Na^+ channel toxins. Two distinct groups of toxins seem to emerge from this comparison, suggesting a structure-function relationship of these peptides towards the recognition of either mammalian or insect tissues.

Scorpion toxin; Na^+channel; cDNA clone; Nucleotide sequence; Peptide processing

1. INTRODUCTION

Scorpion toxins are exquisite tools for the study of ion channels [1]. At least two classes of different families of peptides have been purified and characterized from the venom of the scorpions of the family, Buthidae [2]. Among the toxins specific for the Na^+ channel is toxin II-10 from the Mexican scorpion *Centruroides noxius* Hoffmann. Initially, this peptide was purified by Sephadex G-50 gel filtration followed by ion-exchange chromatography on a carboxymethyl-cellulose column [2,3]. Its effect on Na^+ channels of excitable membranes was documented in squid axon membranes [4,5], and brain synaptosomes [6]. In this communication we describe the determination of its complete amino acid sequence, by automatic Edman degradation [7] of reduced and alkylated toxin, and of its peptides purified by high performance liquid cromatography (HPLC). We also describe the cloning of the cDNA encoding this toxin (CngtV), and discuss a possible mechanism for its post-translational processing.

2. MATERIALS AND METHODS

2.1. *Purification of the toxin Cn4.*

Venom from the scorpion *C. noxius* was obtained in the laboratory by electrical stimulation [3] and toxin Cn4 was purified by the same methods as previously described [3], except for the addition of an HPLC step at the end of the purification procedure, using published protocols [8].

2.2. *Amino acid sequence determination*

Highly purified toxin was reduced and carboxymethylated (RC-toxin) as reviewed [2], prior to sequencing on a Millipore 6600 ProSequencer. RC-toxin Cn4 was digested (66 nmol) using protease V8 from *Staphylococcus aureus* (Boehringer-Mannheim). Another sample (30 nmol) was cleaved with trypsin (TPCK treated) from Sigma Co. (St. Louis, MO). Both digests were separated by HPLC. Several peptides obtained by this procedure were sequenced in order to obtain the primary structure of this toxin. The C-terminal peptide was also analysed by mass spectrometry (data not shown) in order to confirm the amidation of the last residue. Only analytical grade reagents were used, as earlier described [8].

2.3. *Construction and screening of a* C. noxius *cDNA library*

Isolation of total RNA, purification of the poly(A)$^+$ RNA, synthesis and construction of the cDNA library in λgt11 were performed according to the instructions supplied in the commercial kits utilized (Amersham, RPN.1264, RPN.1511, RPN 1256Y/Z, respectively). The screening of the library, and the conditions for pre-hybridization and hybridization were performed as described [9].

3. RESULTS AND DISCUSSION

3.1. *Amino acid sequence determination of toxin Cn4*

Fig. 1 shows the complete amino acid sequence of toxin II-10, which we propose to name Cn4 (fourth toxin completely sequenced from *C. noxius*), following the nomenclature proposed in our recent publication [10]. Direct automatic Edman degradation confirmed the sequence of the N-terminal region of toxin Cn4, as previously published by our group [3,5], permitting unequivocally the identification of the first 45 residues. Tryptic digestion of RC-toxin produced at least 16 pep-

Correspondence address: L.D. Possani, Instituto de Biotecnología-UNAM, Av. Universidad 2001, Apto. Postal 510-3, Cuernavaca 62271, Mexico. Fax: (52) (73) 172 388.

Published by Elsevier Science Publishers B.V.

```
1                            10                           20
Lys-Glu-Gly-Tyr-Leu-Val-Asn-Ser-Tyr-Thr-Gly-Cys-Lys-Tyr-Glu-Cys-Phe-Lys-Leu-Gly-
:...............................>.D...................................>..

21                           30                           40
Asp-Asn-Asp-Tyr-Cys-Leu-Arg-Glu-Cys-Lys-Gln-Gln-Tyr-Gly-Lys-Gly-Ala-Gly-Gly-Tyr-
D...............................>.D...................................>..
                                                          :+++++Tryp+++++>++

41                           50                           60
Cys-Tyr-Ala-Phe-Gly-Cys-Trp-Cys-Thr-His-Leu-Tyr-Glu-Gln-Ala-Val-Val-Trp-Pro-Leu-
D.................>                            :------V8--------------->--
Tryp +++++++++++++++++++++++++++++++Tryp+++++++++++++++++++++++++++++>

61             66
Lys-Asn-Lys-Thr-Cys-Asn(NH2)
V8---------------->  Ms-Aaa
```

Fig. 1. Complete amino acid sequence of toxin Cn4. Results of direct Edman degradation (shown with D separated by points), allowed us to identify the first 45 residues. Tryptic digestion (labeled Tryp separated by crosses) indicates the amino acid sequence corresponding to the peptide eluted at 48.35 in the HPLC. This sequence starts with Gly at position 36 and goes to Trp at position 58. Protease V8 digestion (labeled with V8 separated by dashes) corresponds to the C-terminal region, with amino acid sequence corresponding to Gln at position 54 to the end of the sequence. The last residue (asparagine amide) was positioned based on Ms (mass spectrometry) and Aaa (amino acid analysis).

tides (data not shown), from which the peptide eluting at 48.35 min corresponded to the intermediate overlapping segment of the primary structure. HPLC separation of protease V8-cleaved RC-toxin (data not shown) produced at least 13 different peptides, from which the peptide eluting at 35.09 min was shown to correspond to the C-terminal region of the toxin (Fig. 1). The C-terminal V8 peptide was covalently attached to Sequelon AA (Millipore Corp.) following the manufacturer's instructions, and subjected to N-terminal sequence analysis on a Milligen 6600 ProSequencer using TFA 100 cycles. Amino acid residues 54–65 were unambiguously identified. In order to confirm that the entire peptide was sequenced, a second aliquot was subjected to amino acid analysis. The amino acid composition revealed the presence of a second aspartic acid residue (either Asp or Asn) not detected in the sequence. Our failure to identify the C-terminal residue would suggest that it was not attached to the aryl amide membrane. One would expect that aspartic acid and asparagine to be covalently attached to the membrane and therefore be identified, albeit in low yield. Thus, we surmised that the missing residue was asparagine amide. The peptide was covalently attached to the membrane through the carboxyl group of carboxymethylated-cysteine (CM-cys) of position 65, an attachment site previously identified for other CM-cys containing peptides (B.M. Martin, unpublished). Mass spectral analysis of the peptide with electrospray ionization indicated a molecular mass of 1,557 ± 1, consistent with an asparagine amide at the C-terminal residue. These conclusions are further supported by the results of cloning and nucleotide sequencing of the cDNA encoding for this toxin, as discussed below.

3.2. *Isolation and sequencing of the cDNA*

Using duplicate filters to screen the cDNA library with a degenerate oligonucleotide probe designed from a conserved region of the *Centruroides* toxins, we detected several positive signals with different intensities. One of these clones was isolated, subcloned in pBluescript (Stratagene, La Jolla, CA, USA), and both strands sequenced (Fig. 2), using the Sequenase version 2 kit (USB, USA). The resulting amino acid sequence deduced from this clone corresponded to toxin Cn4, the complete amino acid sequence of which is reported in this work (Fig. 1). The nucleotide sequence shown in Fig. 2 displayed an open reading frame coding for a precursor of 86 amino acid residues with a signal peptide of 19 amino acid residues, the mature peptide of 66

```
GAAA ATG AAC TCG TTG TTG ATG ATC ACT GCT TGT TTG GCC CTG GTC GGA ACA GTG TGG GCA AAG   64
      M   N   S   L   L   M   I   T   A   C   L   A   L   V   G   T   V   W   A   K    20

     GAA GGT TAT CTG GTA AAC TCG TAC ACG GGC TGC AAA TAC GAA TGC TTT AAA TTG GGA GAC  124
      E   G   Y   L   V   N   S   Y   T   G   C   K   Y   E   C   F   K   L   G   D    40

     AAC GAT TAT TGC TTG AGG GAA TGC AAA CAG CAG TAC GGA AAA GGT GCT GGC GGC TAT TGT  184
      N   D   Y   C   L   R   E   C   K   Q   Q   Y   G   K   G   A   G   G   Y   C    60

     TAC GCT TTT GGG TGC TGG TGC ACA CAT TTG TAC GAA CAA GCG GTG GTC TGG CCC CTT AAA  244
      Y   A   F   G   C   W   C   T   H   L   Y   E   Q   A   V   V   W   P   L   K    80

     AAT AAA ACA TGC AAC GGA AAA TAA TCGCAACGACTTTTTATTGCCCACCAACAGAAATATTGTAACGCTTC  318
      N   K   T   C   N   G   K  End                                                   87

     TTAATTTCAATTAAATGAAATAAAATATTATACCTTTAAAAAAAAAA...   CagtV                       365
```

Fig. 2. Nucleotide sequence and deduced amino acid sequence of the cDNA encoding the Cn4 precursor. Signal peptide is underlined. Polyadenylation signal (AATAAA) is indicated in bold characters. This nucleotide sequence will appear in the GenBank nucleatide sequence Database under accession number L05063. Nucleotides and amino acids are numbered at the right side of the corresponding sequences.

GROUP	TOXIN	AMINO ACID SEQUENCE	a	b
I	Cn4	KEGYLVNSYTGCKYECFKLGDNDYCLRECKQQYGKGAGGYCYAFGCWCTHLYEQAVVWPLKNKTCN	100%	(+)
I	Cll1	KEGYLVNHSTGCKYECFKLGDNDYCLRECRQQYGKGAGGYCYAFGCWCTHLYEQAVVWPLPNKTCS	92%	(10)
I	Cn3	KEGYLVELGTGCKYECFKLGDNDYCLRECKARYGKGAGGYCYAFGCWCTQLYEQAVVWPLKNKTCR	89%	(10)
I	CssII	KEGYLVSKSTGCKYECLKLGDNDYCLRECKQQYGKSSGGYCYAFACWCTHLYEQAVVWPLPNKTCN	88%	(14)
I	Cn2	KEGYLVDKNTGCKYECLKLGDNDYCLRECKQQGYKGAGGYCYAFACWCTHLYEQAIVWPLPNKRCS	83%	(10)
I	CssIII	KEGYLVSKSTGCKYECLKLGDNDYCLRECKQQYGKSSGGYCYAFACWCEALPDHTQVW-VPNK-CT	71%	(14)
II	CsEv2	KEGYLVNKSTGCKYGCLKLGENEGCDKECKAKNQGGSYGYCYAFACWCEGLPESTPTYPLPNK-CSS	56%	(14)
II	CsEv1	KEGYLVKKSDGCKYDCFWLGKNEHCNTECKAKNQGGSYGYCYAFACWCEGLPESTPTYPLPNK-CS	56%	(14)
II	CsEv3	KEGYLVKKSDGCKYGCLKLGENEGCDTECKAKNQGGSYGYCYAFACWCEGLPESTPTYPLPNKSC	56%	(14)
II	Cn1	KDGYLVDA-KGCKKNCYKLGKNDYCNRECRMKHRGGSYGYCYGFGCYCEGLSDSTPTWPLTNKTC	56%	(10)
II	CsEI	KDGYLVEK-TGCKKTCYKLGENDFCNRECKWKHIGGSYGYCYGFGCYCEGLPDSTQTWPLPNK-CT	56%	(14)
		****** *** * ** * * ** **** * * * * ** *		

Fig. 3. Amino acid sequence comparison of several toxins from the genus *Centruroides* that affect Na^+ channels Gaps (-) were introduced for maximizing similarities. (a) Percent of similarity taking Cn4 as 100%. The gaps have been taken as differences for the calculus. (b) Reference: +this paper. Cll1, is toxin 1 from *C. limpidus tecomanus*. Cn1, Cn2, Cn3, Cn4 are toxins from *C. noxius* Hoffmann. CsEv1, CsEv2, CsEv3 and CsEI are toxins from *C. sculpturatus* Ewing. CssII and CssIII are toxins from *C. suffusus suffusus*. *Amino acid residues that are invariant in the same position in all the sequences.

residues and two additional residues at the C-terminus. These results support previous reports [11,12] which proposed that scorpion toxins are synthesized as precursors of the mature toxin. Two processing steps seem to occur for toxin Cn4: (i) elimination of the signal peptide, and (ii) processing of the C-terminal region. At the N-terminal region, the signal peptide is certainly excised by a signal peptidase, while at the C-terminal part of the peptide the additional two amino acid residues (Gly and Lys) are removed and the Gly residue donates its amino group for the α-amidation of asparagine, as has been described for the processing of the inactive precursors of secretory peptides that contain sites for proteolysis and α-amidation, including a toxic peptide from a cDNA library of *Tityus serrulatus* [13].

3.3. *Primary structure comparison*

The comparison of the amino acid sequence of toxins purified and sequenced from *Centruroides* scorpion venoms (Fig. 3) shows that they constitute a cluster of closely related peptides, sharing a high degree of similarity in their primary structures (at least 56%), as well as in their functional specificities through their recognition of Na^+ channels [4-6,14]. By comparing the different binding properties and electrophysiological data, the toxins affecting Na^+ channels were classified as: (i) α-toxins, that mainly act at the inactivation mechanism, and (ii) β-toxins that modify the activation mechanism of the Na^+ channel [14]. Toxin Cn4 has been shown by Carbone et al. [5] to affect the activation mechanism of Na^+ channels of squid axon, hence it belongs to the β-type toxins. It is very toxic to mice and rats, and binds to brain synaptosomes with an affinity of 2–10 nM [15]. Toxin Cn2 was poorly effective towards Na^+ channels of squid axon [5], but was shown to be very toxic to mice [10], binding to their brain synaptosomes with K_d's in the range of the low nanomoles [6,15]. Because it affects the activation mechanism, it is also a β-type toxin. The toxins from the venom of *C. sculpturatus* Ewing (CsEv1, 2 and 3) have the characteristic effects of α-type toxins and are very toxic to insects or crustaceans but have a small toxic effect in mammals [16].

Analysis of the sequences shown in Fig. 3 permits us to sort out several general features of the primary structure that can be correlated with the distinct pharmacological effects observed. Since the total number of amino acids and the positions of all cysteines are conserved it is very likely that the general folding of all toxic molecules are the same (the positions of 27 of the 66 amino acid residues are identical). The N-terminal region of all toxins is highly conserved. Larger differences are observed at the C-terminal part of the toxins, which provides support for a tentative classification into two distinct groups: group I, for which Cn4 is the prototype, is composed of peptides mainly toxic to mammals, displaying at least 70% similarity in the primary structure, with only two prolines in the last 15 amino acid residues of the sequence. Group II, consisting of peptides mainly toxic to insects and crustaceans, with about 56% primary structure similarity to Cn4, containing more than 2 proline residues at the C-terminal region, with CsEV2 as a representative toxin. Possibly, the structural constraints caused by the proline residues (α-imino acid) near the C-terminus could hamper an appropriate folding during interaction with the channel proteins, as suggested by others [16]. Final evaluation of this model, distinguishing the toxins based on their specificity for either mammalian or insect targets, depends upon development of more comprehensive structural data for the comparison of these motifs.

Acknowledgements: The Ph.D. scholarship obtained by A.V. from DGAPA-UNAM is greatly acknowledged. We express our gratitude to Dr. Mark Duncan of the Biomedical Mass Spectrometry Unit, University of New South Wales, Australia, for the mass spectral data, and Mr. Bill Eliason of NIMH for helpful discussions. This work was partially supported by grants from Howard Hughes Medical Institute (75191-527104), DGAPA-UNAM (IN202689 and IN300991) to L.D.P., and CONACyT-Mexico (0018-N9105) to L.D.P. and F.B.

REFERENCES

[1] Rochat, H. Bernard, P. and Couraud, F. (1979) In: Advances in Cytopharmacology, vol. 3 (Ceccarelli, B. and Clementi, F. eds.) pp. 325-334, Raven Press, NY.
[2] Possani, L.D. (1984) In: Handbook of Natural Toxins, vol. 3 (Tu, A.T. ed.) pp. 513-550, Marcel Dekker, NY.
[3] Possani, D.L., Dent, M.A.R., Martin, B.M., Maelicke, A. and Svendsen, I. (1981) Carlsberg Res. Commun. 46, 207-214.
[4] Carbone, E., Wanke, E., Prestipino, G. Possani, L.D. and Maelicke, A. (1981) Nature 296, 90-91.
[5] Carbone, E., Prestipino, G., Franciolini, F. Dent, M.A.R. and Possani, L.D. (1984) J. Physiol. (Paris) 79, 179-184.
[6] Sitges, M., Possani, L.D. and Bayon, A. (1987) J. Neurochem. 48, 1745-1752.
[7] Edman, P. and Begg, G. (1967) Eur J. Biochem. 1, 80-91.
[8] Possani, L.D., Martin, B.M., Svendsen, I., Rode, G.S. and Erickson, B.W. (1985) Biochem. J. 229, 739-750.
[9] Becerril, B., Vázquez, A., García, C. Corona, M., Bolivar, F. and Possani, L.D. (1993) Gene (in press).
[10] Zamudio, F.Z., Saavedra, R., Martin, B.M., Gurrola-Briones, G., Herion, P. and Possani, L.D. (1992) Eur. J. Biochem. 204, 281-192.
[11] Bougis, P.E., Rochat, H. and Smith, L.A. (1989) J. Biol. Chem. 264, 19259-19265.
[12] Gurevitz, M., Urbach, D., Zlotkin, E. and Zilberberg, N. (1990) Toxicon 29, 1270-1272.
[13] Bradbury, A., Finnie, M.D. A. and Smith, D.G. (1982) Nature 298, 686-688; see also Martin-Eauclaire, M.F., Ceard, B., Ribeiro, A.M., Diniz, C.R., Rochat, H. and Bougis, P.E. (1992) FEBS Lett. 302, 220-222.
[14] Meves, H., Simard, M.J. and Watt, D.D. (1984) J. Physiol. Paris 79, 185-191; same authors (1986) Ann. NY Acad. Sci. 479, 113-132.
[15] Ramirez, A.N., Gurrola, G.B., Valdivia, H.H. and Possani, L.D. (1991) Proceedings Xth World Congress on Animal, Plant and Microbial Toxins, Singapore. Abstract 301, p. 355.
[16] Loret, E.P., Sampieri, F., Roussel, A., Granier, C. and Rochat, H. (1990) Proteins 8, 164-172.

Gene, 000 (1993) 00–00
© 1993 Elsevier Science Publishers B.V. All rights reserved. 0378-1119/93/$06.00

GENE 07136

29 MAART 1993

Cloning and characterization of cDNAs that code for Na^{+}-channel-blocking toxins of the scorpion *Centruroides noxius* Hoffmann

(Recombinant DNA; λgt11 cDNA library; Southern blot genomic analysis; nucleotide sequence; polymerase chain reaction; toxin gene evolution; introns)

Baltazar Becerril, Alejandra Vázquez, Consuelo Garcia, Miguel Corona, Francisco Bolivar and Lourival D. Possani

Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Morelos 62271, Mexico

Received by L.K. Miller: 23 September 1992; Revised/Accepted: 19 January/20 January 1993; Received at publishers: 1 March 1993

SUMMARY

With the purpose of studying the organization and characteristics of the genes that code for toxins present in the venom of the Mexican scorpion, *Centruroides noxius* Hoffmann (*Cn*H), we prepared a λgt11 cDNA library from the venom glands. Using specific oligodeoxyribonucleotides (oligos) designed according to known amino acid (aa) sequences of *Cn*H toxins (STox), we detected several positive clones, determined their nucleotide (nt) sequences and deduced their aa sequences. A comparative analysis of these sequences with previously reported STox revealed that *Cn*H cDNAs code for a family of very similar STox. The cDNA coding for a known STox, II-10, was cloned. Additionally, three other complete (new) nt sequences were obtained for cDNAs encoding peptides similar to STox 1 from *Cn*H or variants 2 and 3 from *Centruroides sculpturatus* Ewing. Southern blot genomic DNA analysis showed a minimum size of approximately 600 bp as *Eco*RI fragments for elements of this family. PCR amplifications of *Cn*H genomic DNA and hybridization of PCR products with specific probes indicated that the genomic structural regions that code for these genes do not contain introns, or at least not large introns.

INTRODUCTION

Of the 134 species and sub-species of Mexican scorpions, only eight are dangerous to man (Dehesa-Dávila, 1989). Species which cause the most frequent cases of scorpion poisoning in Mexico belong to the *Centruroides* genus. They represent a serious health problem with more than 200 000 people stung per year and a mortality rate of 700–800 people per year (D. López- Acuña and A. Alagón, communicated during the technical session of the Sociedad Mexicana de Salud Pública, Mexico City, March 1979).

Scorpion venoms constitute a rich source of low-M_r peptides toxic to a variety of organisms including man (Miranda et al., 1970; reviews by Zlotkin et al., 1978; Possani, 1984). The most dangerous scorpion venoms so

Correspondence to: Dr. L.D. Possani, Instituto de Biotecnología, UNAM, Av. Universidad 2001, Apartado Postal 510-3, Cuernavaca, Morelos 62271, Mexico. Tel. (52-73) 172799; Fax (52-73) 172388; e-mail: possani@pbr322.ceingebi.unam.mx

Abbreviations: aa, amino acid(s); bp, base pair(s); cDNA, DNA complementary to RNA; *Clt*, *Centruroides limpidus tecomanus*; *Cngt*, cDNA from *Cn*H cloned into λgt11; *Cn*H, *Centruroides noxius* Hoffmann; *Cs*E, *Centruroides sculpturatus* Ewing; *Css*, *Centruroides suffusus suffusus*; ctsd DNA, calf thymus sonicated and denatured DNA; Denhardt's solution, 0.02% each of bovine serum albumin/ficoll/polyvinylpyrrolidone; EtdBr, ethidium bromide; kb, kilobase(s) or 1000 bp; N, any nucleoside; nt, nucleotide(s); oligo, oligodeoxyribonucleotide; PCR, polymerase chain reaction; SDS, sodium dodecyl sulfate; SSC, 0.15 M NaCl/0.015 M Na_3·citrate pH 7.6; SP, signal peptide(s); STox, scorpion toxin(s); T_m, melting temperature (temperature at which 50% of the hybrids formed between two nucleic acids are dissociated).

(right hand col. short)

far studied have been shown to contain two kinds of STox: long-chain polypeptides of 60–70 aa which block Na^+-channels of excitable cells (Catterall, 1977; Couraud et al., 1982) and short-chain peptides of 37–39 aa affecting K^+-channels (Carbone et al., 1982; Possani et al., 1982; Miller et al., 1985; Gimenez-Gallego et al., 1988; Strong et al., 1989). Recently, Valdivia et al. (1991) demonstrated the existence of another class of polypeptides in the scorpion venom that activates the Ca^{2+}-release channel of the sarcoplasmic reticulum.

Nevertheless, the most thoroughly studied STox are those that modify Na^+-channels and were classified as α- and β-STox (Couraud et al., 1982; Wheeler et al., 1983). This classification was proposed based on the existence of two different binding sites on the Na^+-channels for the α- and β-toxins (Jover et al., 1980; Couraud et al., 1982). α-STox modify mainly the inactivation mechanism of the Na^+-channels (Nonner, 1979), while β- STox preferentially alter the activation mechanism (reviewed by Meves et al., 1986; Strichartz et al., 1987; Thomsen and Catterall, 1989).

Our group has contributed during the last decade to the knowledge of the structure-function relationship of the toxins from the venom of Mexican scorpions (Possani, 1984; Zamudio et al., 1992). A few regions involved in toxicity or antigenicity have been mapped in these STox using synthetic peptide and monoclonal antibody approaches (Gurrola et al., 1989; Possani et al., 1991). The possibility of using modern molecular biology strategies for cloning, directed mutagenesis and expression of specific STox-encoding genes or gene fragments in order to use the purified products to verify their interactions with ion channels has greatly motivated our work.

The aim of the present study was the isolation and characterization of several cDNAs from the Mexican scorpion *Centruroides noxius* Hoffmann (*Cn*H) and an initial characterization of the genomic region encoding these cDNAs.

RESULTS AND DISCUSSION

(a) Isolation of cDNA clones encoding STox peptides

A degenerate oligo probe coding for aa 42–48 of STox Cn2, which are well-conserved among *Centruroides* toxins, was used to screen a λgt11 cDNA library prepared from *Cn*H venom glands (legend to Fig. 1 and Fig. 2). From this screening we detected several positive signals with different intensities; we decided to characterize first the clone that gave the strongest signal. This clone was named *CngtI* and had an insert of approximately 230 bp. This insert was subcloned into the *Eco*RI sites of M13mp18 and M13mp19, and its nt sequence was determined. When the aa sequence from this cDNA was deduced, we found the target sequence (aa 42- 48 of Cn2); however, the entire sequence did not correspond to Cn2. *CngtI* was a truncated clone that contained the sequence for a (new) Na^+-channel-blocking STox cDNA coding from aa 25 through the polyadenylation site. A comparative analysis of the deduced aa sequence of clone *CngtI* with reported STox aa primary sequences revealed that from 41 aa encoded by this clone, 40 aa were identical with the STox variant 3 of *C. sculpturatus* (see Fig. 2), whose three-dimensional structure is known (Fontecilla-Camps et al., 1980). To explore the possibility of isolating a complete cDNA that encoded a STox closely related to *Cs*E variant 3, we decided to use the *CngtI* insert as a probe to screen the cDNA library under conditions of high stringency in order to isolate clones highly homologous to the *CngtI* insert. From this second screening, we isolated several positive clones.

The next step consisted of probing Southern blots of digested DNA of the isolated positive clones with a variant 3-specific oligo (designed from aa 10–16; see Figs. 1 and 2). Two positive inserts with a size adequate to encode a complete STox sequence were subcloned into the *Eco*RV site of pBluescript (pKS, Stratagene, La Jolla, CA). They were called *Cngt*II and *Cngt*III.

From this second screening, two less strongly positive clones were isolated, subcloned and sequenced. They were called *Cngt*IV and *Cngt*V. The nt sequences of clones *Cngt*II through *Cngt*V are shown in Fig. 1. These cDNAs are about 350 bp in length and encode STox precursors of 86–87 aa.

It has been proposed, based on the analysis of their cDNA sequences, that STox might be synthesized as precursors with a signal peptide (SP) of 18–19 aa (Bougis et al., 1989; Gurevitz et al., 1991). *Cn*H SP sequences met all the requirements for eukaryotes as proposed by von Heijne (1986).

In the case of STox that affect mammalian Na^+-channels (Bougis et al., 1989) and in the case of a STox that affects insect Na^+-channels (Gurevitz et al., 1991), it has also been observed that they have extensions at their C-terminal ends: Arg, Gly-Arg or Arg-Lys. These basic residues are not present in the mature peptides and when Gly precedes a basic residue, the residue becomes amidated at the C terminus (Bougis et al., 1989).

Of the STox coded by the cDNAs shown in Fig. 1, only *Cngt*V was thoroughly studied. It corresponded to STox II-10 (Possani et al., 1981), a Na^+-channel effector (Carbone et al., 1982; 1984), whose complete primary aa sequence has been determined (A.V., B.B., B. Martin, F. Zamudio, F.B. and L.D.P., manuscript in preparation). As shown in Fig. 2, the deduced aa sequences of these cDNAs are closely related to *Centruroides* STox that have

```
    -60       -50       -40       -30       -20       -10        1        10        20        30        40        50        60
CTTCGAAGATGAACTCGTTGTTGATGATCACTGCTTGTTTGTTCCTGATCGGAACAGTGTGGGCAAAAGAAGGTTATCTGGTAAACAAGAGCACGGGCTGCAAATACGGTTGCCTGCTATTGGGGAAA
     1 M  N  S  L  L  M  I  T  A  C  L  F  L  I  G  T  V  W  A  K  E  G  Y  L  V  N  K  S  T  G  C  K  Y  G  C  L  L  L  G  K
    GATGAACTCGTTGTTGATGATCACTGCTTGTTTGGTCCTATTCGGAACAGTGTGGGCAAAAGAAGGTTATCTGGTAAACAAGAGCACTGGCTGCAAATACGGTTGCTTCTGGTTGGGAAAA
       M  N  S  L  L  M  I  T  A  C  L  V  L  F  G  T  V  W  A  K  E  G  Y  L  V  N  K  S  T  G  C  K  Y  G  C  F  W  L  G  K
       TGAATTCGTTGTTGATCATCACTGCTTGTTTGGTCTTGATCGGAACAGTGTGGGCAAAGGACGGTTATCTGGTGGACGTA---AAGGGCTGCAAAAAGAATTGCTATAAATTGGGAGAA
       M  N  S  L  L  I  I  T  A  C  L  V  L  I  G  T  V  W  A  K  D  G  Y  L  V  D  V     K  G  C  K  K  N  C  Y  K  L  G  E
   GAAAATGAACTCGTTGTTGATGATCACTGCTTGTTTGGCCCTGGTCGGAACAGTGTGGGCAAAGGAAGGTTATCTGGTAAACTCGTACACGGGCTGCAAATACGAATGCTTTAAATTGGGAGAC
       M  N  S  L  L  M  I  T  A  C  L  A  L  V  G  T  V  W  A  K  E  G  Y  L  V  N  S  Y  T  G  C  K  Y  E  C  F  K  L  G  D

             70        80        90       100       110       120       130       140       150       160       170       180
CngtII     AACGAAGGCTGCGATAAGGAATGCAAAGCGAAGAACCAAGGAGGTAGTTACGGCTATTGCTACGCTTTTGGGTGCTGGTGCGAAGGTTTGCCCGAAAGTACACCGACTTATCCCCTTCCT
        41 N  E  G  C  D  K  E  C  K  A  K  N  Q  G  G  S  Y  G  Y  C  Y  A  F  G  C  W  C  E  G  L  P  E  S  T  P  T  Y  P  L  P
CngtIII    AACGAAGGCTGCGATAAGGAATGCAAAGCGAAGAACCAAGGAGGTAGTTACGGCTATTGCTACGCTTTTGGGTGCTGGTGCGAAGGTTTGCCCGAAAGTACACCGACTTATCCCCTTCCT
           N  E  G  C  D  K  E  C  K  A  K  N  Q  G  G  S  Y  G  Y  C  Y  A  F  G  C  W  C  E  G  L  P  E  S  T  P  T  Y  P  L  P
CngtIV     AACGATTATTGCAATAGGGAATGCAAAATGAAACACCGAGGAGGTAGTTACGGCTATTGCTACGGATTTGGGTGCTATTGTGAAGGATTATCCGATAGTACACCGACTTGGCCCCTTCCT
           N  D  Y  C  N  R  E  C  K  M  K  H  R  G  G  S  Y  G  Y  C  Y  G  F  G  C  Y  C  E  G  L  S  D  S  T  P  T  W  P  L  P
CngtV      AACGATTATTGCTTGAGGGAATGCAAACAGCAGTACGGAAAAGGTGCTGGCGGCTATTGTTACGCTTTTGGGTGCTGGTGCACACATTTGTACGAACAAGCGGTGGTCTGGCCCCTTAAA
           N  D  Y  C  L  R  E  C  K  Q  Q  Y  G  K  G  A  G  G  Y  C  Y  A  F  G  C  W  C  T  H  L  Y  E  Q  A  V  V  W  P  L  K

            190       200       210       220       230       240       250       260       270       280       290       300
II         AATAAATCTTGCAGCAAAAAATAATGGCAACGACTTTT-ATTGTTTACCAACAGAAATATTGTAACGCTTGTTAATTGCAGTTAAATGAAAAAAAAAAAAAAA
        81 N  K  S  C  S  K  K end 87
III        AATAAAACATGCAGCAAAAAATAATGGCAACGACTTTTTATTGTCCATAAACAGAAATATTGTAACGCTTCTTAATTGCAGTTAAATGAAATAAAATGCTAATAGCATTAAAAAAAAAA
           N  K  T  C  S  K  K end
IV         AATAAAAGATGCGGCGGAAAATAATGGCAACGACTTTTTATTGTCCACCAACAGAAATAGTGTAACGCTTCTTAATTCCAAGTAAAAAAA
           N  K  R  C  G  G  K end
V          AATAAAACATGCAACGGAAAATAATGGCAACGACTTTTTATTGCCCACCAACAGAAATATTGTAACGCTTCTTAATTTCAATTAAATGAAATAAAATA-TTATACCTTTAAAAAAAAAA..
           N  K  T  C  N  G  K end
```

Fig. 1. Nucleotide sequences of clones *Cngt*II–*Cngt*V and their deduced aa sequences. SP sequences are underlined. The overlined nt sequences [1–21 and 246–266 (complementary)] were used to synthesize PCR primers. Precursor peptides start at 1 (bold number) on the left side of *Cngt*II and finish with aa residue 87 (bold number) to the right of the *Cngt*II sequence. Polyadenylation signals (AATAAA) are doubly underlined. Methods: The screening of the cDNA library was carried out in two steps. In the first step, we analyzed approximately 30 000 plaques with a mixture of oligo probes designed from a conserved region of *Cn*H STox: nt 42–48 in Fig. 2 [5'-TA$^{T}_{C}$GCNTT$^{T}_{C}$GCNTG$^{T}_{C}$TGGTG, where N = A,G,C or T]. Probes were synthesized on a Mycrosyn 1450A DNA synthesizer by β-cyanoethyl phosphoroamidite chemistry and ^{32}P-end-labeled by T4 polynucleotide kinase. Replica filters were prehybridized for 2–8 h at 37°C in 6×SSC pH 7.6 containing 5×Denhardt's/0.1% SDS/100 μg per ml calf thymus sonicated and denatured DNA (ctsd DNA)/0.05% sodium pyrophosphate. Filters were then hybridized in 6×SSC pH 7.6/1×Denhardt's/100 μg per ml ctsd DNA/0.05% sodium pyrophosphate/^{32}P-end-labeled oligo probe for 12–16 h at 37°C. Successive washes were performed in 6×SSC pH 7.6/0.05% Na-pyrophosphate/0.1% SDS at 37°C for 5 min and once at 42°C for 5 min before autoradiography using X-AR film with intensifying screens for 18 h at −70°C. The second step of library screening consisted in the isolation of clones similar to the positive clone *Cngt*I identified in the first step (oligo screening). Positive clones were hybridized with an oligo probe designed from nt 10–16 of *C. sculpturatus* variant 3 (*Cs*Ev3) from Fig. 2 (5'-GA$^{T}_{C}$GGNTG$^{T}_{C}$AA$^{A}_{G}$T$^{T}_{C}$GGNTG). Prehybridization and hybridization conditions for the double-stranded probe (*Cngt*I insert) were similar to those just described except that incubation was at 42°C in the presence of 50% formamide. The pre- and hybridization conditions for *Cs*Ev3 oligos were the same as in the first step of library screening. Positive clone inserts were amplified by PCR using λgt11 forward (5'-GGTGGCGACGACTCCTGGAGCCCG) and reverse (5'-TTGACACCAGACCAACTGGTAATG) primers (New England Bio-Labs). These primers hybridize with the flanking regions of the λgt11 *Eco*RI cloning site. The annealing temperature depended on the T_m of the primers but typically was 50°C. The reactions were performed in a programmable heating chamber (Biosycler) using 30 rounds of temperature cycling (92°C for 1 min, 50°C for 1.5 min and 72°C for 3 min) followed by a final 10-min step at 72°C. As recommended by the manufacturer, we used 300 ng of each primer/500 ng of template DNA/2.5 units of Vent polymerase (New England Bio-Labs) in a final vol. of 100 μl reaction buffer. These PCR products were purified from gel, blunt-ended with T4 DNA polymerase and subcloned into the *Eco*RV site of pBluescript phagemid (Stratagene, La Jolla, CA). The ligation reaction was used to transform competent *E. coli* DH5-α cells. The subcloned DNA was sequenced using the Sequenase kit (US Biochemical, Cleveland, OH) on both strands. The λgt11 forward and reverse oligo primers were used for sequencing. Since the PCR reaction is known to be associated with a relatively high rate of nt misincorporation, we performed the PCR amplification experiments with Vent polymerase instead of *Taq* polymerase, and several independent clones were sequenced. The *Cngt*II-*Cngt*V nt sequences were deposited with the GenBank Nucleotide Sequence Database under accession Nos. L05060–L05063, respectively.

been characterized. As can be seen in Figs. 1 and 2, the four final C-terminal residues for STox II-10 (*Cngt*V) are Cys-Asn-Gly-Lys. It has been determined that the Asn residue is amidated at the C terminus in the mature II-10 STox (A.V., B.B., B. Martin, F. Zamudio, F.B. and L.D.P., manuscript in preparation). Now that the cDNA sequences encoding both North African (Bougis et al., 1989; Gurevitz et al., 1991) and North American (this study) STox are available, we can surmise that C-terminal STox processing follows the rules previously proposed by Bougis et al. (1989). Additional processing consists of the removal of the SP.

Mature peptides encoded by the cDNAs shown in Fig. 1 would be 66 aa residues in length for *Cngt*II, *Cngt*III and *Cngt*V and 65 aa residues for *Cngt*IV. Comparative analyses between these sequences reveal that *Cngt*II and *Cngt*III are 92% similar at the nt level, and the peptides encoded by these cDNAs are 91% similar. Also, *Cngt*IV shares 80% similarity with *Cngt*III at the nt level and 68% at the aa level. Similar results are obtained when *Cngt*V is compared. In spite of a lower similarity at the aa level between the peptides encoded by *Cngt*II (or *Cngt*III) and *Cngt*IV (or *Cngt*V), the similarity at the nt level (at least 80%) suggests that these

```
Group                 10        20        30        40        50        60
        Cn2      KEGYLVDKNTGCKYECLKLGDNDYCLRECKQQYGKGAGGYCYAFACWCTHLYEQAIVWPLPNKRCS
        CssII    KEGYLVSKSTGCKYECLKLGDNDYCLRECKQQYGKSSGGYCYAFACWCTHLYEQAVVWPLPNKTCN
        CssIII   KEGYLVSKSTGCKYECLKLGDNDYCLRECKQQYGKSSGGYCYAFACWCEALPDHTQVW-VPNK-CT
  1     Cll1     KEGYLVNHSTGCKYECFKLGDNDYCLRECRQQYGKGAGGYCYAFGCWCTHLYEQAVVWPLPNKTCS
        Cn3      KEGYLVELGTGCKYECFKLGDNDYCLRECKARYGKGAGGYCYAFGCWCTQLYEQAVVWPLKNKTCR
        CngtV    KEGYLVNSYTGCKYECFKLGDNDYCLRECKQQYGKGAGGYCYAFGCWCTHLYEQAVVWPLKNKTCNGK
        Cons.1   KXGYLVXXSTGCKYXCLKLGDNDYCLRECKQQYGKGAGGYCYAFACWCTHLYXQAVVWPLPNKTCS
                                                             G

                      10        20        30        40        50        60
        CsEv1    KEGYLVKKSDGCKYDCFWLGKNEHCNTECKAKNQGGSYGYCYAFACWCEGLPESTPTYPLPNK-CS
        CngtI                              CDKECKAKNQGGSYGYCYAFACWCEGLPESTPTLPLPNKSCGRK
        CngtIII  KEGYLVNKSTGCKYGCFWLGKNEGCDKECKAKNQGGSYGYCYAFGCWCEGLPESTPTYPLPNKTCSKK
  2     CsEv2    KEGYLVNKSTGCKYGCLKLGENEGCDKECKAKNQGGSYGYCYAFACWCEGLPESTPTYPLPNK-CSS
        CngtII   KEGYLVNKSTGCKYGCLLLGKNEGCDKECKAKNQGGSYGYCYAFGCWCEGLPESTPTYPLPNKSCSKK
        CsEv3    KEGYLVKKSDGCKYGCLKLGENEGCDTECKAKNQGGSYGYCYAFACWCEGLPESTPTYPLPNKSC
        Cons.2   KXGYLVNKSTGCKYGCLXLGXNXGCDKXCKAXNQGGSYGYCYAFACWCXGLPESTPTYPLPNKSCSKK

                      10        20        30        40        50        60
        CsEI     KDGYLVEK-TGCKKTCYKLGENDFCNRECKWKHIGGSYGYCYGFGCYCEGLPDSTQTWPLPNK-CT
        CngtIV   KDGYLVDV-KGCKKNCYKLGENDYCNRECKWKHRGGSYGYCYGFGCYCEGLSDSTPTWPLPNKRCGGK
  3     Cn1      KDGYLVDA-KGCKKNCYKLGKNDYCNRECRWKHRGGSYGYCYGFGCYCEGLSDSTPTWPLTNKTC
        Cons.3   KDGYLVDX-KGCKKNCYKLGENDYCNRECKWKHRGGSYGYCYGFGCYCEGLSDSTPTWPLPNKXCXGK
```

Fig. 2. The aa sequence comparison of principal representatives of *Centruroides* STox. Toxin sequences were grouped according to their similarities. Gaps (-) were introduced to maximize similarities. Below each similarity group, the consensus sequence is shown in bold. The aa sequences deduced from cDNAs are included. The *Cngt*V deduced primary sequence has been corroborated by direct aa sequence of the mature STox (A.V., B.B., B. Martin, F. Zamudio, F.B. and L.D.P., manuscript in preparation). Data are from this study, Meves et al. (1984) and Zamudio et al. (1992). In consensus sequences, X represents a variable residue. The aa are aligned with the last digits of the numerals.

cDNAs originate from mRNAs transcribed from closely related genes. Comparative studies with the North African *Androctonus australis* Hector STox cDNAs show that although the regions encoding mature peptides are more variable, the nucleotide sequences encoding SP and the 3′ non-coding regions are relatively well-conserved (data not shown). The cDNAs encoding STox I, I′, II and III of *A. australis* Hector (see Bougis et al., 1989) have a similar variation in the regions that code for the mature peptides, but those encoding the SP and the 3′ non-coding regions are also well-conserved. These results indicate that both scorpion species (*A. australis* Hector and *Cn*H) have followed a similar strategy to generate variation in their STox: gene duplications and independent evolution of the duplicated genes.

(b) Deduced aa sequences derived from cDNAs. Analysis and comparison with different *Centruroides* primary STox sequences

In order to compare the aa sequences derived from the cloned *Cn*H cDNAs, different *Centruroides* STox were grouped according to their similarities. Gaps were introduced where necessary to maximize similarities. The STox can be clustered into three groups as shown in Fig. 2. The sequences were aligned with respect to Cys residues. Each main group was aligned, and a consensus sequence was proposed for each one of them. *Cngt*V(II-10) is included in the first group, where *Cn*2 can be considered representative. *Cngt*II and *Cngt*III are included in the second group, where *Cs*E variants 2 and 3 are also presented. *Cngt*IV is included in the third group, where *Cn*1 can be considered representative. Some specific differences among the three groups can be observed, especially several deletions of 1 aa residue in different positions. Examples of such deletions or insertions can be deduced from cDNA sequences. The aa residue number 9 that is a Ser in CngtII and CngtIII is missing in CngtIV (Figs. 1 and 2). It is interesting to note, however, that this region in *Centruroides* toxins is very variable (Zamudio et al., 1992), indicating that its encoding DNA could be a hot spot for mutation.

(c) Southern blot genomic analysis

To determine the size of the genomic DNA region containing the genes encoding the peptides shown in Fig. 1, Southern blot hybridization analysis was carried out. When using *Cngt*I or *Cngt*II inserts as probes (Fig. 3), only two DNA fragments are clearly observed: one of approximately 18 kb when the genomic DNA was digested with *Bam*HI, *Hin*dIII or *Bam*HI+*Hin*dIII and a second segment of approximately 0.6 kb when the digestions were performed with *Eco*RI, *Eco*RI+*Bam*HI or *Eco*RI+*Hin*dIII. These results and the fact that *Cngt*II and *Cngt*III share 92% similarity at the nt level suggest that at least two different 0.6-kb *Eco*RI DNA segments should contain the genomic DNA regions encoding these two cDNAs (mRNAs). Since *Cngt*IV (or *Cngt*V) is 80% similar at the nt level to *Cngt*II (or *Cngt*III), and since the bands mentioned above are the only two bands present in the Southern blot experiment (18 kb and 0.6 kb), we propose that the genomic regions that specify these four cDNAs (mRNAs) are contained within these 0.6-kb *Eco*RI DNA fragments. The double digestions with *Eco*RI+*Bam*HI or *Eco*RI+*Hin*dIII suggest that the 18-

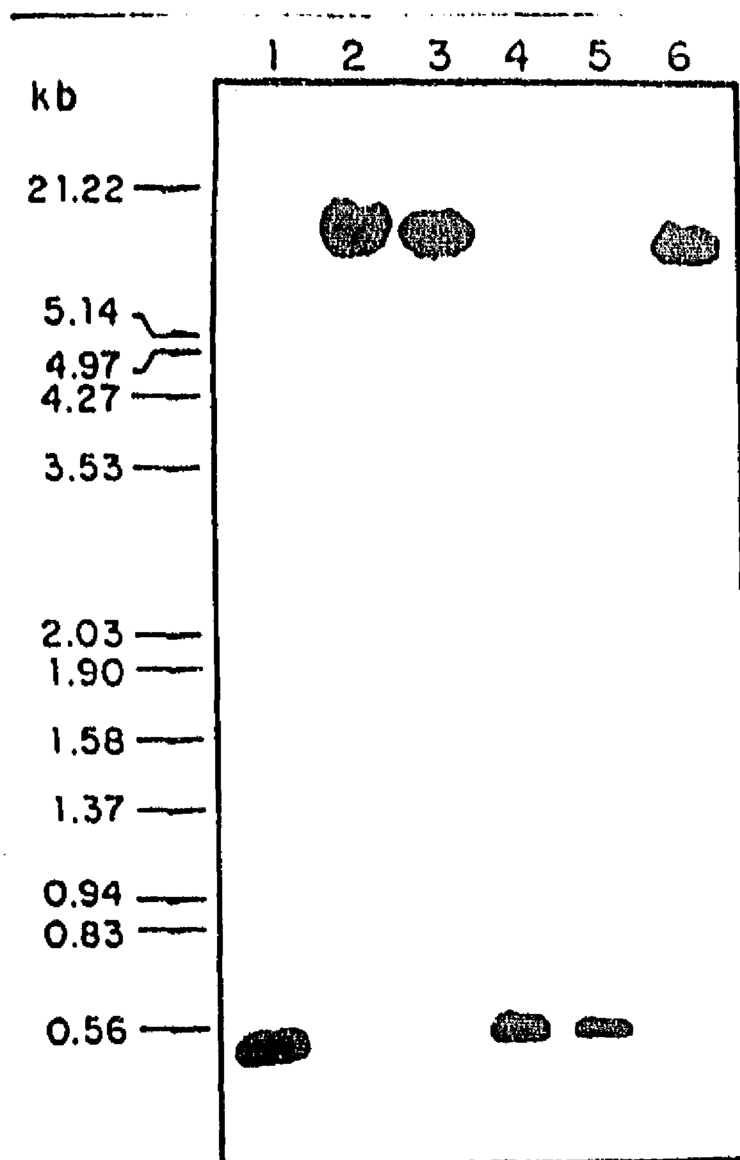

Fig. 3. Southern blot genomic analysis. *CnH* DNA digested with restriction endonucleases was blotted in lanes: 1, *Eco*RI; 2, *Bam*HI; 3, *Hin*dIII; 4, *Eco*RI+*Bam*HI; 5, *Eco*RI+*Hin*dIII; 6, *Bam*HI+*Hin*dIII. Methods: High M_r genomic DNA was prepared from the whole body of *CnH*. After ethanol precipitation the DNA was resuspended in water and subjected to molecular filtration chromatography on an agarose A-50 (Bio-Rad, 0.7 × 28 cm) column equilibrated (and then eluted) with 10 mM Tris pH 7.6/1 mM EDTA/0.75 M NaCl. Twenty fractions of 0.5 ml were collected, and the DNA profile was determined by absorbance at 260 nm. Fractions corresponding to the maximum of absorbance were pooled and ethanol-precipitated. This step was necessary in order to eliminate a pigment that coprecipitates with scorpion DNA. The colored material was shown to inhibit the enzymatic digestion of DNA. After digestion with the indicated endonucleases, 10-μg samples of DNA were electrophoresed in a 0.7% agarose gel, blotted onto a nitrocellulose membrane and probed independently with the inserts derived from clones *CngtI* or *CngtII*. The probes were ^{32}P-labeled using the random primer labeling kit from Dupont. The filters were prehybridized, hybridized and washed as described in the legend to Fig. 1 for double-stranded probes (second step screening), except that an additional wash at 65°C for 10 min was performed.

kb band might contain at least these four 0.6-kb *Eco*RI segments. An alternative explanation is that each cDNA (mRNA) could be encoded by a different 18-kb band. Thus, their corresponding genes may also be encoded in the same or different 18-kb DNA bands, and consequently, they might have the same organization as the 0.6-kb *Eco*RI fragments in which a substantial part of the region transcribed into their respective mRNAs might be present. Furthermore, these results also suggest that the genes encoding these STox lack large introns.

(d) PCR genomic analysis

In order to understand the genomic organization of the structural DNA regions encoding *Cn*H STox, a PCR amplification experiment using genomic DNA and specific primers was performed. The DNA sequences that were used for the synthesis of the PCR primers correspond to well-conserved sequences of *Cn*H cDNA. They are shown overlined in Fig. 1 and flank a stretch of 263–266 bp of DNA that includes the mature part of the STox encoded by these cDNA. As can be seen in Fig. 4, while several DNA bands were visualized after staining with EtdBr, when using 0.5 μg of genomic DNA as the PCR substrate (Fig. 4A, lane 2), only one specific DNA band, of about 270 bp, was observed after hybridization (Fig. 4B, lane 2). This band corresponds to the smallest band stained with EtdBr. When ten times more genomic DNA (5 μg) was used as the substrate for PCR amplification, a second specific band of about 330 bp was detected

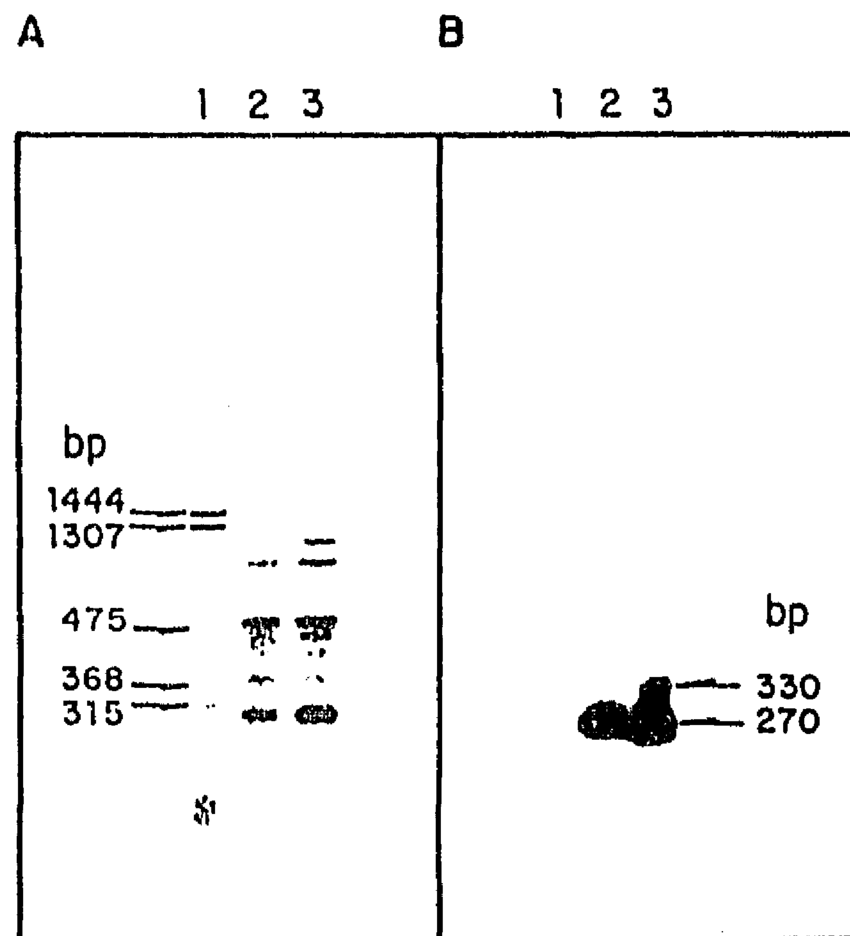

Fig. 4. PCR genomic analysis. Primers for genomic PCR amplifications were synthesized as described in the legend to Fig. 1. The forward primer (5'-AAAGAAGGTTATCTGGTAAAC), corresponds to a well-conserved DNA sequence that codes for the first 7 aa of mature *C. noxius* toxins (see Figs. 1 and 2). The reverse primer (5'-AACTGCAATTAAGAAGCGTTA) is complementary to a well-conserved nt 246–266 sequence of cDNA (see Fig. 1). The source of the DNA was the same as for the Southern blot genomic analysis. PCR conditions were the same as described in the legend to Fig. 1, except that 0.5 μg or 5 μg of template DNA were used. (Panel A) Products of PCR amplification; lanes: 1, pBR322 (Bolivar et al., 1977) digested with *Taq*I as size marker; 2, PCR sample (1/20 of total reaction) using 0.5 μg of genomic DNA as template; 3, same as lane 2 but using 5 μg of template DNA. The gel was stained with EtdBr and photographed under ultraviolet light. (Panel B) Autoradiography of PCR products shown in panel A hybridized with *CngtI* insert. The pBR322 DNA was also transferred and hybridized as a negative control.

(Fig. 4B, lane 3). It is important to observe that the 330-bp band was detected only when high concentrations of genomic DNA were used as a substrate for PCR, suggesting that this band might correspond to a region that might have an at least tenfold lower copy number than the 270-bp band versions. Taking the cDNA sequences shown in Fig. 1 as prototypes of DNA sequences that encode *Cn*H STox, these PCR results indicate that the 270-bp product should code for the mature part of *Cn*H STox closely related, to the ones encoded by the *Cngt*II–*Cngt*V cDNA. These results also indicate that those genomic regions do not contain introns. However, the results in Fig. 4B (lane 3) clearly show the presence of a 330-bp fragment which could accommodate an intron of about 60 bp in its genomic region. Another possibility is that the 330-bp band corresponds to a rare version of a closely related gene. Preliminary results aimed at verifying the existence of introns in the genomic regions encoding the 5′ end of *Cn*H STox cDNAs (including the SP) did not detect introns. We have sequenced some of the cloned PCR products, and we did not find introns at least between the regions coding for the SP and the C terminus (data not shown). An interpretive review of the results presented in this work, including detailed comparative analyses and recent data on the STox structure-function relationship, is now in preparation for publication (B.B., A.V., C.G., M.C., L.D.P. and F.B.).

(e) Conclusions

(1) We have prepared and probed a λgt11 cDNA library from the venom glands of the Mexican *Cn*H scorpion. From this cDNA library we have isolated four cDNAs that code for different STox.

(2) We have determined the nt sequence of these cDNAs and deduced their aa sequences. From these sequences we have shown that they are closely related to the STox *Cn*1 of *Cn*H and to variants 2 and 3 of *Cs*E. *Cngt*V corresponded to STox II-10 of *Cn*H, which has been purified and sequenced in our group.

(3) From the nt sequence data, it seems that these *Cn*H cDNAs derived from a common gene ancestor. We propose that this gene might have duplicated, and each duplication then evolved independently.

(4) From Southern blot genomic analysis, it can be proposed that this family of closely related genes, including at least the genes coding for the four cDNAs reported here, shows a minimal size of about 600 bp when the genomic DNA is digested with *Eco*RI.

(5) PCR genomic amplifications of *Cn*H DNA with specific primers and hybridization with specific probes suggest that the region(s) encoding the mature part of STox closely related to the ones reported here do not seem to contain large introns.

(6) Comparison of STox sequences deduced from *Cn*H cDNAs provided background for their classification, in terms of their similarity, in three different groups.

ACKNOWLEDGEMENTS

This work was supported in part by grants of the Universidad Nacional Autónoma de México project DGAPA No. IN202689 to L.D.P., CONACyT (México) project 0018-N9105 to L.D.P. and F.B. and Howard Hughes Medical Institute No. 75191–527104 to L.D.P. Training received by B.B. in the laboratories of Dr. Ki-Han Kim (Department of Biochemistry, Purdue University) and Dr. Edward Ginns (National Institute for Mental Health, Bethesda) is greatly appreciated. Dr. Fernando López Casillas and Dr. Brian Martin are acknowledged for helpful discussions. Oligos were synthesized at the 'Unidad de Sintesis de Oligonucleótidos' of the Instituto de Biotecnologia, with the assistance of Paul Gaytán. The technical assistance of Fernando Zamudio is also recognized.

REFERENCES

Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heyneker, H.L., Boyer, H.W., Crosa, J.H. and Falkow, S.: Construction and characterization of new cloning vehicles, II. A multipurpose cloning system. Gene 2 (1977) 95–113.

Bougis, P.E., Rochat, H. and Smith, L.A.: Precursors of *Androctonus australis* scorpion neurotoxins. J. Biol. Chem. 264 (1989) 19259–19265.

Carbone, E., Wanke, E., Prestipino, G., Possani, L.D. and Maelicke, A.: Selective blockage of voltage-dependent K⁺ channels by a novel scorpion toxin. Nature 296 (1982) 90–91.

Carbone, E., Prestipino, G., Franciolini, F., Dent, M.A.R. and Possani, L.D.: Selective modification of the squid axon Na currents by *Centruroides noxius* toxin II-10. J. Physiol. Paris 79 (1984) 179–184.

Catterall, W.A.: Activation of the action potential Na⁺ ionophore by neurotoxins. An allosteric model. J. Biol. Chem. 252 (1977) 8669–8676.

Couraud, F., Jover, E., Dubois, J.M. and Rochat, H.: Two types of scorpion toxin receptor sites, one related to the activation, the other to the inactivation of the action potential sodium channel. Toxicon 20 (1982) 9–16.

Dehesa-Dávila, M.: Epidemiological characteristics of scorpion sting in León, Guanajuato, México. Toxicon 27 (1989) 281–286.

Fontecilla-Camps, J.C., Almassay, R.J., Suddath, F.L., Watt, D.D. and Bugg, C.E.: Three dimensional structure of a protein from scorpion venom: a new structural class of neurotoxins. Proc. Natl. Acad. Sci. USA 77 (1980) 6496–6500.

Gimenez-Gallego, G., Navia, M.A., Reuben, J.P., Katz, G.M., Kaczorowski, G.J. and Garcia, M.L.: Purification sequence, and model structure of charybdotoxin, a potent selective inhibitor of calcium-activated potassium channels. Proc. Natl. Acad. Sci. USA 85 (1988) 3329–3333.

Gurevitz, M., Urbach, D., Zlotkin, E. and Zilberberg, N.: Nucleotide sequence and structure analysis of a cDNA encoding an alpha insect toxin from the scorpion *Leiurus quinquestriatus hebraeus*. Toxicon 29 (1991) 1270–1272.

Gurrola, G.B., Molinar-Rode, R., Sitges, M., Bayon, A. and Possani, L.D.: Synthetic peptide corresponding to the sequence of Noxiustoxin indicate that the active site of this K^+ channel blocker is located on its N-terminal portion. J. Neural Transm. 77 (1989) 11–20.

Jover, E., Couraud, F. and Rochat, H.: Two types of scorpion neurotoxins characterized by their binding to two separate receptor sites on rat brain synaptosomes. Biochem. Biophys. Res. Commun. 95 (1980) 1607–1614.

Meves, H., Simard, M.J. and Watt, D.D.: Biochemical and electrophysiological characteristics of toxins isolated from the venom of the scorpion *Centruroides sculpturatus*. J. Physiol. Paris 79 (1984) 185–191.

Meves, H., Simard, M.J. and Watt, D.D.: Interactions of scorpion toxins with the sodium channel. In: Kao, C.Y. and Levinson, S.R. (Eds.), Tetrodotoxin, Saxitoxin and the Molecular Biology of the Sodium Channel. Ann. NY Acad. Sci. 479 (1986) 113–132.

Miller, C., Moczydlowski, E., Latorre, R. and Phillips, M.: Charybdotoxin, a protein inhibitor of single Ca^{2+}-activated K^+ channels from mammalian skeletal muscle. Nature 313 (1985) 316–318.

Miranda, F., Kopeyan, C., Rochat, H., Rochat, C. and Lissitzky, S.: Purification of animal neurotoxins. Isolation and characterization of eleven neurotoxins from the venom of the scorpions *Androctonus australis* Hector, *Buthus occitanus tunetanus* and *Leiurus quinquestriatus quinquestriatus*. Eur. J. Biochem. 16 (1970) 514–523.

Nonner, W.: Effects of *Leiurus* scorpion venom on the 'gating' current in myelinated nerve. In: Ceccarelli, B. and Clementi, F. (Eds.), Advances in Cytopharmacology, Vol. 3. Raven Press, New York, 1979, pp. 345–351.

Possani, L.D.: Structure of scorpion toxins. In: Tu, A.T. (Ed.), Handbook of Natural Toxins, Vol. 2. Marcel Dekker, New York, 1984, pp. 513–550.

Possani, L.D., Dent, M.A.R., Martin, B.M., Maelicke, A. and Svendsen, I.: The N-terminal sequence of several toxins from the venom of the Mexican scorpion *Centruroides noxius* Hoffmann. Carlsberg Res. Commun. 46 (1981) 207–214.

Possani, L.D., Martin, B.M. and Svendsen, I.: The primary structure of Noxiustoxin: a K^+ channel blocking peptide, purified from the venom of the scorpion *Centruroides noxius* Hoffmann. Carlsberg Res. Commun. 47 (1982) 285–289.

Possani, L.D., Gurrola, G.B., Portugal, T.O., Zamudio, F.Z., Vaca, L.D., Calderon, E.S.A. and Kirsch, G.E.: Scorpion toxins: a model for peptide synthesis of new drugs. In: Oliveira, B. and Sgarbieri, V. (Eds.) Proceedings of the First Brazilian Congress on Proteins. Editora da UNICAMP, Campinas, Brazil, 1991, pp. 352–367.

Strichartz, G., Rando, T. and Wang, G.K.: An integrated view of the molecular toxinology of sodium channel gating in excitable cells. Annu. Rev. Neurosci. 10 (1987) 237–267.

Strong, P.N., Weir, S.W., Beech, D.J., Hiestand, P. and Kocher, H.P.: Effects of potassium channel toxins from *Leiurus quinquestriatus hebraeus* venom on responses to cromakalin in rabbit blood vessels. Br. J. Pharmacol. 98 (1989) 817–826.

Thomsen, W.J. and Catterall, W.A.: Localization of the receptor site for α-scorpion toxins by antibody mapping: implications for sodium channel topology. Proc. Natl. Acad. Sci. USA 86 (1989) 10161–10165.

Valdivia, H.H., Fuentes, O., El-Hayek, R., Morrissette, J. and Coronado, R.: Activation of the ryanodine receptor Ca^{2+} release channel of sarcoplasmic reticulum by a novel scorpion venom. J. Biol. Chem. 266 (1991) 19135–19138.

von Heijne, G.: A new method for predicting sequence cleavage sites. Nucleic Acids Res. 14 (1986) 4683–4690.

Wheeler, P.K., Watt, D.D. and Lazdunski, M.: Classification of Na channel receptors specific for various scorpion toxins. Pflügers Arch. 397 (1983) 164–165.

Zamudio, F., Saavedra, R., Martin, B.M., Gurrola-Briones, G., Herion, P. and Possani, L.D.: Amino acid sequence and immunological characterization with monoclonal antibodies of two toxins from the venom of the scorpion *Centruroides noxius* Hoffmann. Eur. J. Biochem. 204 (1992) 281–292.

Zlotkin, E., Miranda, F. and Rochat, C.: Chemistry and pharmacology of Buthinae scorpion venoms. In: Bettini, S.(Ed.), Handbook of Experimental Pharmacology, Vol. 48. Springer-Verlag, Berlin, 1978, pp. 317–369.

3. DATOS NO PUBLICADOS

3.1. OBTENCION Y COMPARACION DE LA SECUENCIA DE NUCLEOTIDOS DE cDNAs

Siguiendo la metodología descrita en los artículos incluidos en la sección anterior, se aislaron y secuenciaron otras clonas de cDNA encontrándose que varias de ellas (incluída la Cngtl, CnH41 y Cn141) son clonas truncadas. Contienen, a partir de lo que correspondería al resíduo de cisteína 25 hasta el sitio de la cola de poliA, la secuencia de un cDNA que codifica para una toxina bloqueadora de canales de sodio aún no reportadas por secuencia directa de los péptidos (figura 3).

```
TGC GAT AAG GAA TGC AAA GCG AAG AAC CAA GGA GGT AGT TAC GGC TAT TGC TAC GCT TTT GCA TGC   66
 C   D   K   E   C   K   A   K   N   Q   G   G   S   Y   G   Y   C   Y   A   F   A   C    22
                                    CAA GGA GGT AGT TAC GGC TAT TGC TAC GCT TTT GCA TGC   39
                                     Q   G   G   S   Y   G   Y   C   Y   A   F   A   C    13
                                    CAA GGA GGT AGT TAC GGC TAT TGC TAC GCT TTT GGG TGC   39
                                     Q   G   G   S   Y   G   Y   C   Y   A   F   G   C    13

TGG TGC GAA GGT TTG CCC GAA AGT ACA CCG ACT TAT CCC CTT CCT AAT AAA TCA TGT GGC AGA AAA  132
 W   C   E   G   L   P   E   S   T   P   T   Y   P   L   P   N   K   S   C   G   R   K    44
TGG TGC GAA GGT TTG CCC GAA AGT ACA CCG ACT TAT CCC CTT TCT AAT AAA TCA TGT GGC AGA AAA  105
 W   C   E   G   L   P   E   S   T   P   T   Y   P   L   S   N   K   S   C   G   R   K    35
TGG TGC GAA GGT TTG CCC GAA AGT ACA CCG ACT TAT CCC CTT CCT AAT AAA ACA TGC AGC AAA AAA  105
 W   C   E   G   L   P   E   S   T   P   T   Y   P   L   P   N   K   T   C   S   K   K    35

TAA TGGCAACGAGTGTTTATTGTCTACCAACAGAAATATTGTAACGCTTCTTAATTGCAGTTAAATAAAATAAAATGTAATATC  216
TAA TGGCAACGAGTTTTTATTGTTTACCAACAGAAATATTGTAACGCTTCTTAATTGCAGTTAAATAAAATAAAATGTAATATC  189
TAA TGGCAACGACTTTTTATTGTCCACCAACAGAAATATTGTAACGCTTCTTAATTGCAGTTAAATGAAATAAAATTAATAGCA  189

TTAAAAAAAAAAAAAA                                                                        232
TAAAAAAAAAAAAAAA                                                                        205
TTATTTTAAATATCTTTATTATGTGTATACAGTTTACAGAAAATAAAACCGTTATGTTTATTCTCTGAAAAAAAAAAA          252
```

Figura 3.
Secuencia de nucleótidos de las clonas: Cngt1, CnH41 y Cn141, respectivamente. En letras oscuras se indica la secuencia de aminoácidos deducida, el codón de terminación (TAA) y la posible señal de poliadenilación (AATAAA).

Además, se encontró una clona que tiene características muy peculiares: dentro de la región 5' se encuentran duplicaciones e inversiones de lo que equivaldría a la secuencia nucleotídica que codificaría para el péptido señal y la correspondiente a los resíduos de aminoácidos de la región amino terminal. El significado que pudiera tener este hallazgo, desde un punto de vista evolutivo aún no se ha podido descifrar, pero se sigue trabajando con esta clona con el fin de determinar si se encuentra presente dentro del genoma, o fue un artificio de las técnicas utilizadas.

Es importante hacer notar que varias de las clonas aisladas y secuenciadas resultaron ser la CngtII, una toxina contra insectos, lo cual podría reflejar la abundancia de dicha toxina dentro del veneno de este alacrán. Se ha mencionado también que dentro del veneno de *C. sculpturatus* las toxinas más abundantes son las tres variantes que corresponden a toxinas contra insectos (Babin *et al.*, 1974). Al parecer, los alacranes presentan una abundancia mayor de toxinas contra insectos que contra mamíferos, lo cual resulta lógico, ya que los alacranes se alimentan principalmente de insectos.

3.2. COMPARACION DEL PEPTIDO SEÑAL Y LA REGION 3' QUE NO SE TRADUCE

Como se puede ver en la figura 4, los péptidos señal de los cDNAs que codifican para diferentes toxinas del veneno de alacranes tanto del nuevo mundo (primer bloque), como del viejo mundo (segundo bloque), son muy parecidas en su estructura general, ajustándose a lo propuesto por von Heijne (1986) para un péptido señal funcional.

```
               -20                -1
CngtII         M-NSLLMITACLFLIGT-VWA
CngtIII        M-NSLLMITACLVLFGT-VWA
CngtIV         M-NSLLIITACLVLIGT-VWA
CngtV          M-NSLLMITACLALVGT-VWA
TsVII          MKGMILFIS-CLLLIGIVVEC

LqhαIT         M-AHLVMISLALLLLLG-VES
AaHI           M-NYLVMISLALLLMIG-VES
AaHIII         M-NYLVNISLALLLMTG-VES
AaHIT          M-KFLLLFLVVLPIMG--VES
```

Figura 4.
Comparación de péptidos señal de algunos cDNAs secuenciados de: Cn *C. noxius*; Ts *Tityus serrulatus*; Lqh *Leiurus quinquestriatus hebraeus*; AaH *Androctonus australis* Hector.

Los resíduos -1 y -3 son los más importantes para el sitio de ruptura del péptido señal, en donde -1 debe ser pequeño (Ala, Ser, Gly, Cys, Thr o Gln) y -3 no debe ser aromático, cargado, grande o polar. Como se observa en la Fig. 4, en todos los péptidos señal comparados, en -1 hay un resíduo de aminoácido pequeño y en -3 se encuentra una valina en todos los casos. Además todos contienen una región central hidrofóbica.

En la figura 5 se muestra la comparación de la región 3' que no codifica y que se extiende desde el codón de terminación, TAA, hasta la cola de poliA. En ésta, podemos observar que existe una gran similitud entre todos los cDNAs que codifican para toxinas del veneno de alacranes descritos hasta el momento, en especial cuando se comparan los del género *Centruroides* entre sí, y los del viejo mundo con ellos mismas. Aún no hemos podido determinar el significado que pudiera tener esto, pero no deja de resultar interesante el hecho de que a pesar de que las toxinas sean tan diferentes en su parte estructural, se parezcan tanto en las regiones que no codifican.

```
CngtI     TGGCAACGAGTGTTTATTGTCTACCAACAGAAATATTGTAACGCTTCTTAATTGCAGTTAAATAAAATAAAATG-TAATATC-TT
CngtII    TGGCAACGAC-TTTTATTGTTTACCAACAGAAATATTGTAACGCTTCTTAATTGCAGTTAAATG
CngtIII   TGGCAACGACTTTTTATTGTCCATAAACAGAAATATTGTAACGCTTCTTAATTGCAGTTAAATGAAATAAAATGCTAATAGCATT
CngtIV    TGGCAACGACTTTTTATTGTCCACCAACAGAAATAGTGTAACGCTTCTTAATTCCAAGT
CngtV     TGGCAACGACTTTTTATTGCCCACCAACAGAAATATTGTAACGCTTCTTAATTTCAATTAAATGAAATAAAATATTA-TACCTTT
TsVII          ATTTGTTTCGCTGAAAATCCTTTACAAATGAACTGTAATAAG----TTTGGA--A----AAATAAAAAAAT----G-TTC

LqhαIT         A--TCTGTAGAACATAAACACAAAGAATGTATCCTAAGAAT-------TGATC------AAATAAA----TAA
AaHI           AC--CTGTAGAGTAAAATCAGAAAGAATGTATCCTAAAAATAAC----TGGT-------AAATAAA--CATAA--GTATA
AaHIII         AC--CTGTAAAGCAAAA-CACAAAGAATGTATCCTGAAAATAAC----TGGT-------AAATAAA--CATAA--GTAAT
AaHIT              TTGTAATAATTAT-------GAA-GTAT---GAA--T-------TGATCTA----AAATAAAATGCACATA
```

Figura 5.
Comparación de la región 3' que no se traduce de algunas secuencias de cDNAs de: Cn *Centruroides noxius*; Ts *Tityus serrulatus*; Lqh *Leiurus quinquestriatus hebraeus*; AaH *Androctonus australis* Hector.

3.3. ANALISIS Y COMPARACION DE LA SECUENCIA DE AMINOACIDOS DEDUCIDA DE LOS cDNAs CON LA SECUENCIA DE AMINOACIDOS REPORTADAS

En la figura 1 (página 6) se muestra la estructura primaria de varias toxinas del veneno de alacranes tanto del nuevo mundo (primer bloque), como del viejo mundo (segundo bloque). Al realizar un alineamiento con base a los resíduos de cisteína, introduciendo guiones para maximizar las similitudes entre todas las secuencias, se observa que existen dos regiones principales en donde se permite la mayor variabilidad en la secuencia de aminoácidos. A estas regiones se les ha denominado asa J y asa B (Meves *et al.*, 1984), ya que se ha visto que éstas coinciden con dos asas dentro de la estructura terciaria (fig. 2, página 8). Dichas asas conectan diferentes partes de los principales motivos estructurales y por ello pueden tener varias longitudes, estructuras y orientación espacial. Es interesante hacer notar que dichas asas coinciden con la división de toxinas tipo α o del viejo mundo y toxinas tipo β o del nuevo mundo, lo cual podría estar indicando que ésta es una característica estructural importante para determinar el modo de acción de la toxina, excepto para el caso de las toxinas que tienen actividad dual, contra insectos y mamíferos en las que se puede ver que presentan parte de ambas asas (AaHIT4; Loret *et al.*, 1991).

Otro motivo estructural importante que destaca en este alineamiento, es el CHS (Cystein-stabilized α-helical) propuesto recientemente por Kobayashi y colaboradores (1991), el cual se ha encontrado que está presente en péptidos neurotóxicos con una actividad común: bloquean canales iónicos. Dicho motivo estructural es un par de cisteínas espaciadas por tres resíduos de animoácidos (Cys-X-X-X-Cys), unidos por puentes disulfuro a un segundo par de resíduos de cisteína con solo un resíduo de aminoácido entre ellas (Cys-X-Cys). En este caso, se trata de los resíduos correspondientes a las posiciones 25-26-27-28-29 y 46-47-48, tomando como referencia a la toxina Cn4 (representado con asteriscos en la figura 1, página 6).

Ménez y colaboradores (1992) señalan que las toxinas de origen animal tienen varias características en común: son generalmente pequeñas (menos de 120 resíduos de aminoácidos), con un alto contenido de disulfuros, estables a condiciones de desnaturalización y ataque enzimático. Aunque expresan múltiples funciones, las toxinas animales generalmente tienen un número muy limitado de patrones de conformación. En el veneno de alacranes se dá principalmente una estructura tridimensional básica y en el veneno de serpientes se han encotrando seis conformaciones. Se ha visto también que la estructura general de las toxinas permite múltiples inserciones, deleciones y/o mutaciones (asa J y asa B). La variabilidad que se encuentra en la estructura de las toxinas relacionada con su función en el canal, podría estar reflejando la variabilidad que se ha visto que existe en los canales de sodio (Noda *et al.*, 1986; Catterall *et al.*, 1988),

ESTA TESIS NO DEBE
SALIR DE LA BIBLIOTECA

De la misma manera, puede observarse que a pesar de la gran similitud que existe en la región amino-terminal en todas las secuencias comparadas, en la región carboxilo terminal existen grandes diferencias, haciéndose evidente la riqueza de resíduos de prolina en algunos casos (3 o 4) en tanto que en otros sólo se presentan uno o dos. Ha sido propuesto por Loret y colaboradores (1990), que la falta de prolinas en la región carboxilo terminal le confiere a la molécula menor rigidéz, lo cual podría guardar relación con la especificidad de acción de la toxina (mamífero o insecto).

A pesar de que estos datos sugieren que existe alguna relación entre la estructura de la toxina y su especificidad (por canales iónicos; insecto o mamífero; α o β), se requieren mas datos para poder llegar a entender la relación que existe entre la estructura y la función de las toxinas del veneno de los alacranes.

4. PERSPECTIVAS FUTURAS

En esta tesis se sientan las bases para el conocimiento de los aspectos de biología molecular del alacrán mexicano *C. noxius* Hoffmann. Las clonas de cDNA obtenidas durante la realización de este trabajo pueden ser utilizadas para llevar a cabo estudios como:

1. Mutagénesis dirigida y expresión genética, con el fin de obtener mayor información con respecto a la relación que hay entre la estructura y la función de las toxinas del veneno de los alacranes (tesis de Doctorado de Consuelo García).

2. Estudiar como se encuentran organizados los genes que codifican para las toxinas del veneno de este alacrán y determinar cual es la estructura general de los mismos (Tesis de Maestría de Miguel Corona).

3. Diseñar una estrategia para poder obtener los cDNAs que codifican para toxinas específicas contra mamíferos (Cn2 y Cn3).

4. Con base en los datos obtenidos en los puntos anteriores, diseñar una vacuna híbrida y comparar su comportamiento inmunológico con los datos que se tienen en nuestro laboratorio con péptidos sintéticos (Calderón, 1992).

5. A partir de las toxinas obtenidas por la combinación de mutagénesis dirigida y expresión genética junto con técnicas de resonancia magnética nuclear y difracción de rayos X, ampliar la información que se tiene referente a la relación que existe entre la estructura de las toxinas del veneno de los alacranes y su función sobre los canales de sodio.

5. REFERENCIAS

Almassy, R. J., Fontecilla-Camps, J. C., Suddath, F. L. and Bugg, Ch. E. (1983). Structure of variant-3 scorpion neurotoxin from *Centruroides scupturatus* Ewing, refined at 1.8 Å resolution. J. Mol. Biol. 170:497-527.

Babin, D. R., Watt, D. D., Goss, S. M. and Mlejnek, R. V. (1974). Amino acid sequence of a neurotoxic protein variant from the venom of *Centruroides sculpturatus* Ewing. Arch. Biochem. Biophis. 164:694-706.

Becerril, B., Vázquez, A., García, C., Corona, M., Bolivar, F. and Possani, L. D. (1993). Cloning and characterization of cDNAs that code fot Na^+-channel blocking toxins of the scorpion *Centruroides noxius* Hoffmann. Gene. (en prensa).

Bontems, F., Roumestand, C., Gilquin, B., Ménez, A. and Toma, F. (1991). Refined structure of charybdotoxin: common motifs in scorpion toxins and insect defensins. Science 254:1521-1523.

Bougis, P. E., Rochat, H. and Smith, L. A. (1989). Precursors of *Androctonus australis* scorpion neurotoxins. J. Biol. Chem. 264:19259-19265.

Briux, M., Jiménez, M. A., Santoro, J., González, C., Colilla, F. J., Méndez, E. and Rico, M. (1993). Solution structure of γ1-H and γ1-P thionins from barley and wheat endosperm determined by ^{1}H-NMR: A structural motif common to toxic arthropod proteins. Biochem. 32:715-724.

Calderón, A. E. (1992). Estudio de la relación estructura-función y capacidad inmunogénica de neurotoxinas y péptidos análogos sintéticos del veneno del alacrán *Centruroides noxius* Hoffmann. Tesis Doctoral. UACPyP - CCH. UNAM.

Carbone, E., Wanke, E., Prestipino, G., Possani, L. D. and Maelicke, A. (1982). Selective blockage of voltage-dependent K^+ channels by a novel scorpion toxin. Nature. 296:90-91.

Catterall, W. A. (1976). Purification of a toxic protein from scorpion venom which activates the action potential Na^+ ionophore. J. Biol. Chem. 251:5528-5536.

Catterall, W. A. (1988). Structure and function of voltage-sensitive ion channels. Science. 241:50-61.

Couraud, F., Jover, E., Dubois, J. M. and Rochat, H. (1982). Two types of scorpion toxin receptor sites, one related to the activation, the other to the inactivation of the action potential sodium channel. Toxicon. 20:9-16.

Darbon, H., Weber, C. and Braun, W. (1991). Two-dimensional ^{1}H nuclear magnetic resonance study of a AaHIT, an anti-insect toxin from scorpion *Androctonus australis* Hector. Biochem. 30:1836-1845.

Eitan, M., Fowler, E., Hermann, R., Duval, A., Pelhate, M. and Zlotkin, E. (1990). A scorpion venom neurotoxin paralitic to insect that affects sodium current inactivation: purification, primary structure and mode of action. Biochem. 29:5941-5947.

Fontecilla-Camps, J. C., Habersetzer- Rochat, C. and Rochat, H. (1988). Orthorhombic crystals and three-dimensional structure of the potent toxin II from the scorpion *Androctonus australis* Hector. Proc. Natl. Acad. Sci. USA. 85:7443-7447.

Gimenez-Gallego, G., Navice, M. A., Reuben, J. P., Katz, G. M., Kaczorowski, G. J. and Garcia, M. L. (1988). Purification, sequence and model structure of charybdotoxin, a potent selective inhibitor of calcium-activated potassium channels. Prpoc. Natl. Acad. Sci. USA. 85:3329-3333.

Gurevitz, M., Urbach, D., Zlotkin, E. and Zilberberg, N. (1991). Nucleotide sequence and structure analysis of a cDNA encoding an α insect toxin from the scorpion *Leiurus quinquestriatus hebraeus.* Toxicon. 29:1270-1272.

Jover, E., Couraud, F. and Rochat, H. (1980). Two types of scorpion neurotoxins characterized by their binding to two separate receptor sites on rat brain synaptosomes. Biochem. Biophys. Res. Commun. 95:1607-1614.

Kirsch. G. E., Skattebol, A., Possani, L. D. and Brown, A. M. (1989). Modification of a Na channel gating by an α scorpion toxin from *Tityus serrulatus*. J. Gen. Physiol. 93:67-83.

Kobayashi, Y., Takashima, H., Tamaoki, H., Kyogoku, Y., Lambert, P., Kuroda, H., Chino, N., Watanabe, T., Kimura, T., Sakakibara, S. and Moroder, L. (1991). The cystein-stabilized α-helix: A common structural motif of ion-channel blocking neurotoxic peptides. Biopolymers. 31:1213-1220.

Kopeyan, Ch., Mansuelle, P., Sampieri, F., Brando, T., Bahraoui, E. M., Rochat, H. and Granier, C. (1990). Primary structure of scorpion anti-insect toxins isolated from the venom of *Leiurus quinquestriatus quinquestriatus.* FEBS Lett. 261:423-426.

Loret, E. P., Sampieri, F., Roussel, A., Granier, C. and Rochat, H. (1990). Conformational flexibility of a scorpion toxin active on mammals and insects: a circular dichroism study. Proteins. 8:164-172.

Loret, E. P., Martin-Euclaire, M. F., Mansuelle, P., Sampieri, F., Granier, C. and Rochat, H. (1991). An anti-insect toxin purified from the scorpion *Androctonus australis* Hector also acts on the α and β- sites of the mammalian sodium channel: sequence and circular dichroism study. Biochem. 30:633-640.

Martin-Euclaire, M. F., Ceard, B., Ribeiro, A. M., Diniz, C. R., Rochat, H. and Bougis, P. E. (1992). Molecular cloning and nucleotide sequence analysis of a cDNA encoding the main β-neurotoxin from the venom of the South American scorpion *Tityus serrulatus*. FEBS Lett. 302:220-222.

Ménez, A., Bontems, F., Roumestand, C., Gilquin, B. and Toma, F. (1992). Structural basis for functional diversity of animal toxins. Proc. Royal Society of Edinburgh. 99B:83-103.

Meves, H., Simard, J. M. and Watt, D. D. (1984). Biochemical and electrophysiological characteristics of toxins isolated from the venom of the scorpion *Centruroides sculpturatus*. J. Physiol. Paris. 79:185-191.

Meves, H., Simard, J. M. and Watt, D. D. (1986). Interactions of scorpion toxins with the sodium channel. Annals New York Acad. Sci. 479:113-132.

Noda, M., Ikeda, T., Suzuki, H., Takeshima, H., Kurasaki, K., Takahashi, H. and Numa, S. (1986). Existence of distinct sodium channel messenger RNAs in rat brain. Nature. 320:188-192.

Nonner, W. (1979). Effects of *Leiurus* scorpion venom on the "gating" current in myelinated nerve. In: Advances in Cytopharmacology. Vol. 3. B. Ceccarelli and F. Clementi Eds. Raven Press N. Y. pp345-352.

Pelhate, M and Zlotkin, E. (1982), Actions of insect toxin and other toxins derived from the venom of the scorpion *Androctonus australis* on isolated giant axons of the cockroach. J. Exp. Biol. 97:67-77.

Possani, L. D. (1984). Structure of the scorpion toxins. In: Handbook of Natural Toxins. Vol. 2. A. Tu Ed. Marcel Dekker, Inc. N. Y. pp513-550.

Possani, L. D. and Dent, M. A. R. (1981). The amino terminal sequence of several toxins from the venom of the mexican scorpion *Centruroides noxius* Hoffmann. Carlsberg Res. Commun. 46:207-214.

Possani, L. D., Martin, B. M. and Svendsen, I. (1982). The primary structure of noxiustoxin: a K^+ channel peptide, purified from the venom of the scorpion *Centruroides noxius* Hoffmann. Carlsberg Res. Commun. 47:285-289.

Possani, L. D., Martin, B. M., Svendsen, I. Rode, G. S. and Erickson, B. W. (1985). Scorpion toxins from *Centruroides noxius* and *Tityus serrulatus*. Biochem. J. 229:739-750.

Rochat, H., Bernard, P. and Couraud, F. (1979). Scorpion toxins: Chemistry and mode of action. In: Advances in Cytopharmacology. Vol. 3. B. Ceccarelli and F. Clementi Eds. Raven Press. New York. pp325-334.

Vázquez, A., Becerril, B., Martin, B. M., Zamudio, F., Bolivar, F. and Possani, L. D. (1993). Primary structure determination and cloning of the cDNA encoding toxin 4 of the scorpion *Centruroides noxius* Hoffmann. FEBS Lett. 320:43-46.

Von Heijne, G. (1986). A new method for predicting sequence cleavage sites. Nucl. Acids. Res.14:4683-4690.

Zamudio, F., Saavedra, R., Martin, B. M., Gurrola-Briones, G., Herion, P. and Possani, L. D. (1992). Amino acid sequence and immunological characterization with monoclonal antibodies of two toxins from the venom of the scorpion *Centruroides noxius* Hoffmann. Eur. J. Biochem. 204:281-292.

Zilberberg, N., Zlotkin, E. and Gurevitz, M. (1991). The cDNA sequence of a depressant insect selective neurotoxin from the scorpion *Buthotus judaicus*. Toxicon. 29:1155-1158.

Zlotkin, E., Martínez, G., Rochat, H. and Miranda, F. (1975). A protein toxic to crustacean from the venom of the scorpion *Androctonus australis*. Insect. Biochem. 5:243-250.

Zlotkin, E., Miranda, F. and Rochat, H. (1978). Chemistry and pharmacology of *Buthinae* scorpion venoms. In: Handbook of Experimental Pharmacology. Vol. 48. S. Bettini Ed. Springer-Verlag. Berlin. pp317-369.

Zlotkin, E., Eitan, M., Bindokas, V. P., Adams, M. E., Moyer, M. and Fowler, E. (1991). Functional duality and structural uniqueness of depressant insect-selective neurotoxins. Biochem. 30:4814-4821.