

03072
UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

UNIDAD ACADÉMICA DE LOS CICLOS PROFESIONAL Y DE
POSGRADO DEL COLEGIO DE CIENCIAS Y HUMANIDADES
INSTITUTO DE BIOTECNOLOGÍA



**CARACTERIZACION DE LA REGION GENOMICA
QUE CODIFICA PARA UNA VARIANTE DE LA
TOXINA C₂ (II-14) DEL VENENO DEL ALACRAN
*Centruroides noxius Hoffmann***

T E S I S
QUE PARA OBTENER EL GRADO DE
MAESTRA EN BIOTECNOLOGIA
P R E S E N T A
JUANA VIRGINIA TAPIA VIEYRA

CUERNAVACA, MOR.

DE 1997.

**TESIS CON
FALLA DE ORIGEN**



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



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 esta 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.

El presente trabajo se llevó a cabo bajo la dirección del Dr. Baltazar Becerril Luján en el laboratorio del Dr. Lourival Domingos Possani Postay, del Departamento de Reconocimiento Molecular y Bioestructura del Instituto de Biotecnología de la Universidad Nacional Autónoma de México.

**A mis padres y mis hermanos:
Antonio y Carmen**

**A ti, porque no preciso tiempo
ni espacio y menos aún, una definición.**

A todos mis amigos, muy especialmente a Nacho.

A G R A D E C I M I E N T O S

- . A la Universidad Nacional Autónoma de México por permitirme realizar en ella mis estudios.
- . Al Dr. Lourival D. Possani Postay por darme la oportunidad de trabajar en su laboratorio, por su confianza, y por todo el apoyo que siempre me ha brindado.
- . Al Dr. Baltazar Becerril Luján por la dirección en el desarrollo del proyecto y por su inquebrantable entusiasmo.
- . Al M. en Biotecnología Miguel Corona Villegas por el tiempo dedicado en el aprendizaje y manejo de algunas estrategias.
- . A todos mis compañeros y amigos del laboratorio por todo su apoyo.

Este trabajo fue financiado en parte por Howard Hughes Medical Institute (No.75191-527104) y el Consejo Nacional de Ciencia y Tecnología (CONACyT) (No. 4734-N), concedidos al laboratorio del Dr. Lourival D. Possani Postay. La alumna recibió beca de maestría del CONACyT No. 90284.

Comité Tutorial

Dr. Baltazar Becerril Luján

Dr. Mario Enrique Zurita Ortega

Dr. Omar Pantoja Ayala

Jurado de Examen

Presidente

Dr. Baltazar Becerril Luján

Secretario

Dr. Enrique Merino Pérez

Vocal

Dr. Vicente Madrid Marina

Suplente

Dr. Mario Enrique Zurita Ortega

Suplente

Dr. Lourival Domingos Possani Postay

INDICE DE FIGURAS

- Fig. 1** Secuencia nucleotídica de las clonas de DNAc CngtI-CngtIX y sus secuencias de aminoácidos deducidas a partir de las mismas.
- Fig. 2** Secuencia nucleotídica del DNAc de la clona CngtVI correspondiente a la toxina II-14 (Cn 1).
- Fig. 3** Secuencia primaria de la toxina Cn1 (II-14).
- Fig. 4** Autorradiografía en donde se muestra la hibridización del producto de PCR (oligos 5'Gen y Ctox) con un tamaño de 800 pares de bases (pb).
- Fig. 5** Producto de amplificación por PCR con un tamaño aproximado al de la clona de DNAc (Cngt VI).
- Fig. 6** Producto de amplificación del DNA genómico con los oligonucleótidos 5'Gen e Int2Cn1, con un tamaño de alrededor de 1100 pb.
- Fig. 7** Secuencia nucleotídica de la clona de alrededor de 1100 pb.
- Fig. 8** Producto de amplificación por PCR, con los oligonucleótidos 5' Gen y Ctox, con un tamaño de aproximadamente 1300 pb.
- Fig. 9** Secuencia nucleotídica de la clona de 1300 pb.
- FIG.10** Alineamiento de secuencia nucleotídica de las clonas de 1100 pb, 1300pb y CngtVI.

INDICE

RESUMEN.....	1
INTRODUCCION.....	2
1. Alacranes.....	2
2. Importancia del Estudio del Veneno de Alacrán.....	2
3. Extracción y Caracterización del Veneno.....	3
4. Las Toxinas del Veneno de Alacrán.....	4
4.1 Clasificación.....	4
ANTECEDENTES.....	7
OBJETIVO.....	11
METODOLOGIA.....	11
1. Extracción del DNA del alacrán <i>Centruroides noxius</i> Hoffmann.....	11
2. Purificación del DNA de <i>C. noxius</i> Hoffmann.....	11
3. Diseño de oligonucleótidos específicos.....	12
4. Resección en Cadena de la Polimerasa (PCR).....	13
5. Transferencia por el Método de Southern de los productos de amplificación por PCR.....	13
6. Ensayos de hibridización de los productos de amplificación que fueron transferidos.....	13
7. Subclonación.....	13
8. Secuencia Nucleotídica.....	13
9. Análisis e Interpretación de la Secuencia.....	13
RESULTADOS Y DISCUSION.....	14
CONCLUSIONES.....	27
PERSPECTIVAS FUTURAS.....	28
REFERENCIAS.....	29
ANEXO.....	34

RESUMEN

A partir del análisis de las secuencias nucleotídicas que han sido obtenidas de las clonas de Acido Desoxirribonucleico Complementario (DNAC), que codifican para toxinas bloqueadoras de canales de Na^+ del alacrán *Centruroides noxius* Hoffmann, ha podido observarse la alta homología de secuencia que existe entre ellas. La secuencia de la clona CngtVI de la que se comprobó que codifica para la toxina Cn 1 (II-14) del alacrán *C. noxius* Hoffmann, presenta algunas características útiles que permitieron el diseño de un conjunto de oligonucleótidos utilizados para la detección y amplificación de la región genómica que codificara para esta toxina, o bien para una toxina semejante, si se toma en cuenta la homología de secuencia entre las mismas. Los oligonucleótidos diseñados, se utilizaron para el empleo de la técnica de Reacción en Cadena de la Polimerasa (PCR), en donde se usó como templado el DNA genómico de la especie mencionada. Después de llevar a cabo varios intentos de amplificación del DNA genómico, finalmente se montaron las condiciones de amplificación específica y se aisló una clona de 1292 pb, que correspondió a la unidad génica estructural que codifica para una variante de la toxina Cn1 (II-14), así como también se aisló y secuenció una clona de alrededor de 1100 pb, la cual aparentemente está truncada, pero considerando las características que presenta, pudiera tratarse de la región genómica codificante para una toxina semejante a la II-14 (Cn1). La región genómica que codifica para la variante de la toxina Cn 1 (II-14) constituye la primera en ser aislada y caracterizada para un alacrán mexicano en este caso *Centruroides noxius* Hoffmann. El análisis de ambas secuencias se discute en este trabajo.

I N T R O D U C C I O N

1. Alacranes

Los alacranes son animales artrópodos pertenecientes a la Clase Arachnida y al orden Scorpionida. El metasoma o cola del cuerpo del alacrán termina en un segmento bulboso llamado telson. Este es el órgano que utiliza para inyectar el veneno a sus presas (Keegan, 1980). Los alacranes se encuentran distribuidos en varias regiones del mundo, tales como: Este y Norte de Africa, el Cáucaso, el Medio Oriente, la India, Brasil, Venezuela y México (Rochat *et al.*, 1979; Keegan, 1980). El orden Escorpionida incluye nueve familias (Sissom, 1990), una de ellas, la familia Buthidae comprende 48 géneros y más de 500 especies y es la familia de alacranes más grande y amplia (Goyffon y Chippaux, 1990). Un número importante de las especies venenosas actualmente estudiadas, pertenecen a esta familia. El género *Centruroides* perteneciente a esta familia, agrupa a las especies venenosas de México y Estados Unidos. Dentro del género *Centruroides* se encuentra el alacrán *Centruroides noxius* Hoffmann originario de Nayarit, el cual es de particular interés por la alta toxicidad de su veneno, ya que es considerado como uno de los más venenosos del mundo (Dent *et al.*., 1980).

2. Importancia del Estudio del Veneno de Alacrán

Los primeros estudios acerca de los venenos de alacrán fueron motivados principalmente por la importancia médica de las picaduras de alacrán en varias regiones del mundo, como lo es el Este y Norte de Africa, el Medio Oriente, México y Brasil (Rochat *et al.*., 1979; Ménez *et al.*., 1992). Los alacranes de la familia Buthidae, hablando de géneros como: *Androctonus*, *Buthus*, *Centruroides*, *Leiurus* y *Tityus*, motivaron la investigación médica sobre la naturaleza y la composición del veneno, así como el desarrollo de antídotos (Possani, 1984). Considerando la cantidad limitada de veneno que un alacrán puede inyectar y el cuadro clínico que resulta de una picadura simple, es obvio que el veneno de alacrán podría contener componentes altamente activos, cuya purificación fuera deseable para determinar su estructura y elucidar su modo de acción (Rochat *et al.*., 1979). Los accidentes por picadura de alacrán constituyen un problema muy importante de salud principalmente para niños cuya edad se encuentra por debajo de los cinco años y gente de edad avanzada (Alagón *et al.*., 1989).

3. Extracción y Caracterización del Veneno

El veneno de alacrán es una mezcla compleja, obtenida en el laboratorio mediante la estimulación eléctrica del telson de los alacranes (Rochat *et al.*; 1979; Possani, 1984), por la maceración del telson, o permitiendo la picadura del alacrán a través de una película de parafina estirada sobre un matraz (Zlotkin and Shulov, 1969). La heterogeneidad del veneno de alacrán, ha sido bien demostrada por electroforesis sobre gel de almidón (Zlotkin *et al.*, 1971), tiras o láminas de cellogel (Ismail *et al.*, 1974), láminas de acetato de celulosa (Ghazal *et al.*, 1975), gel de poliacrilamida (Chhatwal and Habermann, 1981) y cromatografía en columna (Miranda *et al.*, 1962). En la revisión publicada por Possani en 1984, se describen la obtención y separación del veneno del alacrán de Nayarit, *C. noxius* Hoffman. El veneno es resuspendido en agua y centrifugado para remover residuos celulares y material mucoso. El veneno soluble es usualmente liofilizado y mantenido a -20°C, hasta ser usado. El veneno en solución se purifica por cromatografía en columna de exclusión molecular, obteniéndose tres fracciones, de las cuales las fracciones I y III no son tóxicas a ratones, a una concentración de 120 µg/20g de peso corporal, inyectado intraperitonealmente. La fracción I contiene actividad hialuronidásica. Cuando se sometió a análisis, la fracción III se separó en varios componentes positivos a ninhidrina. La fracción II constituye la mayor proporción del veneno recuperado y es letal a ratones. El término "letal" significa que el componente en la dosis inyectada (en general con cantidades de alrededor de 50 microgramos por ratón) fue suficiente para matar el ratón, en el período de veinte horas después de la inyección. Esta fracción es separada sobre columna de carboximetilcelulosa (CM-celulosa). Ya que las toxinas son polipéptidos básicos, una resina de intercambio catiónico, tal como CM-celulosa, es adecuada para el segundo paso de purificación. Usualmente, un tercer paso cromatográfico es necesario para obtener las toxinas en forma homogénea. Arbitrariamente, se han nombrado las toxinas de acuerdo a su patrón de elución durante la cromatografía (Possani, 1984). Por este procedimiento se han obtenido varias toxinas del veneno de *C. noxius* Hoffmann, entre ellas la toxina Cn 1 (II-14) (las dos letras corresponden al género y especie, el número 1 al orden de secuenciación de la toxina con respecto a las demás, II-14 corresponde al patrón de elución cromatográfico), esta es una toxina de cadena larga (64 aminoácidos), bloqueadora de canales de Na⁺, tóxica para insectos y crustáceos. La estrategia seguida para la determinación de las secuencias de aminoácidos de las toxinas de alacrán, es exactamente la misma que para cualquier otra proteína o polipéptido. El primer aspecto importante es el conocimiento de la composición

química de las toxinas, y de sus péptidos derivados, después de una digestión parcial con proteasas. Esta información da el número y tipo de aminoácidos que debe ser esperado durante la secuenciación. La gran mayoría de las toxinas, ha sido sujeta primero a la degradación automática de Edman.

4. Las Toxinas del Veneno de Alacrán

Las toxinas del veneno de alacrán son proteínas de peso molecular bajo, compuestas de una cadena polipeptídica simple, estabilizadas por tres a cuatro puentes disulfuro (Rochat *et al.*., 1979; Debin *et al.*, 1993; Olamendi-Portugal *et al.*, 1996). La presencia de estos puentes disulfuro pudiera determinar la gran estabilidad de las toxinas de alacrán. El grado máximo de homología entre secuencias primarias es obtenido cuando los péptidos son alineados con respecto a la posición de sus residuos de cisteína (Rochat *et al.*., 1979).

4.1. Clasificación

Los venenos de alacrán contienen una variedad de toxinas que actúan sobre canales de sodio o potasio, en una amplia variedad de formas y con especificidades marcadamente diferentes hacia presas diferentes, incluyendo mamíferos y artrópodos (Zlotkin *et al.*, 1978; Ménez *et al.*., 1992). Las toxinas de alacrán pueden ser divididas en dos categorías, con base en sus estructuras primarias conocidas. Estas son: toxinas de cadena larga y las toxinas de cadena corta (Possani, 1984).

Las toxinas de alacrán de cadena larga, contienen una cadena polipeptídica constituida por 60-70 aminoácidos, estructurada por cuatro puentes disulfuro (Kopeyan *et al.*., 1974). Estas tienen un peso molecular aproximado de 7000 daltones (Miranda *et al.*, 1970, Rochat *et al.*, 1979). Esencialmente alteran la conductancia axonal y pueden ser divididas en cuatro grupos principales. El primero, constituido por las alfa-toxinas, que afectan específicamente la conductancia axonal de mamíferos, prolongando los potenciales de acción, como resultado de una disminución o bloqueo de la inactivación del canal de sodio (Rochat *et al.*., 1979; Catterall, 1984., Kharrat *et al.*.; 1989., Eitan *et al.*.; 1990). El segundo grupo está formado por las toxinas-beta, que afectan la activación del canal de sodio en mamíferos. Las alfa y beta toxinas se unen a sitios distintos en el canal, pero sólo las alfa toxinas se unen de manera voltaje dependiente (Couraud *et al.*., 1982; Catterall, 1984) y muestran cooperatividad positiva con alcaloides solubles en lípidos, tales como la veratridina (Catterall, 1984). Las beta toxinas no

interactúan sinérgicamente con la veratridina (Couraud *et al.*, 1982; Couraud and Jover, 1984). El tercer grupo de toxinas comprende las toxinas depresoras selectivas a insecto. Estas inducen una parálisis flácida progresiva, ya que bloquean los potenciales de acción, por depolarización de membranas axonales y supresión de corrientes de sodio (Zlotkin *et al.*, 1985, Lester *et al.*, 1982). El cuarto grupo está compuesto de las toxinas excitatorias selectivas a insecto, las cuales causan una parálisis espástica inmediata en insectos, provocando un disparo repetitivo en los nervios motores (Walther *et al.*, 1976), debido a un incremento de la corriente máxima de sodio y una disminución voltaje-dependiente de la inactivación de la corriente de sodio (Pelhate and Zlotkin, 1981 y 1982). Las toxinas del cuarto grupo se unen a membranas neuronales de insecto, de una forma voltaje-independiente. Las toxinas de cadena corta de alacrán comprenden 35-40 residuos y tres a cuatro puentes disulfuro (Possani *et al.*, 1982; Debin *et al.*, 1993; Olamendi-Portugal *et al.*; 1996) Semejante a las toxinas de cadena larga, ellas pueden ser clasificadas en toxinas selectivas de insecto o mamífero, y tienen como blanco canales de potasio (Dreyer, 1990). El modo de acción de las toxinas de alacrán ha sido extensamente estudiado, usando venenos crudos o toxinas purificadas (Rochat *et al.*, 1979). Hasta el momento una colección grande de péptidos homólogos de alacranes ha sido identificada y se ha encontrado que algunos de estos bloquean dos clases de canales de K^+ los disparados por voltaje (tipo K_v) y los canales de K^+ de alta conductancia activados por Ca^{2+} (tipo BK). Estas toxinas inhiben en un rango de concentración de $1pM$ a $10nM$ y actúan sólo sobre canales de K^+ . Los venenos interrumpen la actividad neuronal mediante la producción de disparos repetitivos seguidos por depolarización de vida larga. Las toxinas contra canales de K^+ se unen bien a los estados abiertos y cerrados de los canales de K^+ . Estas toxinas podrían actuar como iniciadores de depolarización inmediata (Miller, 1995). La caribdotoxina y la noxiustoxina inhiben más de una clase de canal (varios canales de K^+ activados por Ca^{2+} y canales de K^+ dependientes de voltaje, la iberiotoxina parece ser un bloqueador selectivo del canal de K^+ activado por Ca^{2+} de alta conductancia que está presente en músculo y tejido neuroendócrino (García *et al.*, 1991). Es preciso mencionar que existen otras toxinas contra canales de Ca^{2+} , como lo son los dos péptidos aislados del veneno del alacrán *Pandinus imperator*, una de ellas llamada imperatoxina inhibidora e imperatoxina activadora, donde respectivamente la primera y la segunda bloquean o activan los receptores de rianodina de músculo esquelético y cardíaco (Valdivia *et al.*, 1992). Se ha demostrado la existencia de toxinas bloqueadoras de canales de cloro de pequeña conductancia entre las que se puede mencionar a la clorotoxina, la cual es un componente del

veneno del alacrán *Leiurus quinquestratus quinquestratus* (Dabin et al; 1993).

A N T E C E D E N T E S

En el grupo del Dr. Possani se han caracterizado alrededor de nueve clonas de DNAC que codifican para toxinas del veneno del alacrán *Centruroides noxius* Hoffmann (Becerril, 1993a; Tapia, 1994; Vázquez *et al.*, 1995), éstas se han denominado CngtI-CngtIX (nomenclatura arbitraria en la que las dos primeras letras corresponden al género y especie, las letras g y t corresponden al vector en que el inserto del DNAC correspondiente a cada clona fue subclonado, en este caso lambda gt 11 y el número romano hace referencia al orden de obtención de la secuencia nucleotídica con respecto a las otras). El análisis comparativo de estas secuencias con las toxinas de alacrán reportadas reveló que los DNAsc de *C. noxius* Hoffmann codifican para una familia de toxinas de alacrán muy similares (Fig 1). Los resultados de la secuencia del DNAC de la clona CngtVI (Fig 2) indican que se trata de un inserto de 346 nucleótidos y que codifica para un polipéptido precursor de 87 aminoácidos con un péptido señal de 19 aminoácidos y una secuencia de poli A de 16 nucleótidos. El polipéptido maduro es de 68 aminoácidos, esta clona al ser comparada con las secuencias nucleotídicas de ADNc conocidos de *C. noxius* presentó la mayor homología con la clona CngtIV (Fig 1) la cual fué utilizada como sonda radioactiva para aislada (Tapia, 1994, Vázquez *et al.*, 1995). Al comparar la secuencia de aminoácidos inferida de la clona de ADNc CngtVI (Fig. 2), con las secuencias de aminoácidos de las toxinas del mismo alacrán ya conocidas (Fig. 1), se observó que esta secuencia correspondía con la toxina II-14 (*Cn* 1) (Fig. 3). Becerril *et al.* (1993b) lograron caracterizar la región genómica que codifica para la toxina γ y del alacrán *Tityus serrulatus* la cual fue la primera secuencia genómica codificante para una toxina del veneno de alacrán, en ser aislada y caracterizada. Delabre *et al.* (1995); aislaron el gene que codifica para la toxina AaH I' del alacrán *Androctonus australis* Héctor. También fueron aislados y secuenciados los genes que codifican para las toxinas γ -b y γ -st de los alacranes *Tityus bahiensis* y *Tityus stigmurus* respectivamente (Becerril *et al.*, 1996). Se llevaron a cabo intentos de caracterización de genes del alacrán *Centruroides noxius* Hoffmann, pero hasta el momento de iniciar esta tesis no se habían obtenido resultados positivos.

OBJETIVO

Aislar y determinar la secuencia nucleotídica de la región genómica que codifica para la toxina *Cn 1* (II-14) del veneno del alacrán *Centruroides noxius* Hoffmann.

METODOLOGIA

1. Extracción del DNA del alacrán *Centruroides noxius* Hoffmann.

Esta se llevó a cabo mediante el maceramiento de los organismos completos a temperatura de congelamiento. Se utilizaron 5 gr. de estos organismos por cada proceso de preparación. El pulverizado se repartió en dos tubos (partes iguales) y se le agregó a cada tubo 25ml de la solución de extracción de DNA genómico (0.1M EDTA (Etilendinitrilotetracetato disódico), 0.05M Tris (Trihidroximetilaminometano) pH 8.0, 0.5 % SDS (Dodecilsulfato de sodio) y 100 microgramos/ml. de proteinasa K). Se mantuvieron los tubos a 55°C durante cuatro horas. Para llevar a cabo la eliminación de desechos celulares inicialmente se centrifugó a una temperatura de 4°C a 10,000 revoluciones por minuto (r.p.m.) durante 20'. El sobrenadante fué tratado mediante extracciones con fenol, fenol-cloroformo-alcohol isoamílico y cloroformo respectivamente. El sobrenadante fué puesto a dializar en cuatro litros de TE (Tris 10mM y EDTA 1mM) durante toda la noche; como el volumen resultante fue grande éste fue concentrado con butanol. Posteriormente se precipitó con acetato de sodio 3M a pH 4.8 y etanol absoluto, y se resuspendió en 1ml de agua tetradesilada.

2. Purificación del DNA de *C. noxius* Hoffmann.

Fué necesaria la eliminación de un pigmento de color rojizo que se encuentra unido al DNA de este alacrán, el cual se especula que pudiera tener una naturaleza de polisacárido con cargas positivas, por lo que se hizo necesario un paso de cromatografía de filtración molecular, para que sometido a alta fuerza iónica, el pigmento fuera separado. Se utilizó una columna de Bio-Gel A-50m malla 100-200 (Bio Rad, Richmond California.) y el buffer Tris 10mM pH 8, EDTA 1mM, NaCl 0.5M. La columna se cargó con un volumen de 500 microlitros del resuspendido del DNA genómico. El DNA apareció en las fracciones 8-12, éste fué precipitado y resuspendido en agua tetradesilada. Considerando la exigencia en las condiciones de pureza del DNA genómico, fueron necesarios otros pasos posteriores de purificación

después del procedimiento cromatográfico, como fueron: Centricón (centricon concentrators Amicon, Beverly USA), Glass Max (Glass Max DNA Isolation Spin Cartridge System, GIBCOBRL, USA) ó extracciones fenol-cloroformo. El DNA genómico se cuantificó por lectura en el espectrofotómetro a 260 y 280nm. Como el pico máximo de absorbancia de los ácidos nucleicos es a 280nm, y para la mayoría de las proteínas es de 280nm, la relación de las lecturas de una muestra de ADN a estas longitudes de onda nos indica la proporción en que una muestra de ADN está asociada con proteínas (Sambrook *et al*, 1989). La concentración del DNA se verificó por medio de electroforesis en gels de agarosa al 0.7%, donde también pudimos estimar la calidad del DNA.

3. Diseño de oligonucleótidos específicos

El diseño de los oligonucleótidos se llevó a cabo con base en la secuencia del DNAC *CngtVI* que codifica para la toxina *Cn* 1 (II-14), esta clona de ADNc fue obtenida en mi trabajo de tesis de licenciatura (Tapia, 1994, Vázquez *et al.*, 1995). La secuencia de esta clona se muestra en la Fig 2.

Se eligió específicamente esta clona de ADNc por ser la única en comparación con las ya secuenciadas, que presenta el mayor número de nucleótidos hacia el extremo 5', característica aprovechable para el diseño de un oligonucleótido que permitiera una amplificación específica, si se considera la posible localización de un intrón en esta región (ver antecedentes), (Becerril *et al*, 1993b, Delabre *et al*, 1995, Becerril *et al*, 1996). Los oligonucleótidos además de ser utilizados para la amplificación del DNA genómico por PCR, se utilizaron como sonda para la hibridación de los productos del mismo. Las secuencias de los oligonucleótidos utilizados en el presente trabajo se pueden observar indicadas en la fig. 2. A continuación se listan los oligonucleótidos, en donde se señala el número de nucleótidos y la temperatura de fusión (T_m) para cada uno de ellos:

$T_m = 59.9 + 0.41(\%GC) - 500 \# \text{ bases}$

Oligonucleótido	T_m (°C)	Tamaño (nucleótidos)	Posición	Orientación
Ctox Cn1	50.22	22	247	5'-3' de la cadena complementaria
Int.Cn1	57.45	20	15	5'-3' de la cadena codificante
5' Gen	51.3	20	-54	5'-3' de la cadena codificante
N Terminal	57.57	21	1	5'-3' de la cadena complementaria
Int2Cn1	61.47	21	19	5'-3' de la cadena complementaria.

4. Reacción en Cadena de la Polimerasa (PCR)

Esta se llevó a cabo mediante la utilización de oligonucleótidos específicos cuya secuencia como ya se mencionó corresponde a la de regiones específicas de la clona de ADNc CngtVI (Fig 2).

5. Transferencia por el Método de Southern (Sambrook *et al.*, 1989) de los productos de amplificación por PCR.

6. Ensayos de hibridación de los productos de amplificación que fueron transferidos. Se utilizó como sonda radioactiva la clona de ADNc CngtVI o bien, alguno de los oligonucleótidos específicos correspondientes a la secuencia de la misma clona.

7. Subclonación

Para llevar a cabo la subclonación se llevó a cabo la purificación de la banda específica, que hibridizó con la clona de ADNc y con los oligonucleótidos específicos y la reacción de ligación del inserto con el vector pUCBM20 el cual deriva de los plásmidos pUC18/19, y consta de 2722 pares de bases que incluye un sitio de clonación múltiple. Los productos de la PCR se subclonaron en el sitio de EcoRV. Para la transformación en células competentes se utilizó la cepa DH5 α las cuales se prepararon según Hiroaki and Hiroshi, 1990. Las clones recombinantes se seleccionaron a través de los métodos de α -complementación, análisis de restricción del DNA plasmídico (Sambrook *et al.*, 1989), PCR en colonia, hibridación en colonia e hibridación en punto.

8. Secuencia Nucleotídica

Se llevó a cabo la secuencia en doble cadena, utilizando el estuche de Secuensa (United States Biochemical, Cleveland, Ohio, U.S. and Canadá).

9. Análisis e Interpretación de la Secuencia.

Esto se hizo a través del estudio de la secuencia nucleotídica de las clones secuenciadas, así como la comparación de éstas con la secuencia de las clones de DNAc que codifican para las toxinas de *C. noxius* Hoffmann que ya se habían reportado, esto con el propósito de definir el aislamiento de la clones genómicas que codificaran para toxinas del veneno de *C. noxius* Hoffmann

RESULTADOS Y DISCUSION

REACCION EN CADENA DE LA POLIMERASA (PCR)

Para llevar a cabo los experimentos de la reacción en cadena de la polimerasa del DNA genómico de *Centruroides noxius* Hoffmann se utilizaron los oligonucleótidos cuyas secuencias corresponden a la clona CngtVI y se señalan en la Fig 2.

Para cada una de las reacciones de PCR se emplearon 5µg de DNA genómico. Cuando se utilizó una temperatura de alineamiento de 48°C y 25 ciclos, los oligonucleótidos 5' Gen y Ctox (estos se usaron debido a la posición de los mismos, en los extremos 5' y 3' de la clona de DNAC), se obtuvo una banda de 800 pares de bases, la cual hibridó con la clona de DNAC CngtVI y el oligonucleótido Int Cn1 marcados radioactivamente (Fig 4).

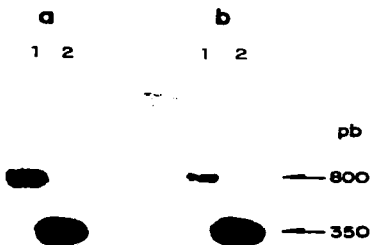


Fig 4. Autorradiografía en donde se muestra la hibridación del producto de PCR (oligos 5' Gen y Ctox) con un tamaño de 800 pares de bases. a) Hibridación con la clona de cDNA Cngt VI. En el primer carril se observa la banda de 800 pares de bases, en el segundo carril el control positivo, el cual corresponde al producto de PCR de la clona Cngt VI. b) Hibridación con el oligonucleótido Int Cn1. En el primer carril encontramos la banda de 800 pares, en el segundo carril el control positivo, el producto de PCR de la clona Cngt VI.

Después de estos resultados se llevaron a cabo otros intentos de PCR, sólo que variando las condiciones, como: la concentración de los oligonucleótidos utilizados y las condiciones de los programas de amplificación.

En estos experimentos de amplificación se utilizaron como primeros o iniciadores los oligonucleótidos 5'Gen y Ctox, se empleó una temperatura de alineamiento de 50°C y 30 ciclos y se obtuvo una banda de un tamaño de 350 p.b. aproximado al de la clona de cDNA CngtVI (Fig 5), pero ésta no hibridizó con la clona de DNAC CngtVI marcada radioactivamente.

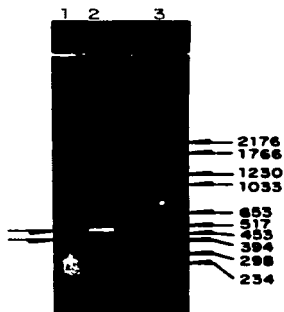


Fig. 5. Producto de amplificación por PCR con un tamaño aproximado al de la clona de DNAC (CngtVI). Electroforesis en gel de agarosa al 1%. En el carril 1 se observa el producto de PCR de la clona Cngt VI, en el carril 2 se observa el producto de PCR con los oligonucleótidos 5'Gen y Ctox; el carril 3 muestra el marcador VI de peso molecular de Boehringer.

Se incrementó el número de ciclos a 30 y con la misma temperatura de alineamiento de 52°C se obtuvo una banda de alrededor 1100 pares de bases, la cual mostró una señal de hibridización con la clona de DNAC CngtVI y el oligonucleótido específico IntCn1(Fig. 6) ambos marcados radioactivamente.

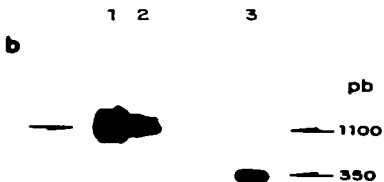
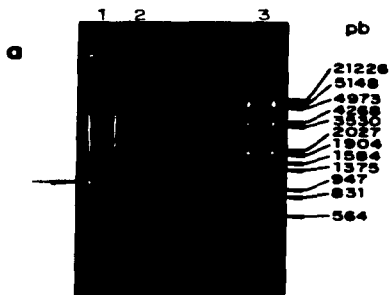


Fig 6. Producto de amplificación del DNA genómico con los oligonucleótidos 5' Gen y Ctox, con un tamaño de alrededor de 1100 pares de bases. a) Electroforésis en gel de agarosa 1%. En los carriles 1 y 2 se observa el producto de amplificación el cual presenta un tamaño de 1100 pares de bases; en el carril 3 el Marcador de Peso Molecular λ Hind III-EcoRI. b) Autorradiografía del producto de amplificación de 1100 pares de bases. En los carriles 1 y 2 se observan los dos productos de amplificación de 1100 pares de bases, en el carril 3 el control positivo, que corresponde al producto de PCR de la clona CngVI.

Se llevó a cabo la purificación de este producto de amplificación, así como la subclonación. Se seleccionaron las clonas recombinantes y se procedió con la secuenciación. La secuencia de este inserto de alrededor de 1100 pares de bases (Fig. 7) mostró casi un cien por ciento de homología con la región del péptido señal de la clona de DNAC CngtVI (Fig. 2); y se observó una interrupción en el aminoácido 15 del péptido señal, donde considerando la secuencia inmediata y que fue comparada con los extremos que definen los intrones para algunos arácnidos (Becerril *et al.*, 1993b) (Tabla 1.) puede tratarse de un intrón.

```

ACCATCGAAGTGAATTCCTGCTGATGATGACACAGCTCTGCTGCTGCTGATCGTGG 57
      M N S L L M I T A C L V L I
gtatgaactataatgatogaaattggtaataatttttaattggccataaatttttaaaagatga 128
aatctgggttatattatacattatattgtatttttgggaastogaaaacctgcagattatagcag 198
ttcttcatatgatgctaaatcgaattcattttccagaaatatttggtaaaaattatagcaactc 268
      1
agataaactgactctggtttcagaaaatagaccgatttgaataattgatgaaacgcataaatg 338
taftaaagctogataattacgatatgaactgcgccgcaagaaggtattcttgcctaccgggggat 407
ctcagagtoactaatcttcatatttttaacttatttcttactacttcaacttcttatttaaggttg 478
      2
acctttaaagctaaattgttttaattgacacactgaccagtgaggctgttttttaagccgctgctgaa 548
tatgttgacctgtataattgagtggttaattattttgatatagtgattgctattgtatttggaaatt 620
ttcagtttttccgctacttctatgctgpgcgaatagttgacctgctaattgagtggaatttattgaa 690
gtgatatttattgatttggaaatttttagtttccctcacttctatggttggtgacttgagcaagtgag 761
tttaagaataaacatcagaaaattogagatttcaatgaaattgcatctggtgaagettccaattataat 831
gtctgttctggtagatagctgocaaataaaacaaagattgggaaccgacaccttaccocctgtttta 900
      3
aaatagcaattatttaaatggactgttttctgttttaacttacttctgtctcatalataataatactttt 970
tgcacacttttgatgctacagctataaaattttaaattttatgtgttaataagttatctgtttgaagt 1041
gacaaagcaatttttggatggcagaaatcagctgttggcga
      T V W A K E G Y L
GTGGACGCCAAAGGCGCTGCAAA 1101
      V D A K G C K 1122

```

Fig. 7. Secuencia nucleotídica de la clona de alrededor de 1100 pares de bases. El primero y segundo exones se encuentran en letras mayúsculas. La región subrayada en letras mayúsculas corresponde a la región que codifica para el péptido señal. Con letras minúsculas encontramos la región del intrón. Las secuencias que se observan en letras minúsculas y subrayadas, así como señaladas con los números en negritas corresponden a los oligonucleótidos que se diseñaron para avanzar en esta secuencia. 1) oligonucleótido sec1 cn1, 2) oligonucleótido sec2 cn1, y 3) oligonucleótido sec3 cn1.

Tabla 1. Comparación de bordes intrónicos de diferentes genes de arácnidos (ver referencias en el pie de la tabla)

secuencia de bordes intrónicos			
Intrón	Extremo 5'.....	tamaño del intrón....	Extremo 3'
1	gtaagtcctca.....	6.1 kpb	tttttttttag
2	gtaagttacat.....	8.6 kpb	ttttttttatag
3	gtaagttccaa.....	6.3 kpb	gtaacctctag
4	gtaagttatgt.....	2.1 kpb	tatttttttag
5	gtaagtagtc.....	5.0 kpb	tcattcccgag
6	gtattttctgc.....	6.9 kpb	catttttacag
7	gtatgtttaag.....	14.3 kpb	tttatttttag
8	gtaagtgcta.....	2.7 kpb	cttttttcag
9	gtgagtataa.....	963 pb	tttggttcag
10	gtgagtatga.....	1,019 pb	tttggttcag
11	gtaagattta.....	425 pb	ctgactacag
12	gtaagctgaa.....	475 pb	gttaacatag
13	gttaggtccc.....	474 pb	gttaacatag
14	gttaggtccc.....	465 pb	attaacatag
15	gtaagatttt.....	347 pb	taaatatttag
16	gttaggtccc.....	58 pb	taataatcag
17	gtaagtttga.....	60 pb	aaattgctag
18	gtatgttttg.....	60 pb	ttatttttag
19	gtaagtttga.....	65 pb	aaaatgcttag
20	gtacaaaatg.....	60 pb	catttttttag
21	gtaagtttga.....	65 pb	aaatttatag
22	gtaagtttgc.....	61 pb	atttttttag
Consenso	gtaagttana		naatttttag

Tabla 1. Comparación de bordes intrónicos de diferentes genes de arácnidos. Comparación de bordes 5' y 3' de intrones de diferentes genes de arácnidos. Solamente los primeros 10 y los últimos 10 nucleótidos de los intrones se muestran. Consensos propuestos para bordes de intrón 5' y 3' se muestran en la parte inferior de esta tabla (n significa cualquier nucleótido). Los intrones 1-8 corresponden al gene que codifica la subunidad de hemocianina de la tarántula *Eurypelma californicum* (Voll et al, 1990). Los intrones 12, 13 y 14 corresponden a los genes que codifica la toxina gamma de los alacranes sudamericanos *Tityus serrulatus*, *T. stygurus* and *T. bahiensis* respectivamente (Becerril et al; 1993b; Becerril et al; 1996). El intrón 15 corresponde al gene que codifica para la toxina IV-5 de *Tityus serrulatus* (Corona et al, 1996). Los intrones 16-22 corresponden a los genes que codifican cisteína proteasas de los ácaros *Dermatophagoides farinae* (intrones 16-18), *Euroglyphus maynei* (intrones 19-20) y *Dermatophagoides pteronyssinus* (intrones 21-22) (Kent et al;). Los intrones 9 y 10 Cn1, este estudio. El intrón 11 corresponde al de la toxina AaH1 del alacrán *Androctonus australis* Héctor (Delabre et al, 1995).

Sin embargo se terminó con la secuencia de la clona, donde para llevarla a cabo, hubo que diseñar tres oligonucleótidos a partir de las secuencias más distales de las que se fueron obteniendo y esto nos permitió avanzar con la secuencia. El análisis de esta secuencia reveló 1122 nucleótidos totales, en donde 1019 nucleótidos corresponden al intrón, se observa parte de la región genómica que codifica para la región del amino terminal de la toxina y se trunca en el codón que estaría codificando para la tercera lisina del polipéptido maduro en esta región. Esta amplificación se llevó a cabo debido a una probable contaminación del oligonucleótido Ctox con el oligonucleótido Int2Cn1 (ver la localización de los oligonucleótidos en la Fig. 2), y esto porque el segundo oligonucleótido se había estado utilizando con el objeto de amplificar en la región hacia el extremo 5' cercana al péptido señal, esto si se consideran los antecedentes del presente proyecto (Becerril *et al* ; 1993b; Becerril *et al* ; 1996; Corona *et al* ; 1996; Delabre *et al* ; 1995). Al efectuar la comparación de esta secuencia se propone que codifica para una toxina muy semejante a la II-14 (Cn 1), ya que existe la diferencia de que en el polipéptido maduro y en la clona de cDNA que codifica para esta toxina CngtVI, el segundo aminoácido es aspártico y en la región genómica esta posición la ocupa un glutámico.

Considerando estos resultados, hubo que volver a iniciar con la búsqueda de un inserto nuevo; entonces, utilizando 10 microgramos de DNA genómico, 60 pmoles de cada uno de los oligonucleótidos utilizados como primeros o iniciadores para la amplificación 5' Gen y Ctox, se utilizó una temperatura de alineamiento de 57°C y 30 ciclos, y se obtuvo un producto de amplificación de alrededor de 1300 pares de bases., este hibridizó con la clona de DNAc marcada radioactivamente y con un oligonucleótido específico de la región amino terminal llamado N-terminal de la clona CngtVI (Fig. 8).

Se llevó a cabo la subclonación y la secuenciación del inserto, para lo cual se utilizaron los tres oligonucleótidos que correspondieron a la clona de 1100 pares de bases, estas secuencias se localizaron dentro del intrón de la clona de 1100 pares de bases, lo cual puso en evidencia la homología entre esta clona de alrededor de 1300 pares de bases (ver fig. 9) y la clona de 1100 pares de bases que ya había sido secuenciada (Fig. 7).



Fig. 8. Producto de amplificación por PCR, con los oligonucleótidos 5' Gen y Ctox, con un tamaño de aproximadamente 1300 pares de bases. a) Electroforésis en Gel de Agarosa 1%. En los carriles 1 y 2 se puede observar de acuerdo al orden de migración en el gel, la banda 1 de 1100 pares de bases y la banda 2 de 1300 pares de bases. El carril 3 corresponde al Marcador ϕ x174/Hae III. b) Hibridización en Punto de las bandas 1) de 1100 pares de bases y 2) de 1300 pares de bases. Esta hibridización se llevó a cabo con el oligonucleótido N terminal ver la secuencia del oligonucleótido en la fig. 2.

CCATCGAAGATGATTCCTGTTGTTGATGATCACTGCTGTTGTTGGTCGGGATCGGGg	68
M N S L L M I T A C L V L I	
ataaactataaatacgaactggggaatnnaaattgaatggcataaattattaaagctatgaact	129
cggttatataatacactataatgataatgggggaatgaaaacaclogcagattatagaactgac	201
ttaataataagcttaaacgcgaactcaattacagaatattgggtggaatataaaactcagat	271
1	
aaactaataatgatgattcagaactagaacagattgataaattctogaactgcataaattctat	341
aaagtgaattattacgattatgagactgcactgactggtggatggcaatgacacaaagggca	410
tcggcaagaaggattctgacaggaggaatactacagagctcactaatctattatatactctat	480
tcctctactctactctaaactctattaaagttgaccctgttaagctaaattggaaatgatgc	551
2	
tactacttaacagtgaggctgattataaagtgaatattgggaccctgctaaattgagtggaatt	622
atttgatattgtaattgattattgattatagaatttttaagatttgcactctatgtgtgact	694
taagcaagtgagttatagaaataacactataaactcaaatfacaatgcaattgcattcgaagatttc	764
attatagctgctctcctggctagtaagctggcaataaatacaagaatacagaacgacaglatccocct	833
agtttaaaatgcaatattataacaggactattttagtttataactactatgttgtaacaataat	904
3	
atacttttggcacactttgagtgcaataogataaattgtaattattgtgtaataagttatgcttt	976
tgaagtgcaacagactattctgattgtagcagaAACAGTGTGGGGCAAGGAAGGTTA	1040
T V W A K E G Y	
TCTGTGTGGAACGGAAAGGGCTGCAAAAAGAATTGTATAAATTGGGAAAAAACGAT	1095
L V D A K G C K K N C Y K L G K N D	
TATTGCAATAGGAAATGACAGAAATGAACAACCGAGGAGGTAGTACGGCTATTGCT	1150
Y C N R E C R M K H R G S Y G Y C	
ACGGATTTGGTCTATTGTGAAAGATTGTCCGATAGTACACCGACTTGGCCOCTT	1205
Y G F G C Y C E G L S D S T P T W P L	
CCTAATAAAACATGACGGCGAAATTA TGGCAATGACTTTTATTGTGCCACCA	1261
P N K T C S G K E N D	
CAGAAATAGTGTAAACGCTTTTAAATTGCAAGT	1293

Fig. 9. Secuencia nucleotídica de la clona de 1300 pares de bases. Las secuencias en letras mayúsculas corresponden a los exones 1) y 2). La región subrayada en mayúsculas corresponde al péptido señal. La secuencia en letras minúsculas corresponde al intrón. Las secuencias subrayadas en letras minúsculas corresponden a los oligonucleótidos 1) sec1 on1, 2) sec2 on1, 3) sec3 on1.

Este inserto de tamaño mayor, después de ser secuenciado, mostró tener 1292 pares de bases, con un intrón de 963 pares de bases que interrumpe la región del péptido señal en el aminoácido 15 de la misma manera que en la clona de 1100 pares de bases. Es preciso mencionar que hasta el momento se trata de la secuencia intrónica de mayor

tamaño, junto con el intrón de la clona de 1122 pares de bases, que han sido encontradas en genes que codifican para una toxina del veneno de alacrán. De la secuencia genómica pudo derivarse la secuencia de aminoácidos de la proteína para la que ésta codifica. Se llevó a cabo la comparación de esta secuencia de aminoácidos con las secuencias de aminoácidos de las toxinas de *C. noxius* Hoffmann ya caracterizadas, y se observó que comparte el mayor porcentaje de homología con la toxina Cn 1 (II-14), a excepción de algunas diferencias encontradas en la región del péptido señal ver Fig. 9, donde el aminoácido once que es leucina es codificado por un codón cuya tercera base cambia de citosina a guanina, esto no origina el cambio de aminoácido, considerando la degeneración del código genético; pero en el aminoácido trece sí existe un cambio de leucina a valina. Sin embargo también se encontró una diferencia en la región estructural, la cual es aún de mayor importancia. Esta se localiza en el segundo aminoácido del polipéptido maduro, donde en la secuencia primaria de aminoácidos de la toxina Cn 1 (II-14) y en la secuencia que corresponde a la clona de DNAC CngtVI que codifica para esta toxina, aparece como ácido aspártico, y en la secuencia genómica esta misma posición es ocupada por ácido glutámico, lo cual nos induce a pensar que se trata de la región genómica que estaría codificando para una variante de la toxina Cn 1 (II-14). Por todo lo anterior es preciso mencionar que la región genómica codificante para la variante de la toxina Cn 1 (II-14) presenta dos exones, donde el primero ocupa la posición 1-43, y el segundo la posición 1017-1292; y un intrón de 963 pares de bases. En el laboratorio se han obtenido algunos polipéptidos que se encuentran juntos con la fracción que contiene a la toxina II-14 (Cn1), estos son péptidos minoritarios, los cuales pudieran ser variantes de esta toxina, hasta ahora se ha llevado a cabo la secuenciación de algunos de estos, pero considerando que esta secuencia corresponde sólo a algunos residuos de las mismas, no es posible poder definir hasta el momento si las clonas genómicas de alrededor de 1100 pares de bases y la de 1300 pares de bases codifican para estos péptidos; para ello todavía es necesario un estudio exhaustivo de la secuencia de los mismos.


```

1100      620      630      640      650      660      670
tatggaaatttttcagtatcttcccgctcaactctatgctggcgaaatcgttgacctgctaa
111 1 1 1 1
1300.e tatcttagtattgtgattgctatttgattatagaaatttttaagtattctgagtcacct
      680      690      700      710      720      730
ttgagtgaaattttatctagaagtgatgatttgatttatggaatttttttagtcttccccgt
1300 ctctgtgttgccttaagcaagtgagttaatagaaacatcataaattcaaatattcaat
      740      750      760      770      780      790
cactctatgtggttgacttgagcaagtgagtttaatagaaataacatcagaaaattcgag
1300 gaattgcatctctgaagattttcattatagctgttctcccggactcaaaagtcgccaata
      800      810      820      830      840      850
attcacatgaaattgcatctcgtggaaagattttcaattataaactcgtttctctggtagata
1300 aatacaagatcagcaacgacagctatccccctcagtggtttcaaatagcactatttatactc
      860      870      880      890      900      910
agtgcacaataaaaaaagatttaggaacgacacagctatccccctcagtggtttcaaatagcat
1300 atttataatcaggactatttttagtctttttatactctcctgtgtgtacaaataataatacttt
      920      930      940      950      960      970
atttataataggactgtttctcgttttttatactctcctgtgtcatatataataatactttt
1300 ttttgacacatcttgagtgcatcagattaaatgtttttaattttattgtgttbaaagtat
      980      990      1000      1010      1020      1030
tttgacacatcttgagtgcatcagattaaatatttttaattttattgtgttbaaagtat
1300 gctttttgaggttgaaacagactattctctttgtatcttctgaggaagAACAGTGTGGCCAAA*
CngtVI T V W A K
AACAGTGTGGCCAAA*
T V W A K
1100      1040      1050      1060      1070      1080      1090
cgtttttgaggttgaaacagcatattctctttgtattcttctgaggaagAACAGTGTGGCCAAA*
1300 GGAAGGTTATCTGGTGGACGCCAAAGGCGCTGCCAAAAGAATTGCTATAAAATGGGAAAAMAAAC
E G Y L V D A K G C K K N C Y K L G K N
CngtVI GGACGGTTATCTGGTGGACGCCAAAGGCGCTGCCAAAAGAATTGCTATAAAATGGGAAAAMAAAC
D G Y L V D A K G C K K N C Y K L G K N
      1100      1110      1120      1130      1140      1150
GGAAGGTTATCTGGTGGACGCCAAAGGCGCTGCCAAA
E G Y L V D A K G C K
1300 GATTAATTOCAATAAGGAAATOCAGAAATGAAACACCGAGGAGGTAGTTACCGCTTATTOCTAC
D Y C N R E C R M K N R G G S Y G Y C Y
CngtVI GATTAATTOCAATAAGGAAATOCAGAAATGAAACACCGAGGAGGTAGTTACCGCTTATTOCTAC
D Y C N R E C R M K N R G G S Y G Y C Y
      1160      1170      1180      1190      1200      1210
GGAATTOGGTGCTATTGTGAAGGATGTGCTCGATAGTACACCGACTTGGCCCTTCCCTAAT
G F G C Y C E G L S D S T P T W P L P N
CngtVI GGAATTOGGTGCTATTGTGAAGGATGTGCTCGATAGTACACCGACTTGGCCCTTCCCTAAT
G F G C Y C E G L S D S T P T W P L P N

```



```

          1220      1330      1240      1250      1260      1270
1300  AAAACATGCAGCGGAAAATAATGGCAATGACTTTTTATTGTCCACCAACAGAAATAGTGT
      K T C S G K End
CngtVI AAAACATGCAGCGGAAAATAATGGCAACGACTTTTTATTGTCCACCAACAGAAATAGTGT
      K T C S G K End

          1280      1290
1300  AACGCTTTTAAATGCAAGT
CngtVI AACGCTTTTAAATGCAAGTAAAAAAAAAAAAAAAAAAAA

```

FIG. 10. Alineamiento de secuencia nucleotídica de las clonas de 1100 pares de bases, 1300 pares de bases y CngtVI. El alineamiento entre las clonas de 1100 y 1300 pares de bases, se llevó a cabo con el programa Fasta localizado en el paquete de software (Genetics Computer Group, GCG) de la universidad de Wisconsin (Madison, WI, U.S.A) de análisis de secuencias. Es observable como el programa alinea en la primera parte de los intrones las regiones de mayor homología, y también se muestra la diferencia de tamaño entre los intrones. La primera parte de las secuencias nucleotídicas que codifican para el segundo exón en las tres clonas, se encuentran señaladas con un punto en negrilla a la derecha del renglón donde se localizan, permitiendo evidenciar la homología de las mismas. Los sitios donador y aceptor de los intrones propuestos por los programas de análisis de secuencia de DNA para la clona de 1300 pares de bases, se encuentran subrayados. Los sitios de poliadenilación se señalan con asteriscos y con el número 1 y 2 en negrilla.

Se utilizó el programa BCM Gene Finder para análisis de secuencias nucleotídicas de DNA, éste contiene los programas Search for Potential Splice Sites (HSPL) y Recognition of 3'-end Cleavage and Polyadenilation Region (POLYAH) los cuales fueron empleados y nos permitieron obtener información complementaria de la clona de 1300 pares de bases, la cual se menciona a continuación:

1.- La presencia de dos posibles intrones de la secuencia de 1300 pares de bases, en donde se definen dos sitios donadores y dos sitios aceptores (ver Fig. 10). Estos permitirían especular acerca de un probable procesamiento diferencial o "splicing" diferencial de la región genómica, aunque no se han aislado hasta el momento toxinas que apoyen el hecho de "splicing" (empalme) diferencial. Pero cabe mencionar que uno de los dos intrones que muestra este análisis de la secuencia, parece funcionar como tal.

2.- Inexistencia de secuencias repetidas.

Este análisis reveló que en la secuencia nucleotídica no existen secuencias repetidas.

3.- La presencia de dos sitios posibles de poliadenilación (ver Fig. 10).

Además se empleó el programa Fasta para efectuar la comparación

entre las clonas de 1100 pares de bases y 1300 pares de bases, y esto dió como resultado la identidad entre ambas de un 75.5% en 642 bases sobrelapantes, localizadas fundamentalmente en la primera parte de las secuencias, que formarían parte de lo que corresponde a las regiones del péptido señal y la primera parte del intrón (ver Fig. 10). Esto nos induce a pensar en la existencia de una serie de familias de genes que están codificando para variantes de toxinas, las cuales como se ha observado presentan un alto porcentaje de homología. Cabe mencionar que a diferencia de lo sucedido en los genes que codifican para las toxinas bloqueadoras de canales de Na⁺ del género *Tityus*, tales como la γ de *T. serrulatus*, la γ -st de *T. stigmurus* y la γ -b de *T. bahiensis* en donde el primer exón es 100% idéntico (Becerril *et al* : 1996), en las clonas de 1300 pares de bases y la de 1100 pares de bases, no ocurre lo mismo como anteriormente ya se había analizado; aunque para ello habría que pensar que la clona de 1100 pares de bases es un gene truncado. Sin embargo los límites de secuencia nucleotídica del extremo 5' de ambos intrones son prácticamente iguales si se consideran las diez primeras bases; esto es a excepción de la base 9, donde en la clona de 1300 pares de bases hay un nucleótido adenina y en la de 1100 pares de bases el nucleótido es guanina. Al comparar la secuencia conocida que codifica para la primera parte de la región amino terminal de la toxina, en la clona de 1100 pares de bases con la clona de 1300 pares de bases, se observa que es idéntica, aunque presentan diferencias en el primer exón, pero solo en la región codificante para el péptido señal, los intrones de ambas clonas también presentan diferencias, tanto en tamaño como en secuencia, por todo esto no es posible definir aún si la región codificante para el polipéptido maduro es la misma o bien difiere y codifica para toxinas distintas; para ello es indispensable la obtención de la secuencia completa de esta clona de alrededor de 1100 pares de bases.

CONCLUSIONES

. A partir de la reacción en cadena de la polimerasa (PCR) fue posible aislar y determinar la secuencia nucleotídica de la región genómica que codifica para una variante de la toxina Cn 1 (II-14). Esta es la primera unidad génica transcripcional que codifica para una toxina de un alacrán mexicano, caracterizada hasta el momento, en este caso de *Centruroides noxius* Hoffmann, y es la clona completa que contiene el intrón de tamaño mayor.

. La unidad génica transcripcional codificante para esta variante de la toxina Cn 1 (II-14), está compuesta por dos exones y un intrón, sujetándose al modelo de organización propuesto (Corona *et al.*, 1996) para los genes que codifican para toxinas del veneno de otras especies de alacrán que hasta el momento ya han sido aislados y caracterizados (Becerril *et al.*, 1993; Delabre *et al.*, 1995; Becerril *et al.*, 1996; Corona *et al.*, 1996).

. Considerando los datos obtenidos de los programas de análisis de secuencia de DNA, podría proponerse un procesamiento diferencial para estos genes, ya que al observar las secuencias límite 5' y 3' de los intrones que se indican, uno de ellos es aparentemente funcional, y el otro aunque no se ha demostrado su funcionalidad, podría especularse su uso potencial, o bien que en algún momento del proceso evolutivo de la especie, este cumplió una función como tal.

. Es preciso definir la existencia de familias de genes que codifican para una serie de variantes de toxinas del veneno de alacrán, esto a partir de los resultados de la comparación de secuencia efectuada entre la clona de 1100 pares de bases y la clona de 1300 pares de bases.

PERSPECTIVAS FUTURAS

Los resultados obtenidos en el presente trabajo dan la pauta para proseguir con investigaciones futuras que permitirán el avance en el conocimiento acerca de la regulación y expresión de los genes que codifican para las toxinas del veneno de alacrán en ésta y otras especies, para ello de manera inmediata se propone:

- . La demostración con resultados concluyentes de la existencia de familias que codifican para una serie de variantes de toxinas del veneno de alacrán, esto, através del aislamiento y caracterización de la clona completa de 1100 pares de bases.
- . El mapeo de la región promotora del gene que codifica para la variante de la toxina Cn 1 (II-14).
- . Llevar a cabo amplificaciones con oligonucleótidos específicos de las regiones variables.
- . La definición de un arreglo posible de los genes en tandem.
- . La búsqueda de otros genes que codifiquen para toxinas de la misma especie, con especificidad diferente, utilizando la secuencia aislada y caracterizada en este trabajo, con el propósito de la comparación y análisis de las mismas, utilizando para ello como patrón de comparación el resultado del análisis en secuencias genómicas como las correspondientes a las especies del género *Tityus* (Becerril et al, 1996).

REFERENCIAS

- . Alagón, A., Dehesa, M. and Possani, L. 1989. Monograph Scorpions from the Genus *Centruroides*., O.M.S., Ginebra, Suiza. 1-20.
- . Becerril, B., Vázquez, A., García, C., Corona, M., Bolívar, F and Lourival D. Possani. 1993a. Cloning and characterization of cDNAs that code for Na⁺-channel- blocking toxins of the scorpion *Centruroides noxius* Hoffmann. *Gene*. 128, 165-171.
- . Becerril, B., Corona, M., Mejía, M., Martín, B., Lucas, S., Bolívar, F. and Lourival D. Possani. 1993b. The genomic region encoding toxin gamma from the scorpion *Tityus serrulatus* contains an intron. *Febs Letters*. 335. 1, 6-8 (b).
- . Becerril, B., Corona, M., Coronas, F., Zamudio, F., Calderón, E., Fletcher, P., Martín, B., and Lourival D. Possani. 1996. Toxic peptides and genes encoding toxin γ of the Brazilian scorpions *Tityus bahiensis* and *Tityus stigmurus*. *Biochem. J.* 313. 753-760.
- . Catterall, W.A. 1984. The molecular basis of neuronal excitability. *Science*. 223, 653-61.
- . Corona, M; Zurita, M; Possani, L and B. Becerril (1996) Cloning and Characterization of The Genomic Region Encoding Toxin IV-5 From The Scorpion *Tityus serrulatus* Lutz and Mello. *Toxicon*.34.2. 251-256.
- . Couraud, F., Jover, E., Dubois, J.M. and H. Rochat. 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.
- . Couraud , F., E, Jover. 1984. Mechanism of Action of Scorpion Toxins. *Handbook of Natural Toxins*, Marcel Dekker, New York. 2, 659-671.
- . Chhatwal, G.S. and E. Habermann. 1981. Neurotoxins, protease inhibitors and histamine releasers in the venom of the indian red scorpion (*Buthus tamulus*). Isolation and partial characterization. *Toxicon*. 19, 807-823.
- . Debin, J., Maggio, J., and Strichartz, G. 1993. Purification and characterization of chlorotoxin, a chloride channel ligand from the venom of the scorpion. *Am. J. Physiol.* 264. C361-C369.

- . Delabre, M., Passero, P., Marilley, M., and P.E. Bougis. 1995. Promoter Structure and Intron-Exon Organization of a Scorpion α -Toxin Gene . *Biochemistry*. 34. 6729-6736.
- . Dent, M., Possani, L.D., Ramírez, G.A. and P.Fletcher Jr. 1980. Purification and Characterization of Two Mammalian Toxins from The Venom of The Mexican Scorpion *Centruroides noxius* Hoffmann. *Toxicon*. 18. 343-350.
- . Dreyer, F. 1990. Peptide toxins and potassium channels. *Reviews of Physiology, Biochemistry and Pharmacology* 115, 93-136.
- . Eitan, M., Fowler, E., Hermann, R., Duval, A., Pelhate, M and E.A. Zlotkin. 1990. Scorpion Neurotoxin Paralytic to Insects That Affects Sodium Current Inactivation: Purification, primary Structure, and Mode of Action. *Biochemistry*. 29, 5941-5947.
- . García, M., Galvez, A., García, M., Frank, V., Vázquez, J., and Kaczorowski, J .1991. Use of Toxins to Study Potassium Channels. *Journal of Bioenergetics and Biomembranes*. 23,4. 615-646.
- . Ghazal, A., Ismail, M., Abdel-Rahman, A.A. and M.F. El-Asmar. 1975. Pharmacological studies of scorpion (*Androctonus amoreuxi* Aud And Sav) venom. *Toxicon*. 13, 253-259.
- . Goyffon, M et Chippaux, J.P. 1990. Animaux venimeux terrestres, Editions techniques, Encycl. Méd. Chir. Paris France. Intoxication, pathologie du travail 16078 A¹⁰ 14-1990, 14p.
- . Kent, N., Hill, M.R., Holland, P.W. and Hart, B.J. Dept. Biochem. Univ. Oxford, South Parks Road, Oxford OX1 3QU, UK.
- . Kharrat, R., Darbon, H., Rochat, H. and Granier, C. 1989. Structure activity relationships of scorpion alfa-toxins. Multiple residues contribute to the interaction with receptors. *European Journal of Biochemistry* 181, 381-90.
- . Keegan, H.L. 1980. Scorpions of medical importance. Jackson, USA, University Press of Mississipi, 140pp.
- . Kopeyan, C., Martínez, G., Lissitzky, S., Miranda, F. and H. Rochat. 1974. Disulfide bonds of toxin II of the scorpion *Androctonus australis* Hector. *European Journal of Biochemistry*. 47, 483-9.

- . Lester, D., Lazarovici, P., Pelhate, M. and E. Zlotkin. 1982. Purification, Characterization and Action of Two Insect Toxins From The Venom of The Scorpion *Buthotus judaicus* . *Biochim. Biophys. Acta* 701, 370-381.
- . Liang, P and Pardee, A (1992) Differential Display of Eukaryotic Messenger RNA by Means of the Polymerase Chain Reaction. *Science*. 257. 967-970.
- . Liang, P, Averboukh, L and A.B. Pardee (1993) Distribution and cloning of eucaryotic mRNAs by means of differential display: refinements and optimization. *Nucleic Acids Research*. 21: 14. 3269-3275.
- . Ménez, A., Bontems, F., Roumestand, Ch., Gilquin, B. and F. Toma. 1992. Structural basis for functional diversity of animal toxins. *Proceedings of the Royal Society of Edinburgh*. 99, 83-103.
- . Miller, Ch. 1995. The Charybdotoxin Family of K⁺ Channel-Blocking Peptides. *Neuron*. 15. 5-10.
- . Miranda, F., Rochat, H., S, Lisitzky. 1962. Propriétés échangeuses d'ions du gel de dextrane (Sephadex), Application à la mise au point d'une technique de rétention réversible des protéines basiques de faible poids moléculaire. *J. Chromatogr.*, 7, 142-154.
- . Miranda, F., Kopeyan, C., Rochat, H., Rochat, C., and S, Lisitzky. 1970. Purification of animal neurotoxins. Isolation and characterization of eleven neurotoxins from the venoms of the scorpions *Androctonus australis* Hector, *Buthus occitanus tunetanus* and *Lelurus quinquestriatus quinquestriatus*. *Eur. J. Biochem.* 16, 514-523.
- . Olamendi-Portugal, T; Gómez-Lagunas, F; Gurroia, G.B; and Posani, L. 1996. A novel structural class of K⁺-channel blocking toxin from the scorpion *Pandinus imperator*. *Biochem. J.* 315, 977-981.
- . Pelhate, M. and Zlotkin. 1981. Voltage-dependent slowing of the turn off of Na⁺ current in the cockroach giant axon induced by the scorpion venom "insect toxin". *Journal of Physiology (London)* 319, 30P-31P.
- . Pelhate, M and E. Zlotkin. 1982. Actions of Insect Toxin and Other Toxins Derived From The Venom of The Scorpion *Androctonus australis* on Isolated Giant Axons of The Cockroach (*Periplaneta americana*). *J. Exp. Biol.* 97, 67-77.

- . Possani, L.D., Dent, M.A.R., Martin, B.M., Maelicke, A. and I, Svendsen. 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; Svendsen, I.B. 1982. The Primary Structure of Noxustoxin: A K⁺ Channel Blocking Peptide, Purified From The Venom of The Scorpion *Centruroides noxius* Hoffmann. Carlsberg Res. Commun. 47, 285-289.
- . Possani, L.D. 1984. Structure of Scorpion Toxins. En: A. Tu (Ed), Handbook of Natural Toxins. Marcel Dekker, New York. 2, 513-550.
- . Rochat, H., Bernard, P and Couraud, F. 1979. Scorpion Toxins: Chemistry and Mode of Action. Advances in Cytopharmacology. 3, 325-334.
- . Sambrook, J., Fritsch, E., T, Maniatis. 1989. Molecular Cloning A Laboratory manual. Cold Spring Harbor laboratory Press, Segunda Edición. Estados Unidos.pp.
- . Sissom, W.D. 1990. Systematics, biogeography and paleontology. In the biology of scorpions. ed. Polis, G.A. Stanford University Press. 65-160.
- . Tapia, J.V. 1994. Determinación de la Secuencia Nucleotídica de cDNAs que Codifican Para Toxinas del Veneno del Alacrán *Centruroides noxius* Hoffmann. Tesis de Licenciatura. Facultad de Ciencias. Universidad Autónoma del Estado de México. México. 50pp.
- . Valdívia, H., Kirby, M., Lederer, J and Coronado, R. 1992. Scorpion toxins targeted against the sarcoplasmic reticulum Ca²⁺ - release channel of skeletal and cardiac muscle. Proc. Natl. Acad. Sci. 89. 12185-12189.
- . Vázquez, A., Tapia, J.V., Ellison, W., Martin, B., Lebreton, F., Delepiere, M., Possani, L.D and Baltazar Becerril. 1995. Cloning and Characterization of the cDNAs Encoding Na⁺ Channel-Specific Toxins 1 and 2 of the Scorpion *Centruroides noxius* Hoffmann. Toxicon. 33,9. 1161-1170.
- . Voll, W., and Volt, R. (1990) Proc. Natl. Acad. Sci. USA 87, 5312-5316.
- . Walther, C., Zlotkin, E and Rathmayer, W. 1976. J. Insect Physiol. 22, 1187-1194.

. Zlotkin, E. and A. Shulov. 1969. A simple device for collecting scorpion venom. *Toxicon* 7, 331-332.

. Zlotkin, E., Miranda, F., Kopeyan, C. and S. Lisitzky. 1971. A new toxic protein in the venom of the scorpion *Androctonus australis* Hector. *Toxicon* 9, 9-13.

. Zlotkin, E; Miranda, F; and H. Rochat. 1978. Chemistry and Pharmacology of Buthinae scorpion venoms. en *Handbook of Experimental Pharmacology*. 48. 317-369. Ed. Sergio Bettini, Spinger verlag. Berlin.

ANEXO

Publicación en Revista Internacional

Vázquez, A., Tapia, J.V., Elason, W., Martín, B., Lebreton, F., Delepiere, M., Possani, L.D. and Baltazar Becerril. 1995. Cloning and Characterization of the cDNAs Encoding Na⁺ Channel-Specific Toxins 1 and 2 of the Scorpion *Centruroides noxius* Hoffmann. *Toxicon*. 33,9. 1161-1170.



CLONING AND CHARACTERIZATION OF THE cDNAs ENCODING Na⁺ CHANNEL-SPECIFIC TOXINS 1 AND 2 OF THE SCORPION *CENTRUROIDES NOXIUS* HOFFMANN

ALEJANDRA VAZQUEZ,¹ JUANA V. TAPIA,¹ WILLIAM K. ELIASON,²
BRIAN M. MARTIN,² FLORENCE LEBRETON,³ MURIEL DELEPIERRE,³
LOURIVAL D. POSSANI^{1*} and BALTAZAR BECERRIL¹

¹Department of Molecular Recognition and Structural Biology, Biotechnology Institute, National Autonomous University of Mexico, Avenida Universidad, 2091, Apartado Postal 510-3, Cuernavaca, Morelos 62271, Mexico; ²National Institute of Mental Health, Unit on Molecular Structure, Clinical Neurosciences Branch, Building 49 B1EE16, Bethesda, MD 20892, U.S.A.; and ³Institut Pasteur, 25 Rue de Dr Roux, 75724 Paris Cedex 15, France

(Received 20 January 1995; accepted 20 March 1995)

A. Vazquez, J. V. Tapia, W. K. Eliaison, B. M. Martin, F. Lebreton, M. Delepiere, L. D. Possani and B. Becerril. Cloning and characterization of the cDNAs encoding Na⁺ channel-specific toxins 1 and 2 of the scorpion *Centruroides noxius* Hoffmann. *Toxicon* 33, 35-40, 1995.—Using a cDNA library prepared from venomous glands of the Mexican scorpion *Centruroides noxius* Hoffmann the genes that encode toxins 1 and 2 were identified, cloned and sequenced. In view of the proposed mechanism for processing the mature peptides coded by these two genes, the corresponding peptide-toxins were sequenced *de novo*. Mass spectrometric and ¹H-NMR analyses of the C-terminal peptide produced by enzymatic digestion of both toxins indicated that the last residue is asine-amide. Sequence comparison revealed that these two genes have a similarity of 56% and 80% at the amino acid and nucleotide levels, respectively. Small corrections to the published primary structures were introduced: Cn toxin 1 has an extra asine residue at position 63 and the residue in position 60 is a proline, while the amino acids at positions 34 and 35 of Cn 2 are, respectively, tyrosine and glycine. Sequence comparison of toxins from the genus *Centruroides* suggests the presence of at least three classes of distinct peptides in these venoms.

INTRODUCTION

Scorpion venoms contain an assortment of low mol. wt peptides toxic to a variety of organisms, including humans (Miranda *et al.*, 1970; Possani, 1984; Zlotkin *et al.*, 1978). Scorpion toxins have been considered valuable tools for the study of ion channels (Catterall, 1976). Two main families of distinct peptides have been described: (1)

*Author to whom correspondence should be addressed.

long-chain polypeptides containing 60–70 amino acid residues, which affect Na⁺ channels of excitable cells (Catterall, 1977; Couraud *et al.*, 1982), and (2) short-chain peptides of 31–39 amino acid residues, blockers of K⁺ channels of excitable cells (Carbone *et al.*, 1982; Gimenez-Gallego *et al.*, 1988; Miller *et al.*, 1985; Posaani *et al.*, 1982; Strong *et al.*, 1989). We have previously described the purification, characterization and complete primary structure of *C. noxius* toxin 1 (abbreviated Cn1), previously called component II-14 (Posaani *et al.*, 1985) and *Centruroides noxius* toxin 2 (abbreviated Cn2), previously called component II-9.2.2 (Zamudio *et al.*, 1992), from *C. noxius* venom. This communication reports the nucleotide sequence of the cDNAs encoding these toxins. Because the peptides encoded by these cDNAs are modified during the maturation process at the N-terminal site (cleavage of signal peptide) and at the C-terminal part (amidation of a serine residue for both Cn1 and Cn2), the amino acid sequence of the corresponding peptides was performed *de novo*. A couple of minor corrections to the published primary structure of Cn1 and Cn2 are now consequently included.

Comparative analysis conducted thereafter suggests the existence of three distinct families or groups of peptides in the venom of scorpions of the genus *Centruroides*.

MATERIALS AND METHODS

Cloning and nucleotide sequencing

Inserts of cDNA from clones Ceg1V and Ceg1Y (Bocerril *et al.*, 1993) were amplified by PCR using lambda gt11 forward (5'-GGTGCCGACGACTCTGGAGCCCG) and reverse (5'-TTGACACACACCACTGTGTAATG) primers (New England Bio-Labs, Beverly, MA, U.S.A.). These primers hybridize with the flanking regions of the lambda gt11 EcoRI cloning site. The reactions were performed using a Perkin Elmer 9600 using 30 rounds of temperature cycling (93°C for 1 min, 60°C for 1 min and 72°C for 2 min) followed by a final 15 min step at 72°C. These PCR products were purified from gel, labelled using [³²P]dCTP (New England Nuclear, Boston, MA, U.S.A.) by the random primer extension method (Dupont, kit NEP-103, Boston, MA, U.S.A.) following the manufacturer's instructions. The labelled PCR products were used to screen the cDNA library. The construction of the venom gland cDNA library, the screening conditions and the sequencing of the desired inserts were performed as described elsewhere (Bocerril *et al.*, 1993). The lambda gt11 forward and reverse and the M13-20 and M13 reverse oligo primers were used for sequencing.

Amino acid sequence determination

Toxins Cn1 and Cn2 were purified as described by Posaani *et al.* (1985) and Zamudio *et al.* (1992). For Cn1 (Posaani *et al.*, 1985), an additional chromatographic step on HPLC using a C4 reverse-phase column was introduced. Peptide eluting at 37.33 min corresponded to pure component Cn1 (data not shown). Both toxins were reduced and carboxymethylated (RC-Toxin) (Zamudio *et al.*, 1992). RC-toxin Cn1 was cleaved with protease Asp-N (Boehringer Mannheim, Germany), while RC-toxin Cn2 was digested with protease V8 from *Staphylococcus aureus* (Boehringer Mannheim, Germany). The resulting fragments were separated on HPLC by means of an analytical column (reverse-phase C18 from Vydac, Hysperia, CA, U.S.A.). The ProSequencer model 6400/6600 (MilliGen/BioSearch/Millipore Division, Burlington, MA, U.S.A.) was used to determine the corresponding sequences (Vazquez *et al.*, 1993).

Mass spectrometry and ¹H-NMR spectroscopy

Mass spectrometry of the C-terminal fragments of both RC-toxin Cn1 and Cn2 was obtained as earlier described (Vazquez *et al.*, 1993), using both a 3-sector mass spectrometer from Scia and a Hewlett Packard electrospray mass spectrometer.

For the proton nuclear magnetic resonance (¹H-NMR) experiments, freeze-dried native toxin Cn2 was dissolved in H₂O/D₂O, 9:1 (v/v). The peptide concentration was 1.5 mM, and the solution pH was adjusted at 3.5. The NMR spectra were run on a Varian Unity 500 MHz spectrometer at 303 K. Data were processed with the VNMR 4.3 (Varian Inc.) software. P. COSY (correlated spectroscopy; Marion and Baer, 1988), clean TOCSY (total correlated spectroscopy; Grzesinger *et al.*, 1988) and NOESY (nuclear Overhauser effect spectroscopy; Kumar *et al.*, 1980) experiments were used for the identification of spin systems, as described recently (Lebetron *et al.*, 1994).

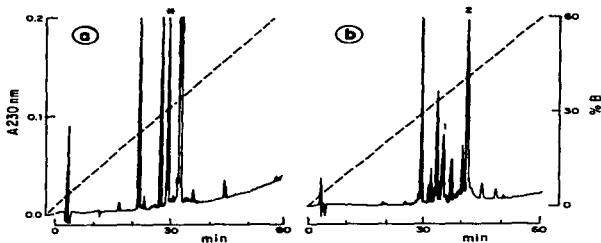


Fig. 3. HPLC separation of peptides generated by endopeptidase cleavage of Cn1 and Cn2. (a) RC-toxin Cn1 (approx. 100 μ g) was digested with Asp-N endopeptidase and separated in a C18 reverse-column, using a linear gradient of 0% solvent A (0.10% trifluoroacetic acid in water) to 60% solvent B (0.12% trifluoroacetic acid in acetonitrile), for 60 min. *Peptide corresponding to the C-terminal segment of Fig. 2. (b) RC-toxin Cn2 (approx. 100 μ g) was digested with *Staphylococcus aureus* protease V8 and separated in the same system as (a). Numbers 1 and 2 correspond, respectively, to the overlapping peptides V8₁ and V8₂ of Fig. 2.

amino acids at positions 33 and 34 (bold in Fig. 2), tyrosine being the one at position 33 and glycine the one at position 34; hence confirming the nucleotide sequence found in Fig. 1. Toxin Cn2 was also analysed by NMR, as discussed below, and unequivocally showed that the last residue in position 66 was amidated.

Mass spectrometry and ¹H-NMR spectroscopy

The pure peptide corresponding to the C-terminal region of Cn1, after cleavage with the protease Asp-N (labelled with an asterisk in Fig. 3a), was additionally submitted to mass spectroscopy analysis showing the presence of a mol. mass of 1602.7 indicating that the serine at the end, position 65, might be amidated. Similarly, the pure C-terminal peptide of RC-toxin Cn 2 (with sequence QAIWVPLPNKRCS) was analysed by mass spectrometry, giving a mol. mass of 1568.4, consistent with the presence of an amidated serine at the end of this peptide. Thus, these results confirmed both the amino acid sequence that was directly determined and the sequence inferred from the nucleotide sequence. Further confirmation of the results found for Cn2 toxin came from ¹H-NMR spectroscopy analysis. Using a combination of 2D (two-dimensional)-NMR experiments, most of the protons in the Cn2 toxin have been assigned, and the NMR three-dimensional structure of Cn2 is under refinement (Lebreton, F., Gurrola, G. B., Possani, L. D. and Delapierre, M. in preparation). Owing to the characteristic pattern of serine spin system (Wuthrich, 1986) and to the uniqueness of this residue in the Cn2 sequence, the protons of the serine at position 66 (Ser66) were assigned unambiguously (with δ HN: 8.79 ppm; δ Ha: 4.49 ppm; δ H β 1: 3.91 ppm; δ H β 2: 3.98 ppm). In the NOESY spectra, cross-peaks were observed between a proton at 7.47 ppm and Ser66: Ha, HN, and H β 1 protons. A smaller NOE

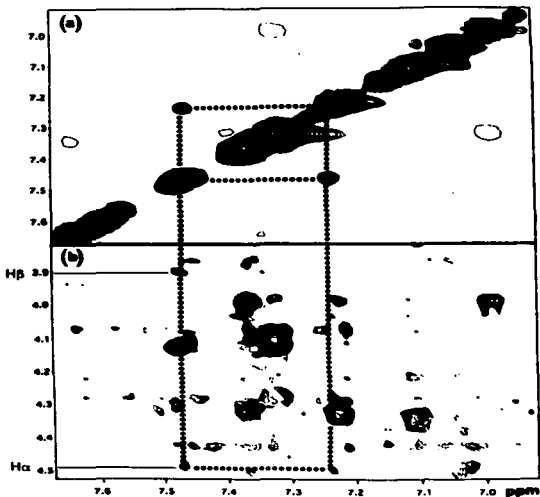


Fig. 4. Connectivities observed in the ROESY (a) and NOESY (b) spectra allowing the assignment of the two amide protons of the amidated-serine at the C-terminal position of Cn2. (a) ROESY experiment run with a mixing time of 200 msec at 303 K. The positive peaks due to chemical exchange are shown with numerous contour levels, and the negative peaks due to dipolar interaction are drawn with only one contour level. The two amide protons of the amidated-Ser66 resonating at 7.47 ppm and 7.24 ppm display a positive ROE cross-peak. (b) NOESY experiment run with a mixing time of 60 msec at 303 K shows the correlations observed between one of the amide proton at 7.47 ppm and the Ser66: H α and H β 1 protons. A smaller NOE cross-peak is also present between the second amide proton resonating at 7.24 ppm and the Ser66: H α proton.

cross-peak was also present between a proton resonating at 7.24 ppm and Ser66: H α proton. These two protons (at 7.47 and 7.24 ppm) are scalarly coupled as shown in the TOCSY and P. COSY experiments. The ROESY experiment (Kessler *et al.*, 1987), run with a mixing time of 200 msec at 303 K, indicates that these two protons are in chemical exchange, as expected for the protons of an amide function. Some of the connectivities

observed in the ROESY and NOESY spectra are shown in Fig. 4. Therefore, the two protons unambiguously assigned at 7.47 ppm and 7.24 ppm demonstrate the presence of the amidated serine at the C-terminal position.

DISCUSSION

The structure of the cDNAs reported (Fig. 1) shows the presence of an open reading frame encoding a signal peptide (19 aa for toxin Cn1) followed by a region encoding the mature peptide (65–66 aa) and two additional codons that code for a basic residue (Lys) preceded by a glycine residue, and finally a stop codon and 60–90 bp as a 3' non-coding region in which the polyadenylation site should be included (the latter is evident in the cDNA reported for toxin Cn2). A similar canonical structure has been reported for cDNAs encoding toxins of the North African scorpion *Androctonus australis Hector* (Bougis *et al.*, 1989) and the Mexican scorpion *Centruroides noxius* (Becerril *et al.*, 1993). The isolated and sequenced peptides from these scorpion venoms showed that they are processed in both ends: the signal peptides are eliminated and the C-terminal residues are processed given either a free carboxyl group or an amidated end-terminal amino acid residue. The basic residue situated in front of the codon of termination is not present in the mature peptides and when Gly precedes one or two basic residues, the residue preceding Gly becomes amidated (Bougis *et al.*, 1989; Vazquez *et al.*, 1993). According to these rules the carboxyl end of toxins Cn1 and Cn2 should be: ... cysteinyl-serinamide. This seems to be true for both toxins, as demonstrated by the NMR and mass spectroscopy data presented here. Unfortunately, the NMR spectroscopy data are not yet available for Cn1. However, the mass spectrometry and NMR data obtained for Cn2 leave no doubt that the rule discussed previously (Becerril *et al.*, 1993; Bougis *et al.*, 1989; Vazquez *et al.*, 1993) holds true, at least for the latter toxin, where the Ser66 is now definitively established to be amidated.

The similarity at the nt level between *CngtVI* and *CngtVII* suggests that these cDNAs originate from mRNAs transcribed from closely related genes. This last premise can be generalized for CnH toxin genes. We have proposed that toxins of scorpions of the *Centruroides* genus, specific for Na⁺ channels, can be divided into three groups (Fig. 5) according to their primary sequences (Becerril *et al.*, 1993). It has also been suggested that there is a relationship between the various groups of distinct amino acid sequences and the pharmacological properties (specific for mammals, crustaceans and insects) of the toxins included in each group or family of different peptides.

The similarity among the cDNAs we have reported thus far, at the aa level fluctuates between 56% and 91%, but at the nt level oscillates between 80% and 92%. Even for cDNAs with 56% similarity at the aa level, the similarity in the region encoding the signal peptide and the 3' non-coding region is remarkable. This analysis led us to conclude that the genes encoding toxins of the CnH venom have been generated from an ancestor gene which experienced several duplications, and each duplicated version evolved independently. This same phenomenon might have occurred in other species of the *Centruroides* genus. A similar event has been observed in the case of the North African scorpion *A. australis Hector* (Bougis *et al.*, 1989).

Returning to the discussion of the species specificity of the scorpion toxins, this subject is still under intense investigation. The groups in Fig. 5 were initially proposed based on sequence comparison (Becerril *et al.*, 1993), but they also correspond to results obtained with the purified peptides using mice, crickets and crustaceans as bioassays. Examples of

the first group, which include peptides from scorpions of the genus *Centruroides* toxic to mice, were reported by Rochat *et al.* (1979) for *C. suffusus suffusus* toxin II (Fig. 5); Possani *et al.* (1982), Zamudio *et al.* (1992), Vazquez *et al.* (1993) for *C. noxius* toxins 2 (same as *Cngt*VII), 3 and 4 (same as *Cngt*V in Fig. 5), respectively; and finally, for *C. l. tecomanus* toxin 1 (*Clt*I in Fig. 5) by Ramirez *et al.* (1988) and Martin *et al.* (1988). For purposes of comparison the sequences deduced from cDNAs (labelled *Cngt*X, where X is a Roman number in Fig. 5) include the unprocessed end terminal residues. The consensus residues (Cons. 1, bottom part of group 1 in Fig. 5), meaning identical amino acids at the same position in all sequences, are very high. Furthermore, toxins Cn2 and Cn3 were shown to be non-toxic to crustaceans, at doses lethal to mice (Serrano, S. and Possani, L. D., unpublished observations).

The second group contains examples of representative toxins specific to insects, such as the variants 1, 2 and 3 (*Csv*1-3 in Fig. 5) from the scorpion *C. sculpturatus*, as described by Babin *et al.* (1974), and minor corrections for variant 3 by Fontcolla-Camps *et al.* (1980). This group includes two amino acid sequences deduced from cDNAs (*Cngt*II and *Cngt*III, Fig. 5), from which only *Cngt*II has already been isolated as an expressed peptide. It turns out to be toxic to insects and crustaceans, but definitively non-toxic to mammals. Also included is *CIII*, a new crustacean toxin isolated from *C. limpidus limpidus* scorpion (Lebreton *et al.*, 1994). The consensus sequence (Cons. 2) for this group is striking.

GROUP 1		10	20	30	40	50	60
<i>Cngt</i> VII		KESYLVRKSTGCKYECLELGGDDSYCLAECHQYQEGAGDGYCYAFACNCHLVEGAVVWFLPSEKCSK					
<i>Csv</i> 1		KESYLVRKSTGCKYECLELGGDDSYCLAECHQYQEGAGDGYCYAFACNCHLVEGAVVWFLPSEKCSK					
<i>CJ</i> 1		KESYLVRKSTGCKYECLELGGDDSYCLAECHQYQEGAGDGYCYAFACNCHLVEGAVVWFLPSEKCSK					
<i>Cn</i> 3		KESYLVRKSTGCKYECLELGGDDSYCLAECHQYQEGAGDGYCYAFACNCHLVEGAVVWFLPSEKCSK					
<i>Cngt</i> V		KESYLVRKSTGCKYECLELGGDDSYCLAECHQYQEGAGDGYCYAFACNCHLVEGAVVWFLPSEKCSK					
Cons. 1		KESYLVRKSTGCKYECLELGGDDSYCLAECHQYQEGAGDGYCYAFACNCHLVEGAVVWFLPSEKCSK					
GROUP 2		10	20	30	40	50	60
<i>Csv</i> 2		KESYLVRKSTGCKYECLELGGDDSYCLAECHQYQEGAGDGYCYAFACNCHLVEGAVVWFLPSEKCSK					
<i>Cngt</i> III		KESYLVRKSTGCKYECLELGGDDSYCLAECHQYQEGAGDGYCYAFACNCHLVEGAVVWFLPSEKCSK					
<i>Csv</i> 3		KESYLVRKSTGCKYECLELGGDDSYCLAECHQYQEGAGDGYCYAFACNCHLVEGAVVWFLPSEKCSK					
<i>CIII</i>		KESYLVRKSTGCKYECLELGGDDSYCLAECHQYQEGAGDGYCYAFACNCHLVEGAVVWFLPSEKCSK					
<i>Csv</i> 3		KESYLVRKSTGCKYECLELGGDDSYCLAECHQYQEGAGDGYCYAFACNCHLVEGAVVWFLPSEKCSK					
Cons. 2		KESYLVRKSTGCKYECLELGGDDSYCLAECHQYQEGAGDGYCYAFACNCHLVEGAVVWFLPSEKCSK					
GROUP 3		10	20	30	40	50	60
<i>Csv</i> 1		KESYLVRKSTGCKYECLELGGDDSYCLAECHQYQEGAGDGYCYAFACNCHLVEGAVVWFLPSEKCSK					
<i>Cngt</i> IV		KESYLVRKSTGCKYECLELGGDDSYCLAECHQYQEGAGDGYCYAFACNCHLVEGAVVWFLPSEKCSK					
<i>Cngt</i> VI		KESYLVRKSTGCKYECLELGGDDSYCLAECHQYQEGAGDGYCYAFACNCHLVEGAVVWFLPSEKCSK					
<i>Csv</i> 2		KESYLVRKSTGCKYECLELGGDDSYCLAECHQYQEGAGDGYCYAFACNCHLVEGAVVWFLPSEKCSK					
Cons. 3		KESYLVRKSTGCKYECLELGGDDSYCLAECHQYQEGAGDGYCYAFACNCHLVEGAVVWFLPSEKCSK					

Fig. 5. Amino acid sequence comparison of main representatives of *Centruroides* scorpion toxins. These sequences were grouped according to their similarities. Gaps (-) were introduced where necessary to maximize similarities. Below each group a consensus sequence is shown in bold. For consensus sequences X means variable residue. Amino acid sequences deduced from cDNAs (denoted as *Cngt* and a Roman number) are included. The last two residues at the carboxyl end of the peptides encoded by *Cngt* II, V, VI and VII cDNAs, are not present in the mature toxin. The last two residues of the peptides encoded by cDNAs III and IV (peptides not yet isolated from the venom) are presumably not present either. The aa are aligned with the last digits of the numbers. The terminal ends of the mature peptides encoded by *Cngt* V (Vazquez *et al.*, 1993), *Cngt* VI and *Cngt* VII (this study), are asparagine-amide, serine-amide and serine-amide, respectively. Data are from this study, Boctrill *et al.* (1993), Maves *et al.* (1984), Zamudio *et al.* (1992) and Lebreton *et al.* (1994).

The third group contains toxin Cn1 (name as *Cngt*VI, Fig. 5) which was originally said to be toxic to mice (Possani *et al.*, 1985), but has turned out to be rather toxic to crustaceans (Serrano, S. and Possani, L. D., unpublished observations). The initial purified toxin Cn1 (Possani *et al.*, 1985) contained a small contaminant of toxin Cn7, which is one of the most toxic peptides to mammals, isolated from the venom of *C. noxius*, with $K_D = 40$ pM to rat brain synaptosomes (Valdivia *et al.*, 1994). Thus, toxin Cn1 was assayed in crustaceans, showing a considerable degree of toxicity, but was non-toxic to mice at doses up to 100 μ g/20 g mouse.

The peptide coded for *Cngt*IV (Fig. 5) is still unknown. Toxin CsEI, from *C. sculpturatus* (Babin *et al.*, 1974), is toxic to cockerels. Thus, the latter group, which also has an important consensus sequence (Cons. 3, Fig. 5) is less well studied and is not strictly specific to crustaceans.

In conclusion, the various groups of toxins compared in Fig. 5 seem to show, *grosso modo*, a certain degree of species specificity. Although more experiments are under way to refine the data, it is worth noting that similar primary structures are apparently conditioning similar pharmacological effects *in vivo*.

Note added in proof—While this manuscript was under revision a new set of experiments was conducted with toxin Cn1 by an independent operator, showing unequivocally that this peptide is amidated at the end terminal residue (serine-amide). RC-toxin I from *C. noxius* was cleaved *de novo* with endopeptidase Asp-N, separated by HPLC (similarly to Fig. 3a) and the peptide corresponding to the C-terminal segment (residues 52–63) was analysed by mass spectrometry. Eight spectra were obtained and using three ions per spectrum (the 12 C, the 13 C1 and 13 C2 doubly charged species) gave 24 data points which gave an average mol. wt. of 1602.73, consistent with an amidated C-terminal amino acid residue. The authors wish to thank Dr Carl James, from N.I.M.H., for running these spectra and calculating the mol. wt. of the peptide.

Acknowledgments—This work was partially supported by grants from Howard Hughes Medical Institute (7519-327104), DGAPA-UNAM IN205893 and CONACYT-Mexico (0018-N9105) to L.D.P. and by funds from the Pasteur Institute (Paris) and the French National Center of the Scientific Research to M.D.

REFERENCES

- Babin, D. R., Watt, D. D., Goos, S. M. and Mlejnek, R. V. (1974) Amino acid sequences of neurotoxic protein variants from the venom of *Centruroides sculpturatus* Ewing. *Arch. Biochem. Biophys.* **164**, 694-706.
- Baccerli, B., Vazquez, A., Garcia, C., Corona, M., Bolivar, F. and Possani, L. D. (1993) Cloning and characterization of cDNAs that code for Na⁺-channel-blocking toxins of the scorpion *Centruroides noxius* Hoffmann. *Gene* **120**, 165-171.
- Bougis, P. E., Rochat, H. and Smith, L. A. (1989) Precursors of *Androctonus australis* Hector. *J. Biol. Chem.* **264**, 19,259-19,265.
- Carbone, E., Wanke, E., Prestipino, G., Possani, L. D. and Maglicke, 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**, 5328-5336.
- Catterall, W. A. (1977) Activation of the action potential Na⁺ ionophore by neurotoxins. An allosteric model. *J. Biol. Chem.* **252**, 8669-8676.
- Couraud, F., Jover, E., Dubois, J. M. and Rochat, H. (1982) Two types of scorpion receptor sites, one related to activation, the other to inactivation of the action potential sodium channel. *Toxicon* **20**, 9-16.
- Fontecille-Camps, J. C., Almassy, R. J., Suddath, F. L., Watt, D. D. and Bugg, C. E. (1980) Three dimensional structure of a protein from scorpion venom: a new structural class of neurotoxins. *Proc. Natl. Acad. Sci. U.S.A.* **77**, 6496-6500.
- Gimenez-Galligo, G., Navia, M. A., Ruben, J. P., Katz, G. M., Kaczorowski, G. J. and Garcia, M. L. (1982) Purification sequence, and model structure of charybotoxin, a potent selective inhibitor of calcium-activated potassium channels. *Proc. Natl. Acad. Sci. U.S.A.* **80**, 3329-3333.
- Griesinger, C., Otting, G., Wuthrich, K. and Ernst, R. R. (1988) Clean TOCSY for ¹H spin system identification in macromolecules. *J. Am. Chem. Soc.* **110**, 7870-7872.

- Kestler, H., Grisinger, C., Kernsebaum, R., Wagner, K. and Ernst, R. R. (1987) Separation of cross-relaxation and J cross-peaks in 2D rotating-frame NMR spectroscopy. *J. Am. chem. Soc.* **109**, 607-609.
- Kumar, A., Ernst, R. R. and Wüthrich, K. (1980) A two-dimensional nuclear overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross relaxation networks in biological macromolecules. *Biochem. Biophys. Res. Commun.* **98**, 1-6.
- Lebrun, F., Delepierre, M., Ramirez, A. N., Balderas, C. and Possani, L. D. (1994) Primary and NMR three-dimensional structure determination of a novel crustacean toxin from the venom of the scorpion *Centruroides limpidus limpidus* Karsch. *Biochemistry* **33**, 11,135-11,149.
- Marion, D. and Baes, A. (1988) F. COSY: A sensitive alternative for double-quantum-filtered COSY. *J. Magn. Reson.* **88**, 528-533.
- Martin, B. M., Carbone, E., Yatani, A., Brown, A. M., Ramirez, A. N., Gurrola, G. B. and Possani, L. D. (1988) Amino acid sequence and physiological characterization of toxins from the venom of the scorpion *Centruroides limpidus tecomae* Hoffmann. *Toxicon* **26**, 785-794.
- Meyer, H., Simard, M. J. and Watt, D. D. (1984) Biochemical and electrophysiological characteristics of toxins isolated from the venom of the scorpion *Centruroides sculpturatus*. *J. Physiol. Paris* **79**, 183-191.
- Miller, C., Moczydlowski, E., Letorre, R. and Phillips, M. (1985) Charybdotoxin, a protein inhibitor of single Ca^{2+} -activated K^{+} channels from mammalian skeletal muscle. *Nature* **313**, 316-318.
- Miranda, F., Kopeyan, C., Rochat, H., Rochat, C. and Lissitzky, S. (1970) Purification of animal neurotoxins. Isolation and characterization of eleven neurotoxins from the venom of the scorpions *Androctonus australis* Hector, *Buthus occitanus tarbientis* and *Leiurus quinquestratus quinquestratus*. *Eur. J. Biochem.* **16**, 514-523.
- Possani, L. D. (1984) Structure of scorpion toxins. In: *Handbook of Natural Toxins*, Vol. 2, pp. 513-550 (Tu, A. T., Ed.). New York: Marcel Dekker.
- Possani, L. D., Martin, B. M. and Svendsen, I. (1982) The primary structure of noxiustoxin: a potassium channel blocking peptide, purified from the venom of the scorpion *Centruroides noxius* Hoffmann. *Carlsberg Res. Commun.* **41**, 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*. Primary structures and sequence comparison by metric analysis. *Biochem. J.* **228**, 739-750.
- Ramirez, A. N., Gurrola, G. B., Martin, B. M. and Possani, L. D. (1988) Isolation of several toxins from the venom of the scorpion *Centruroides limpidus tecomae* Hoffmann. *Toxicon* **26**, 773-783.
- Rochat, H., Bernard, P. and Coussaud, F. (1979) Scorpion toxins: chemistry and mode of action. In: *Advances in Cytothermology*, Vol. 3, pp. 325-334 (Coccarilli, B. and Clementi, F., Eds). New York: Raven Press.
- Strong, P. N., Weir, S. W., Bacch, D. J., Hiestand, P. and Kocher, H. P. (1989) Effects of potassium channel toxins from *Leiurus quinquestratus* *Arboreus*. *Br. J. Pharmacol.* **98**, 817-826.
- Vazquez, A., Boerri, 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.* **328**, 43-46.
- Wüthrich, K. (1985) *NMR of Proteins and Nucleic Acids*, pp. 1-292 (Wüthrich, K., Ed.). New York: Wiley.
- 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.* **206**, 281-292.
- Zlotkin, E., Miranda, F. and Rochat, C. (1978) Chemistry and pharmacology of Euthinas scorpion venoms. In: *Handbook of Experimental Pharmacology*, Vol. 48, pp. 317-369 (Battini, S., Ed.). Berlin: Springer.