



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS QUÍMICAS

SISTEMAS DE EXPRESIÓN PARA LA OBTENCIÓN DE TOXINAS RICAS EN
ENLACES DISULFURO Y SU CARACTERIZACIÓN ESTRUCTURAL POR
RESONANCIA MAGNÉTICA NUCLEAR.

TESIS

PARA OPTAR POR EL GRADO DE

DOCTOR EN CIENCIAS

PRESENTA

M. en C. GUSTAVO ALFREDO TITAU DELGADO

TUTOR: DR. JOSÉ FEDERICO DEL RÍO PORTILLA
INSTITUTO DE QUÍMICA, UNAM

CIUDAD DE MÉXICO, FEBRERO 2020



UNIVERSIDAD NACIONAL
AUTÓNOMA DE MÉXICO

**PROGRAMA DE MAESTRÍA Y DOCTORADO EN
CIENCIAS QUÍMICAS**

SISTEMAS DE EXPRESIÓN PARA LA OBTENCIÓN DE TOXINAS RICAS EN
ENLACES DISULFURO Y SU CARACTERIZACIÓN ESTRUCTURAL POR
RESONANCIA MAGNÉTICA NUCLEAR.

TESIS

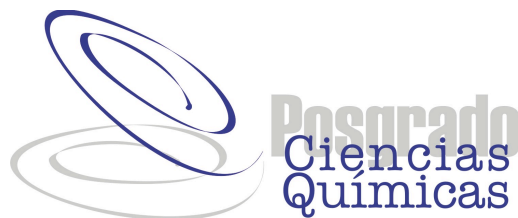
PARA OPTAR POR EL GRADO DE

DOCTOR EN CIENCIAS

PRESENTA

M. en C. GUSTAVO ALFREDO TITAUZ DELGADO

TUTOR: DR. JOSÉ FEDERICO DEL RÍO PORTILLA
INSTITUTO DE QUÍMICA, UNAM



CIUDAD DE MÉXICO, 2020

Jurado Asignado

Presidente: Dra. Adela Rodríguez Romero

Vocal: Dr. Armando Ariza Castolo

Vocal: Dr. Gerardo Corzo Burguete

Vocal: Dra. Nuria Esturau Escofet

Secretario: Dra. Karina Martínez Mayorga

Sitio donde se desarrolló el proyecto:

Instituto de Química, UNAM

Laboratorio 1 del departamento de Química de Biomacromoléculas y
Laboratorio de Biología Molecular.

Sustentante

M. en C. Gustavo Alfredo Titaux Delgado.

Asesor

Dr. José Federico del Río Portilla.

AGRADECIMIENTOS

A cada uno de los miembros del jurado por el tiempo y esfuerzo invertido en la revisión de esta tesis: Dra. Adela Rodríguez Romero, Dr. Armando Ariza Castolo, Dr. Gerardo Corzo Burguete, Dra. Nuria Esturau Escofet y a la Dra. Karina Martínez Mayorga

A mi comité tutor, que semestre a semestre me acompañaron en el desarrollo de esta investigación: Dra. Nuria Esturau Escofet y Dr. Luis Fernando Olguín Contreras

Quiero agradecer especialmente al Dr. José Federico del Río Portilla por aceptarme en su laboratorio, apoyarme siempre en todos los aspectos y mostrarme lo magnífica que es la Resonancia Magnética Nuclear.

Al departamento de supercómputo de la DGTIC por los recursos computacionales ocupados en este trabajo (Proyecto LANCAD-UNAM-DGTIC-145).

Este proyecto fue posible gracias al apoyo otorgado por el Consejo Nacional de Ciencia y Tecnología (CONACyT), número de becario 273490 y parcialmente financiado por la Dirección General de Asuntos del Personal Académico (DEGAPA) a través del proyecto IN210319.

I. RESUMEN.....	3
II. PUBLICACIONES Y PRESENTACIONES	5
III. LISTA DE FIGURAS.....	6
IV. LISTA DE TABLAS.....	7
V. SÍMBOLOS Y ABREVIATURAS.....	7
1. MARCO TEÓRICO.....	8
1.1. Neurotoxinas.....	9
1.1.1. ¿Qué son las neurotoxinas?.....	9
1.1.2. Canales iónicos.....	9
1.1.3. Neurotoxinas como herramientas farmacológicas.....	12
1.1.4. Obtención de neurotoxinas recombinantes.....	13
1.2. Determinación estructural de proteínas por resonancia magnética nuclear.....	18
1.2.1. Resonancia magnética nuclear como herramienta para la determinación estructural de proteínas.....	19
1.2.2. Intercambio químico.....	21
1.2.3. Desplazamiento químico (δ)	21
1.2.4. Acoplamiento (J)	23
1.2.5. Efecto nuclear Overhauser (NOE)	23
1.2.6. Experimentos de RMN	24
1.2.6.1. COSY.....	26
1.2.6.2. TOCSY.....	27
1.2.6.3. NOESY.....	28
1.2.7. Asignación de una proteína con experimentos bidimensionales.....	29
1.2.7.1. Asignación secuencial de aminoácidos.....	29
1.2.7.2. Asignación de NOE's a larga distancia.....	33
1.2.7.3. Cálculo de la estructura tridimensional.....	33
2. PROYECTO 1. TS-BKTX: UN PÉPTIDO DISFUNCIONAL.....	38
2.1. KTx- β y TsTx-K β	39
2.2. Hipotesis.	42
2.3. Objetivos: TsTx-K β	42
2.3.1. Objetivo general.....	42
2.3.2. Objetivos particulares.....	42
2.4. Diseño experimental: TsTx-K β	43
2.4.1. Diseño de los sistemas de expresión.....	43
2.4.2. Expresión y purificación de rTsTx-K β	45
2.4.3. Expresión, purificación y replegamiento de rHTsTx-K β y rTsHTx-K β	46
2.4.4. Expresión y purificación de rTs37Tx-K β	47



2.4.5. Reacción de bromuro de cianógeno.....	47
2.4.6. Caracterización de los péptidos tipo TsTx-K β	48
2.5. Resultado y discusión.....	49
2.5.1. Purificación de TsTx-K β	49
2.5.2. Purificación de rHTsTx-K β y rTsHTx-K β	51
2.5.3. Purificación de rTs37Tx-K β	53
2.5.4. Caracterización de las construcciones de TsTx-K β	54
2.6. Conclusiones: TsTx-K β	60
2.7. Perspectiva :TsTx-K β	61
3. PROYECTO 2. MAGI3: UNA PROTEÍNA RICA EN PUENTES DISULFURO.....	62
3.1. Macrothele Gigas y Magi3.....	63
3.2. Objetivos: Magi3.....	65
3.2.1. Objetivo general.....	65
3.2.2. Objetivos particulares.....	65
3.3. Diseño experimental.....	66
3.3.1. Expresión de la proteína de fusión.....	66
3.3.2. Purificación de la proteína de fisión.....	67
3.3.3. Replegamiento de la proteína de fusión.....	67
3.3.4. Corte enzimático con trombina.....	68
3.3.5. Obtención y caracterización de la Magi3.....	68
3.3.6. Proteólisis enzimática para la determinación de enlaces disulfuro.....	69
3.3.7. Obtención y análisis de los experimentos de RMN- ¹ H.....	69
3.3.8. Pruebas de actividad en grillos y ratones.....	69
3.4. Resultados y discusión.....	71
3.4.1. Expresión y purificación de la proteína de fusión de la Magi3.....	71
3.4.2. Replegamiento de la proteína de fusión y corte enzimático con trombina.....	71
3.4.3. Purificación y caracterización de Magi3.....	72
3.4.4. Caracterización estructural de Magi3 por RMN.....	74
3.4.5. Actividad biológica de Magi3.....	79
3.5. Conclusiones: Magi3.....	82
3.6. Perspectiva: Magi3.....	83
4. REFERENCIAS.....	84
5. ANEXOS.....	88
5.1. Artículo: “Successful refolding and NMR structure of rMagi3: A disulfide-rich insecticidal spider toxin.....	88



Las neurotoxinas son moléculas de gran relevancia para el estudio y entendimiento de los canales iónicos. Entre sus características principales destacan su estabilidad generada por la formación de enlaces disulfuro y una gran selectividad hacia sus blancos moleculares. Sin embargo, para poder estudiarlas se deben extraer y purificar de sus medios naturales como son los venenos de animales tales como alacranes, arañas y serpientes. Estos venenos son mezclas complejas de una gran diversidad de compuestos entre los que se encuentran las neurotoxinas, mismas que aparecen en concentraciones muy bajas dificultando así su purificación y estudio. Una alternativa es usar técnicas de biología molecular para obtenerlas de manera recombinante en sistemas heterólogos. Estas metodologías pueden producir una gran cantidad de muestra, lo que permite manipular los sistemas para facilitar su obtención y purificación. El principal reto es encontrar las condiciones óptimas para producir la proteína con el plegamiento nativo, enfocándose principalmente en la correcta conectividad y formación de los enlaces desulfuro.

En este trabajo, nos dedicamos al estudio de dos neurotoxinas con características particulares. TsTx-K β (60 a.a) es una toxina de alacrán que pertenece al grupo de las β -Ktx caracterizadas por ser toxinas de cadena larga (48 a 78 a.a.) y actividad frente a canales de potasio. Por su secuencia de aminoácidos se postula que este tipo de proteínas presentan dos dominios estructurales, un extremo amino-terminal asociado a una actividad citolítica y un carboxilo-terminal con capacidad de bloquear canales iónicos de potasio.

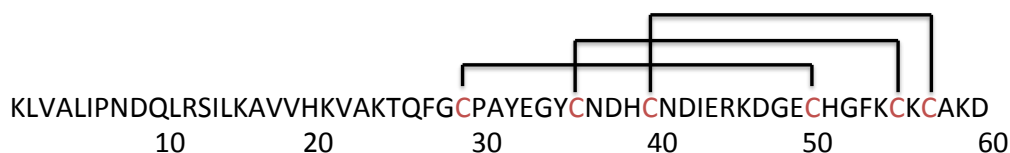


Figura I. Secuencia de aminoácidos de TsTx-K β y los puentes disulfuro predichos (Uniprot, P69940).



Por otro lado, Magi3 (46 a.a.) es una toxina extraída del veneno de araña la cual contiene en su secuencia de aminoácidos diez cisteínas, formando 5 enlaces disulfuro. La toxina presenta actividad sobre canales de sodio provenientes de insecto, haciéndola un péptido con potenciales aplicaciones biotecnológicas.

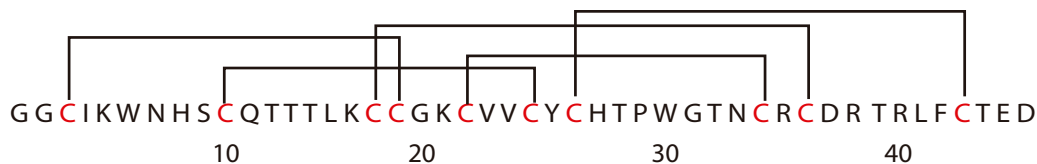


Figura II. Secuencia de aminoácidos de la Magi3 y los puentes disulfuro predichos (Uniprot, P83559)

Hasta la fecha no se cuenta con estudios estructurales de estas dos proteínas, lo cual brindaría información relevante para comprender su naturaleza y poder entender la forma en que estas interaccionan con sus blancos moleculares. Además Magi3 y TsTx-K β representan ejemplos de sistemas complejos con dificultades de expresión, resaltando el objetivo de diseñar sistema que nos permita obtenerlas de manera recombinante, generar suficiente muestra para poder realizar una caracterización estructural por RMN y evaluar su actividad biológica.

Este trabajo esta dividido en tres capítulos, el primero consta de un marco teórico en donde se habla de los dos temas principales en los que se basa esta tesis, las neurotóxicas y la resonancia magnética nuclear de proteínas. En los otros dos capítulos se desarrollan individualmente los proyectos mencionados anteriormente.



Los resultados de este trabajo fueron publicados en la revista "*PROTEIN SCIENCE*" con el título "Successful refolding and NMR structure of rMagi3: A disulfide-rich insecticidal spider toxin", DOI: 10.1002/pro.3363. Diciembre 2017.

También se presentaron en los siguientes eventos académicos:

- 59th Experimental Nuclear Magnetic Resonance Conference, Asilomar Conference Center, Pacific Grove, California, USA. "Biosynthesis and refolding of rMagi3, a toxin with five disulfide bonds and its structural determination by NMR" March 19-24, 2017.
- VI simposio de jóvenes investigadores de IQFR-CSIC celebrado el 10 de junio del 2019 en Madrid, España.



LISTA DE FIGURAS

- Figura 1.1. Estructura del canal de potasio KcsA (PDB 1J95).
- Figura 1.2. Estructura del canal de sodio NavAB (PDB 5VB8).
- Figura 1.3. Representación esquemática del fundamento de RMN.
- Figura 1.4. Espectro de protón donde se indican los desplazamientos químicos comunes de los diferentes protones en los aminoácidos.
- Figura 1.5. Representación esquemática de la obtención de un FID.
- Figura 1.6. Etapas de un experimento de dos dimensiones.
- Figura 1.7. Sistema COSY de la alanina y treonina.
- Figura 1.8. Sistema TOCSY de la alanina y treonina.
- Figura 1.9. Sistema de espín de leucina, valina, cisteína y ácido aspártico.
- Figura 1.10. Asignación de NOE's secuenciales y larga distancia.
- Figura 1.11 Red de NOE's
- Figura 2.1. Secuencia de aminoácidos de la proteína TsTx-K β .
- Figura 2.2. Alineamiento de las diferentes toxinas con similitudes a la TsTx-K β .
- Figura 2.3. Construcciones para la obtención TsTx-K β .
- Figura 2.4. Purificación de rGMTsTx-K β .
- Figura 2.5. Cromatograma de rGMTsTx-K β .
- Figura 2.6. Expresión y purificación de rHTsTx-K β
- Figura 2.7. Cromatograma de rHTsTx-K β y rTsHTx-K β .
- Figura 2.8. Producción y purificación de rTs37Tx-K β .
- Figura 2.9. Espectros de RMN de rTsTx-K β y EhMybS3v3.
- Figura 2.10. Comparación de los espectros monodimensionales de protón de rTsTx-K β , rhTsTx-K β , rTshTx-K β y rTs37Tx-K β .
- Figura 2.11. Dicroísmo circular de rTs37Tx-K β .
- Figura 3.1. Secuencias primarias de las Magi3 (1-6).
- Figura 3.2. Composición de la proteína de fusión que contiene a la Magi3.
- Figura 3.3. Purificación y corte enzimático de la proteína de fusión de Magi3.
- Figura 3.4. Purificación por cromatografía de fase reversa de GSAAAAM-Magi3.
- Figura 3.5. Espectro de masas MALDI-TOF de Magi3.
- Figura 3.6. Espectro monodimensional de hidrógeno de Magi3.
- Figura 3.7. Espectro bidimensional TOCSY de Magi3.
- Figura 3.8. Espectro bidimensional NOESY de Magi3.
- Figura 3.9. Estructura tridimensional de Magi3.
- Figura 3.10. Familias de toxinas de araña con actividad sobre canales de sodio.
- Figura 3.11. Comparación secuencial y estructural de Magi3.



LISTA DE TABLAS

Tabla 3.1. Fragmentos posibles de la proteólisis de Magi3 con Tripsina.

Tabla 3.2. Estadísticas de RMN de Magi3

Tabla 2.3. Pruebas de actividad de Magi3 en grillos y ratones.

SÍMBOLOS, ACRÓNIMOS Y ABREVIATURAS

a.a.	Aminoácidos
ACN	Acetonitrilo
CAMQ	Cromatografía de afinidad metal quelato
CYANA	Combined assignment and dynamics algorithm for NMR applications
COSY	Correlated Spectroscopy
DTT	Dithiothreitol
δ	Desplazamiento químico
DPFGE	Double Pulsed Field Gradient Spin Echo
FID	Free Induction Decay
GndCl	Cloruro de guanidinio
IPTG	Isopropil-tio-β-galcto-piranósido
LB	Lauria Bertani
MALDI-TOF	Matrix Assisted Laser Desorption-Ionization/time of light
MHz	Mega Hertz
NOE	Nuclear Overhauser Effect
NOESY	Nuclear Overhauser Effect Spectroscopy
PF	Proteína de fusión
ppm	Parte por millón
RMN	Resonancia magnética nuclear
RMSD	Root Mean Square Deviation
rpm	Revoluciones por minuto
SDS-PAGE	Sodium Dodecyl Sulfate-Polyacrylamide Gel
TFA	Trifluoroacetic acid
TFE	Trifluoroethanol
TOCSY	Total Correlation Spectroscopy



1.1 Neurotoxinas



1.1.1 ¿Qué son las neurotoxinas?

Las neurotoxinas son un grupo particular de moléculas que actúan como ligandos altamente específicos de proteínas presentes en la membrana celular, principalmente del sistema nervioso (Salceda et al. 2009). El origen biológico de estos compuestos es muy diverso y se pueden encontrar en múltiples organismos tanto unicelulares como pluricelulares. Estos organismos las utilizan para ahuyentar a los depredadores o capturar presas, algunos ejemplos de estos animales son las arañas, serpientes, escorpiones y conos marinos (Israel et al. 2018). Los venenos son mezclas complejas de múltiples componentes como sales, lípidos, compuestos orgánicos de bajo peso molecular, carbohidratos y una fracción proteica, en la que destacan los péptidos con actividad neurotóxica. El tamaño de estas neurotoxinas es menor a los 12 kDa y se caracterizan por su alto contenido en cisteínas formando de 2 a 8 enlaces disulfuro, lo que les confiere una gran estabilidad (King et al. 2013).

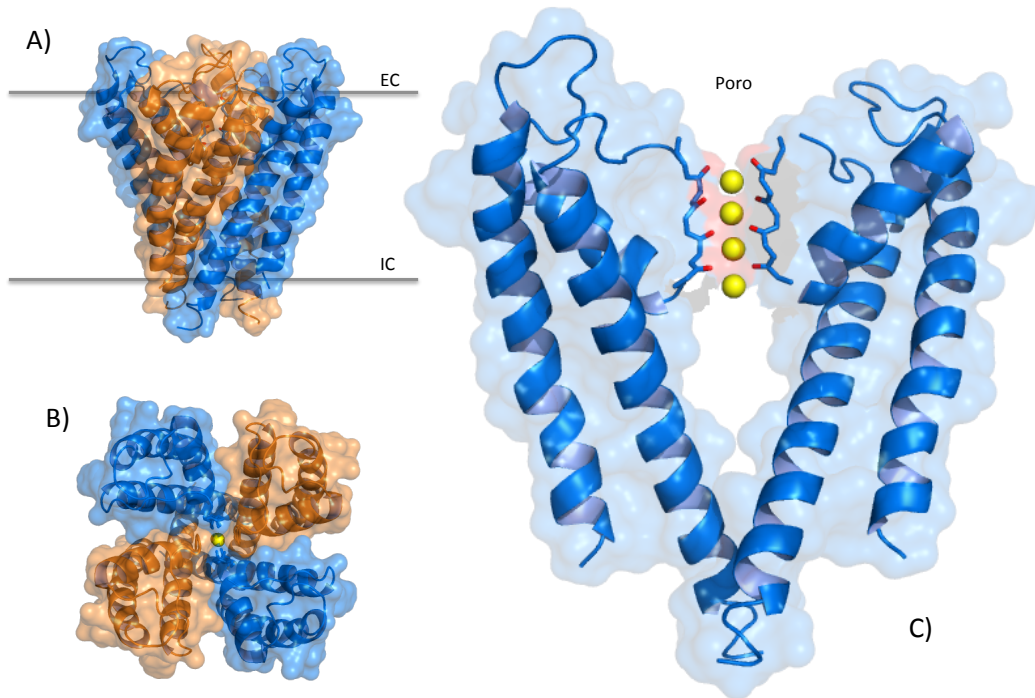
1.1.2 Canales iónicos

Las neurotoxinas por su tamaño, su toxicidad para los humanos o su especificidad biológica, siendo este último el más usado (Tytgat et al. 1999). Como se mencionó anteriormente las neurotoxinas bloquean o regulan proteínas específicas asociadas a las membranas celulares. Estas proteínas son conocidas como canales iónicos y permiten el paso de iones a través de la membrana plasmática de forma selectiva y a favor de gradiente electroquímico. Entre sus muchas funciones, destaca el mantenimiento del potencial de membrana, la contracción muscular, permiten la regulación del volumen celular, participan en diversas vías de señalización y son los encargados de la generación y propagación del impulso nervioso (Tamayo, 2007). Al igual que las enzimas, los canales iónicos tienen sustratos específicos: potasio, sodio, calcio y cloro. Durante los últimos años se han logrado impresionantes avances en la comprensión de su estructura y función. Utilizando canales iónicos purificados de bacterias, se ha encontrado que este tipo de proteínas forman un poro a través del cual los iones pueden pasar. Los canales de potasio por ejemplo son proteínas tetraméricas (Figura 1.1), cada subunidad contiene dos hélices transmembranales entre las cuales hay un bucle formado por una tercera hélice y una región conocida como filtro de selección que contiene la secuencia TVGYG. La región del filtro es estrecha y contiene sitios de unión a iones formados por átomos de oxígeno



orientados hacia el centro del poro (Sansom et al .2002).

Figura 1.1. Estructura del canal de potasio KcsA (PDB 1J95). A) y B) Representación en listones y



superficie de la naturaleza tetramérica del canal, en la figura se muestra el espacio extracelular (EC) e intracelular (IC). C) Muestra dos de las cuatro subunidades de KcsA, vistas de manera perpendicular al eje del poro. En rojo se muestran los oxígenos de los carboxilos en los residuos que forman el filtro de selección (TVGYG) y en amarillo los iones de potasio.

Algunos canales están regulados por ligandos, lo que significa que la apertura de los poros está acoplada enérgicamente a la unión de un ion, una pequeña molécula orgánica, o incluso en algunos casos una proteína (Wickman et al. 1995). Otros canales están regulados por voltaje, en cuyo caso la apertura del poro se acopla energéticamente al movimiento de un sensor de voltaje cargado dentro de la membrana. Por lo tanto, los distintos tipos de canales se abren en respuesta a estímulos como; un cambio en la concentración intracelular de calcio, al nivel de ciertas subunidades de proteína en la célula o al valor del voltaje de la membrana. Debido a esto, los canales tienen diversos dominios estructurales unidos a la unidad de poro. Los canales de potasio, por ejemplo, que son activados por ligando tienen típicamente dominios extracelulares o citoplásmicos para la unión del ligando; en el caso de los canales regulados por voltaje tienen dominios



de membrana integrales para detectar diferencias en el potencial de membrana (MacKinnon, 2003).

Por otro lado, los canales de sodio sensibles a voltaje de organismos procariontes están formados por cuatro subunidades idénticas (DI-DIV) y cada una esta compuesta por seis segmentos transmembranales (S1-S6). Los primeros cuatro segmentos (S1-S4) constituyen el módulo de detección de voltaje, que detecta cambios en el potencial de membrana y los transduce al poro. Los segmentos S5, S6 y un bucle entre ellos (asa P) forman el poro del canal como se observa en la Figura 1.2 (El-Din et al. 2018).

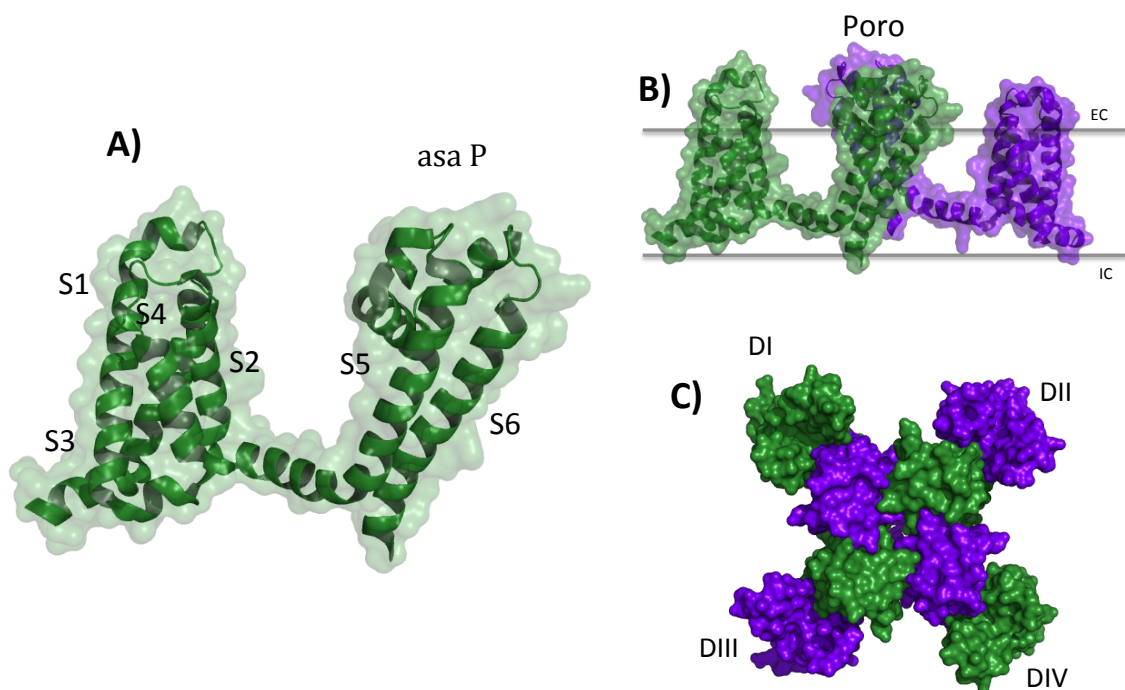


Figura 1.2. Estructura del canal de sodio NavAB (PDB 5VB8). A) Representación en listones y superficie de una subunidad del canal NavAB. Se observan los seis segmentos transmembranales (S1-S6). B) Dos de las cuatro subunidades de NavAB vistas de manera perpendicular al eje del poro, espacio extracelular (EC) e intracelular (IC). C) Conjunto de las cuatro subunidades del canal vista desde arriba del poro.

Existe una gran variedad de patologías asociadas al mal funcionamiento de los canales iónicos o las proteínas que los regulan, en la literatura se han nombrado aproximadamente 200 genes de canales iónicos como objetivos farmacológicos (Alexander et al., 2013). Aunque hay trabajos para la modificación del genoma en células individuales (Long et al, 2014; Maeder y Gersbach, 2016; Naldini, 2015) no se conoce el



efecto que puedan tener este tipo de terapias a largo plazo. Hoy en día es más seguro usar fármacos que se dirijan a las formas mutadas de las proteínas que constituyen al canal, desafortunadamente, en la práctica cuando se utilizan estos fármacos la especificidad al canal normal no desaparece.

En los últimos años, muchas investigaciones se han enfocado en la síntesis y modificación de compuestos orgánicos, con la esperanza de encontrar moléculas activas frente a las diversas patologías; sin embargo, las moléculas pequeñas mantienen un número restringido de puntos activos y presentan múltiples dianas debido a la baja especificidad. Las moléculas peptídicas en contra parte, pueden formar un plegamiento espacial único de interacción y sus metabolitos tienen una baja toxicidad (Kozlov, 2018).

1.1.3 Neurotoxinas como herramientas farmacológicas

Las neurotoxinas extraídas de venenos naturales, pueden ser la mejor alternativa para combatir las enfermedades relacionadas con el mal funcionamiento de los canales iónicos. Dichas toxinas están empaquetadas firmemente y algunas de ellas son extremadamente estables. En los venenos se encuentra una variedad inmensa de estas moléculas y cada una de ellas es específica para un grupo estrecho de receptores. Teniendo en cuenta, por ejemplo, el número de especies de arañas (~46,400) sus venenos representan la biblioteca más grande de moléculas biológicamente activas editadas de forma natural (Escoubas, 2006). Además, estas toxinas no solo pueden servir como moléculas bioactivas sino también, pueden ayudar a comprender mejor la farmacología de los canales iónicos. Actualmente existen varios ejemplos en la literatura de toxinas utilizadas para mejorar nuestra comprensión de la relación estructura-función de dichos canales. A pesar de la alta homología de secuencia (> 50%) entre los subtipos de canales de sodio, por ejemplo, se han aislado varias toxinas que son activas para uno o más subtipos de estos, lo que las convierte en herramientas útiles para diferenciar nuevos subtipos e identificar sitios alostéricos, lo cual, proporciona información esencial para el diseño racional de toxinas selectivas (Salceda, 2009; Israel, 2018; y Israel et al. 2017).



1.1.4. Obtención de neurotoxinas recombinantes

El uso de toxinas como herramientas de investigación puede estar limitado por la cantidad de muestra que se obtiene, actualmente existen tres formas principales de obtener neurotoxinas. La primera es extrayéndolas de su fuente natural, lo cual, involucra la recolección en el campo o la crianza en el laboratorio de especímenes para obtener sus venenos crudos; sin embargo, la falta de especímenes vivos para la colección de veneno, el alto precio, el pequeño catálogo de venenos liofilizados disponibles y los varios pasos de la purificación, representan cuellos de botella que limitan las cantidades de muestra con las que se disponen (Hernández et al. 2011). La segunda forma es la síntesis química de los péptidos, opción que evita la purificación de fuentes naturales. Sin embargo, esta técnica es muy costosa; la complejidad que incrementa conforme aumenta el tamaño del péptido y son necesarios pasos de plegado *in vitro* para que los péptidos sean activos, reduciendo considerablemente el rendimiento de los compuestos sintetizados químicamente (Lecomte et al., 1998). La tercera opción es la ingeniería genética, una alternativa que se ha convertido en una de las herramientas más importantes para producir grandes cantidades de péptidos y proteínas en general, con un costo relativamente bajo y la posibilidad de producir no solo el péptido nativo sino también variantes con cambios en la secuencia de aminoácidos para enriquecer su estudio.

Con estas metodologías se pueden llegar a tener mayores cantidades de muestra lo que hace posible realizar mas estudios, como investigar el mecanismo de acción de los péptidos, determinar sus dosis letales y obtener su estructura tridimensional (Hernández et al. 2011). Normalmente cuando se habla de proteínas recombinantes existen muchos puntos a considerar para su producción y purificación exitosa. A continuación se describirán algunos de los criterios más relevantes.

La primera decisión que se debe tomar al hacer una proteína recombinante es la elección del sistema en el que se va a expresar. Dependiendo si la proteína deseada, contiene o no modificaciones postraduccionales, se podría elegir un sistema eucarionte o procarionte, siendo este último el más usado para producir neurotoxinas, principalmente por su bajo costo. En el caso de las neurotoxinas, la característica más relevante es la correcta formación de enlaces disulfuro, para lo cual, ya se cuenta con varias estrategias y es posible llevarlo a cabo utilizando en algunos casos proteínas chaperonas que ayuden



a su formación o produciéndolas en bacterias con citoplasmas modificados para promover la oxidación (Collet y Bardwell, 2002).

Un problema recurrente con la producción de proteínas eucariotas en sistemas procariontes es el uso de codones diferentes. Sabemos que existe más de un codón para la inserción de un mismo aminoácido y cada organismo utiliza preferentemente ciertos codones, por lo tanto, un codón que es utilizado por humanos puede ser poco usado por bacterias. Afortunadamente, con la posibilidad de sintetizar químicamente el ADN que utilizaremos para la producción de la proteína, es posible optimizar el gen de tal forma que nos aseguremos que los codones que codifican para los aminoácidos de nuestra proteína sean los que la bacteria utiliza (Burgess-Brown et al. 2008).

La siguiente decisión importante es la elección del vector de expresión que se usará. Se deben tomar en cuenta aspectos como el promotor a partir del cual se expresará el gen de interés, el marcador de selección que permite el mantenimiento del plásmido en la célula huésped y el origen de replicación que regula el número de copias del plásmido en cada célula huésped.

Los promotores suelen estar regulados por un inductor químico que activa la transcripción cuando se agrega al cultivo, permitiendo el control temporal de la producción de la proteína. El ejemplo más conocido es el sistema pET; este sistema se basa en un diseño de *E. coli* que lleva una copia cromosómica del gen que codifica para la ARN polimerasa del bacteriófago T7. Este gen de la ARN polimerasa suele estar bajo el control de un promotor inducible por isopropil β -D-1-tiogalactopiranosido (IPTG); esta construcción se incorpora de forma estable en el cromosoma bacteriano. En ausencia de IPTG, el promotor lac está unido al represor LacI, que reprime la transcripción de la ARN polimerasa T7. Tras la adición de IPTG, LacI se libera de la región promotora lac y el gen de la ARN polimerasa T7 se transcribe y traduce. La ARN polimerasa de T7 ahora puede activar la transcripción del promotor T7 ubicado en el plásmido de expresión pET, en el cual se encuentra el gen recombinante de interés. De manera crucial, el promotor T7 en el plásmido de expresión pET no está fuertemente activado por la ARN polimerasa de *E. coli*, por lo que la expresión del gen de interés depende solamente del IPTG (Studier, 1991 y Reznikoff, 1992).



Otros aspectos importantes que involucran al vector de expresión son las modificaciones que puede llevar para facilitar la purificación o mejorar el plegamiento de las proteínas recombinantes. Las secuencias de nucleótidos que codifican estas etiquetas se fusionan directamente con el gen recombinante, puede ser en el extremo amino o carboxilo, por lo tanto, cuando se transcriben y traducen se generan como una sola cadena polipeptídica llamada normalmente proteína de fusión. Este tipo de construcciones, además de permitir que las proteínas se purifiquen de manera más simple, ayudan a fusionarlas con proteínas que pueden aumentar las solubilidad o hacer que se dirijan a compartimentos celulares particulares (Overton, 2014).

Una vez producida la proteína en la bacteria, se debe planear la forma de purificación. Una manera fácil es añadiendo una secuencia de histidinas a la construcción de la proteína. Estos residuos tienen la capacidad de acomplejarse con iones metálicos, como el níquel (II) o el cobre (II). Mediante el uso de una resina de cromatografía que contiene estos iones metálicos, la proteína recombinante se puede separar específicamente de otros componentes celulares en un proceso de cromatografía de afinidad metal quelato (Cheung et al. 2012).

La producción exitosa de una proteína recombinante no es garantía de que ésta tenga el plegamiento correcto para llevar a cabo su función biológica. Recordemos que estamos produciendo una proteína en un ambiente diferente al nativo y este puede ser crucial para su plegamiento. Normalmente, cuando una proteína no tiene su conformación nativa, tiende a formar agregados insolubles, partículas densas compuestas de proteína desplegada y proteína plegada en proporciones variables (Singh y Panda, 2005).

Los agregados insolubles a menudo se consideran indeseables; sin embargo, su formación puede ser ventajosa, ya que su aislamiento del homogeneizado celular es una forma conveniente y eficaz de purificar la proteína de interés. Las principales ventajas asociadas son las siguientes:

- Es relativamente más fácil su aislamiento, debido a las diferencias en su tamaño y densidad en comparación con los contaminantes celulares.



- Existe una menor degradación de la proteína expresada ya que al no ser soluble se reduce el contacto con las proteasas celulares.
- Contiene menos contaminantes lo que nos ayuda a reducir el número de pasos de purificación.

Por lo tanto, a pesar de que la proteína insoluble normalmente no tiene actividad deseada, su purificación directa puede ser una opción. Hay una gran cantidad de ejemplos en la literatura de proteínas recombinantes expresadas como agregados insolubles en *E. coli* (Singh y Panda, 2005).

Una gran parte de las neurotoxinas producidas de manera recombinante han sido purificadas a partir de agregados insolubles. En general, esta metodología consta de tres pasos: 1. El aislamiento de los agregados insolubles; 2. La disolución de estos y la parte más importante y delicada; 3. El replegamiento de la proteína, es decir, pasar de su forma desplegada a la forma plegada con la actividad deseada.

Normalmente estos agregados semipuros se disuelven utilizando altas concentraciones de reactivos desnaturizantes (6-8 M) como la urea o el cloruro de guanidinio; en otros casos se pueden utilizar también detergentes como el dodecilsulfato de sodio (SDS). Cuando la proteína tiene enlaces disulfuro es conveniente agregar agentes reductores adicionales (β -mercaptoetanol, ditioneitol o cisteína) rompiendo así los enlaces incorrectos inter o intramoleculares (Fischer et al. 1993).

Ya obtenida la proteína soluble y desplegada comienza el proceso de replegamiento para alcanzar su estado nativo; esto se lleva a cabo al eliminar el agente desnaturizante de manera controlada. Los procesos como diálisis o dilución son utilizados en esta etapa. Una de las razones de la mala recuperación de la proteína replegada es su agregación durante el proceso; para evitarlo se utilizan concentraciones de proteína en el intervalo de 10 a 50 $\mu\text{g/mL}$ y en ocasiones también se usan agentes estabilizadores como la arginina, la cual se ha demostrado mejora la solubilidad e inhibe la agregación (Reddy et al. 2005).

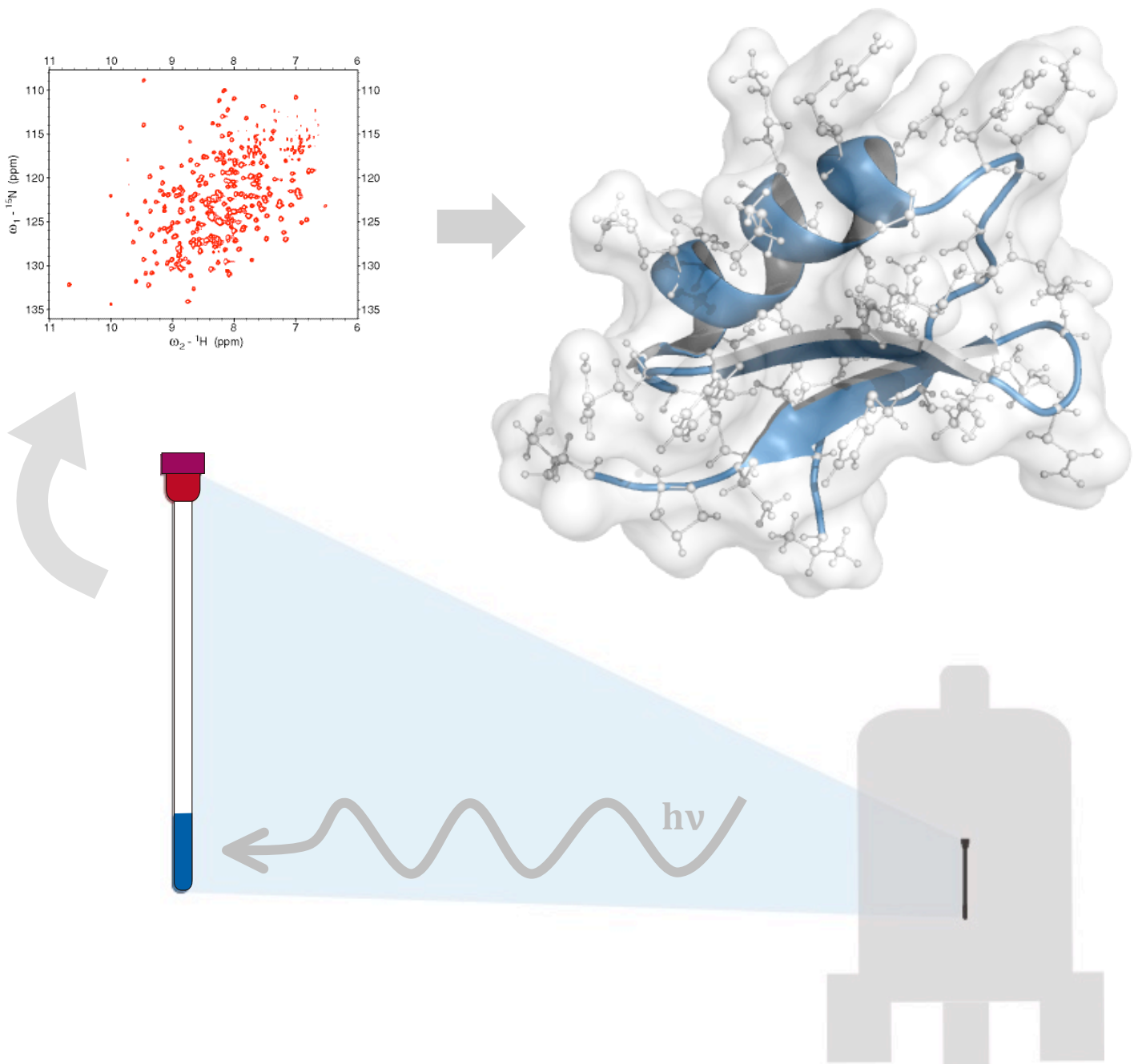


En el caso de las proteínas que tienen múltiples enlaces disulfuro como es el caso de la neurotoxinas, se agrega una complicación al sistema debido a la probabilidad de las diferentes combinaciones para formar estos enlaces disulfuro en la toxina. Por ejemplo, una neurotoxina con 6 cisteínas forma 3 enlaces disulfuro, pero existen 15 combinaciones posibles. En el caso de una secuencia con 10 cisteínas el número de posibles isoformas es de 945. Este tipo de proteínas requiere un repliegamiento más elaborado que se debe realizar en presencia de agentes oxidantes y reductores que ayuden a la formación de los enlaces. Los reactivos más utilizados son el glutatión oxidado/reducido (GSH/GSSH), DTT/GSSH, cisteína/cistina y cisteamina/cistamina a una concentración total de 1–15 mM con una relación molar de compuestos reducidos a oxidados de 1:1 a 5:1, respectivamente (Vallejo y Rinas, 2004).

Una vez obtenida la proteína, ya sea de forma soluble o extraída de los agregados insolubles, se caracteriza y evalúa para corroborar que sea la proteína deseada y que cuente con el plegamiento adecuado.



1.2. DETERMINACIÓN ESTRUCTURAL DE PROTEÍNAS POR RESONANCIA MAGNÉTICA NUCLEAR.



1.2.1. Resonancia magnética nuclear como herramienta para la determinación estructural de proteínas

La resonancia magnética nuclear es una técnica que permite de determinar la estructura tridimensional de péptidos y proteínas con una resolución a nivel atómico. Entre sus ventajas se encuentran que las proteínas se pueden estudiar en disolución permitiendo que estén en condiciones fisiológicas o en estado sólido si el objeto de estudio es fibroso (Mompeán et al., 2018). Esta herramienta aprovecha las propiedades magnéticas de los núcleos (espín nuclear) y los manipula con radiofrecuencias para obtener información acerca de sus ambientes químicos, relacionados directamente con su estructura molecular.

En condiciones terrestres, las orientaciones posibles asociadas al momento magnético de los núcleos son energéticamente equivalentes, generando un conjunto de estados degenerados. Si aplicamos un campo magnético externo, los momentos magnéticos se alinean y adquieren un valor energético particular generando diferentes estados poblaciones (Figura 1.3). La diferencia energética entre estos estados es determinada por $\Delta E = \hbar \gamma B_0$, donde \hbar es la constante de Planck dividida entre 2π , γ es la constante giromagnética de núcleo en cuestión y B_0 el campo magnético. Esta diferencia entre los estados energéticos es directamente proporcional al campo magnético externo y su magnitud se encuentra en el rango de las radiofrecuencias (Figura 1.3).

Con lo dicho anteriormente, la energía necesaria para llevar a cabo estas transiciones energéticas o poblacionales debería ser igual, si estamos hablando de núcleos del mismo isotopo sometidos al mismo campo; sin embargo, si estos núcleos tienen diferencias en sus densidades electrónicas provocados por átomos vecinos o por su posición en la molécula, éstas actuarán como pequeños campos magnéticos locales generando diferencias magnéticas y se podrán diferenciar entre ellos. Cada núcleo entonces, se observaran a una frecuencia de resonancia particular (desplazamiento químico, δ) que podremos interpretar para determinar su estructura molecular.



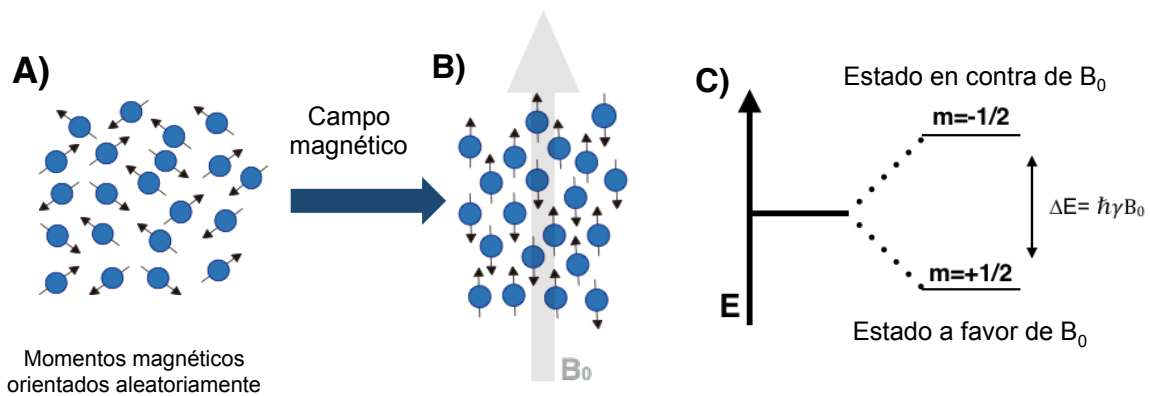


Figura 1.3. Representación esquemática del fundamento de RMN. Momentos magnéticos de núcleos con espín $\frac{1}{2}$ y su orientación en ausencia y presencia de un campo magnético, A) y B) respectivamente. C) Transición energética de los estados de un núcleo con espín $\frac{1}{2}$ sometido a un campo magnético.

Entre los átomos más utilizados para el estudio de proteínas se encuentran el hidrógeno, el carbono y el nitrógeno, de estos elementos, los isótopos ^1H , ^{13}C y ^{15}N con espín nuclear $\frac{1}{2}$ dan lugar a señales observables. El isótopo ^{12}C tiene espín 0 por lo que no es observable en resonancia, mientras que los isótopos ^2H y ^{14}N con espín 1 son núcleos con momento eléctrico cuadrupolares, los cuales tienen señales que relajan muy rápido y no son observables en disolución en condiciones normales.

Históricamente la RMN de macromoléculas estaba limitada debido a la gran complejidad de los espectros y la baja sensibilidad de la técnica, lo cual, en los últimos años se superó con la generación de equipos con campos más intensos, desarrollo de criosondas y mejores formas de preparación de muestra. Cada día es más frecuente encontrar estructuras resueltas de mayor tamaño. Las estructuras de moléculas menores a los 10 kDa se pueden determinar con experimentos ^1H homonucleares, menores a 30 kDa con experimentos de ^{13}C y ^{15}N heteronucleares y moléculas alrededor de 40 a 50 kDa combinando experimentos de ^{13}C , ^{15}N y ^2H heteronucleares (Cavanagh, 1996).

Existen varios parámetros espectrales que reflejan tanto la estructura como la conformación de la molécula, a continuación describiré los más importantes para el estudio de proteínas en disolución.



1.2.2. Intercambio químico

El intercambio químico involucra todos los procesos en los cuales un átomo intercambia entre dos o más entornos en los que sus parámetros de RMN (desplazamiento químico, constante de acoplamiento, tiempo de relajación) son diferentes. Este tipo de procesos pueden ser tanto intra como intermoleculares. Entre los intermoleculares destacan los movimientos de cadena lateral y los equilibrios conformacionales que una proteína puede adoptar en disolución. Entre los intramoleculares están la unión de moléculas pequeñas a la proteína, interacción proteína-proteína y el equilibrio de protonación-desprotonación de los grupos ionizables que existan, así como los átomos lábiles, como es el caso de los protones de amida.

Estos efectos de intercambio afectarán a los parámetros de RMN dependiendo de su velocidad, en el caso del desplazamiento químico, por ejemplo, en un intercambio lento se observan dos señales separadas para un núcleo en dos estados conformacionales. Por lo contrario en el intercambio rápido, se observará una única señal promediada. En el caso de intercambios intermedios existen varios cambios en las señales que van aproximándose y ensanchándose hasta parecer una señal única (Jiménez, 1997).

Este tipo de fenómeno se debe tomar en cuenta a la hora de la adquisición y análisis de los espectros de RMN.

1.2.3. Desplazamiento químico (δ)

El desplazamiento químico (δ) es la frecuencia de resonancia donde se observa la transición entre los estados energéticos de cada núcleo. Esta diferencia energética es directamente proporcional al campo magnético; por lo tanto, una manera de simplificar la interpretación y poder comparar señales entre espectros obtenidos en equipos de diferentes campos es convirtiendo el valor de la frecuencia en partes por millón con la siguiente ecuación: $(\delta) = 10^6 [(v_i - v_{ref}) / v_{ref}]$ donde v_{ref} es la frecuencia de resonancia de una referencia que se toma como cero en la escala.

Al efecto local que generan los electrones cercanos al núcleo frente al campo magnético externo se conoce como apantallamiento (es la razón principal por la cual podemos



observar diferencias en desplazamiento químico entre núcleos de la misma especie). Por lo tanto, este parámetro depende del entorno químico de cada núcleo (densidad electrónica). Cambios pequeños en el entorno electrónico de un núcleo o en la conformación de la molécula podría afectar de forma apreciable el valor del desplazamiento químico del núcleo en cuestión (Keeler, 2005).

Aunque los desplazamientos químicos pueden variar dependiendo del ambiente de cada núcleo, también es cierto que cada tipo de núcleo saldrá en zonas específicas del espectro. Por ejemplo, en la figura 1.4 se ejemplifican los desplazamientos típicos de los diferentes tipos de protones que conforman a los aminoácidos.

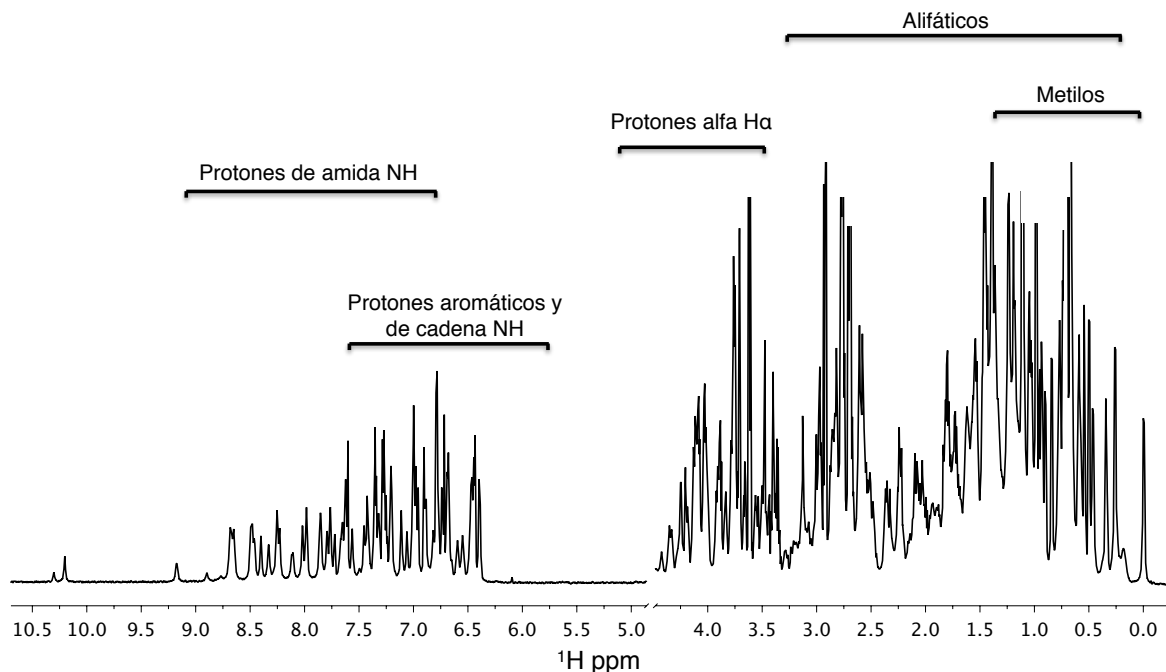


Figura 1.4. Espectro de protón donde se indican los desplazamientos químicos comunes de los diferentes protones en los aminoácidos.

No debe sorprendernos que en moléculas complejas puedan existir núcleos que tengan desplazamientos químicos iguales, sobre todo si hablamos de núcleos que observamos en ventanas espectrales pequeñas como es el caso de ^1H .

Análisis estadísticos de los δ de ^1H , ^{13}C y ^{15}N del gran número de estructuras obtenidas en los últimos años han puesto de manifiesto la existencia de una correlación entre el tipo de estructura secundaria (hélice α o lamina β) y los δ observados con respecto a los



valores medidos de péptidos modelos sin estructura, por lo tanto, es posible la predicción de estructura secundaria de una proteína con solo los datos del δ de núcleos como carbono y protón en la posición α principalmente (Laws et al., 1993).

1.2.4. Acoplamiento

El acoplamiento se define como la influencia que ejerce un núcleo sobre otro; este tipo de interacciones puede ser provocada a través de los enlaces (acoplamiento escalar) o simplemente a través del espacio (acoplamiento dipolar). Para que dos núcleos se acoplen deben tener espín nuclear. El acoplamiento es un parámetro muy valioso con el cual podemos obtener información estructural; por ejemplo, en el caso del acoplamiento escalar, su magnitud dependerá la densidad electrónica y de la geometría de los núcleos, obteniendo información acerca de los ángulos diedros de la molécula. En cambio, los efectos observados por el acoplamiento dipolar nos brindarán la distancian entre lo núcleos que están interaccionando.

El acoplamiento escalar será fundamental para relacionar los átomos que constituyen a cada aminoácido, además, si contamos con los valores de las constantes de acoplamiento, seremos capaces de conocer los distintos ángulos diedros que adopta una cadena polipeptídica, permitiéndonos delinear los diferentes elementos de estructura secundaria, diferenciando una hélice α de una hélice 3_{10} o reconociendo una laminas β (Cavanagh, 1996). Por su parte, con la manifestación del acoplamiento dipolar podremos conocer la distancia entre los núcleos, determinar la secuencia de los aminoácidos y finalmente el plegamiento tridimensional que la proteína adopta.

1.2.5. Efecto nuclear Overhauser (Nuclear Overhauser Effect, NOE)

El efecto nuclear Overhauser es la transferencia de magnetización entre núcleos próximos en el espacio. Esta transferencia tiene lugar por un acoplamiento dipolar generado por efectos de relajación, su eficacia depende de un factor de movilidad (tiempo de correlación) y de la inversa de la distancia a la sexta potencia entre los núcleos que interaccionan (Keeler, 2005).



La observación de un NOE entre dos protones implica que estos se encuentran próximos, a una distancia inferior o igual a 5 Å. En una proteína, los protones que se encuentran muy alejados en su secuencia de aminoácidos pueden estar cerca en el espacio y, por lo tanto, tener un NOE. Una característica del NOE, es que mientras los demás parámetros espectrales se promedian linealmente sobre las distintas conformaciones en el equilibrio, la magnitud del efecto NOE tiene una dependencia no lineal con las distancias interprotónicas, con lo que tienen más peso las conformaciones con distancias cortas haciendo posible la detección e identificación de dichas conformaciones aunque estas sean poco pobladas (Cavanagh, 1996).

1.2.6. Experimentos de RMN

En la figura 1.5 se ilustra las diferentes etapas en la que se adquiere un experimento de RMN de protón. Una vez que los momentos magnéticos de los núcleos se alinean al campo, éstos generan un vector de magnetización, el cual es paralelo al campo magnético (eje Z). Con un pulso de radiofrecuencias este vector es forzado a salir de su estado de equilibrio y es llevado al plano XY. En este punto comienza la detección y se observa como la magnetización regresa a su estado de equilibrio.

Como se observa en la figura 1.5 B), toda la magnetización se encuentra en el eje Y cuando inicia la detección; posteriormente ésta precesa y regresa al eje Z disminuyendo su magnitud en la componente Y hasta que se hace cero. La función de decaimiento que describe este proceso (free induction decay, FID) conlleva la información de todos los núcleos que componen la magnetización; sin embargo, la información adquirida esta en función del tiempo, por lo cual, es necesaria una transformación matemática (transformada de Fourier) que transfiera la información al dominio de las frecuencias. En este dominio, los espectros son interpretables. Al fenómeno en el que la magnetización desaparece del plano XY se le conoce como relajación transversal y cuando la magnetización reaparece en el eje z se le conoce como relajación longitudinal.



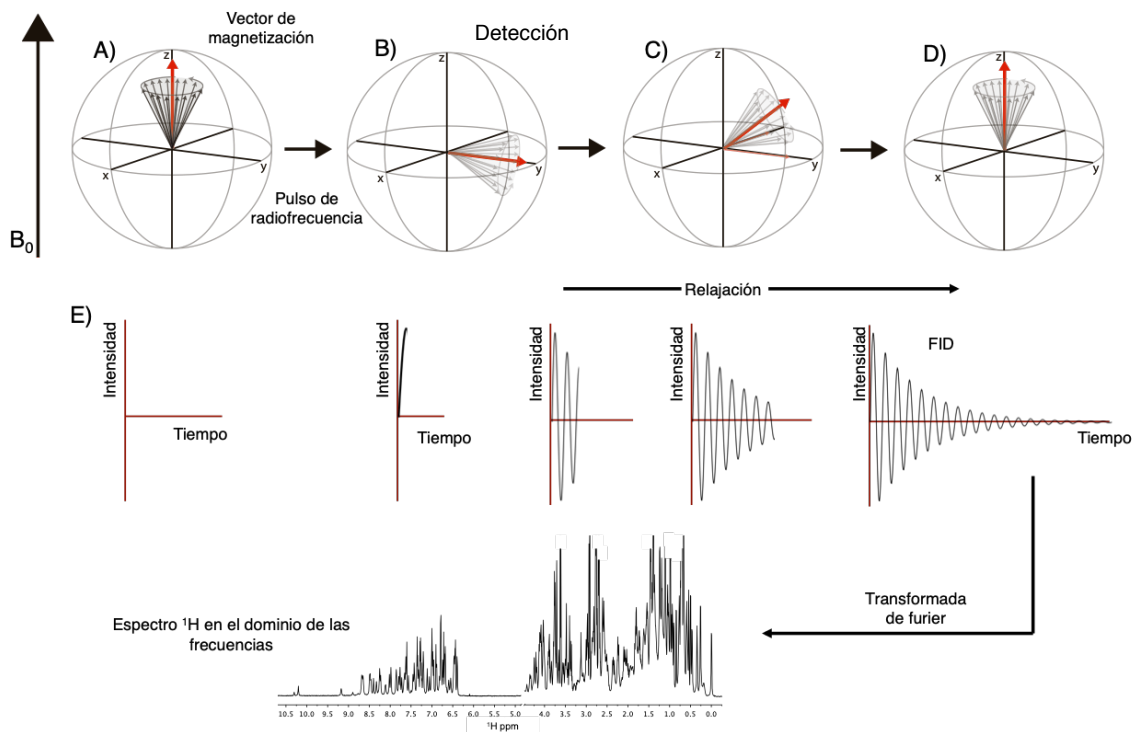


Figura 1.5. Representación esquemática de la obtención de un FID. A) Magnetización orientada al campo magnético, eje Z. B) Magnetización en el plano XY después de un pulso de 90°. C) Magnetización regresando a su estado basal de energía. D) Magnetización en equilibrio. E) Generación del FID, representación grafica de la intensidad de la magnetización en el eje Y con respecto a intervalos de tiempo observados en los incisos A,B ,C y D.

Las señales observadas en los espectros de RMN son el conjunto de todos los átomos que constituyen a la molécula. No es de extrañarse que entre más grande sea el sistema (como en el caso de las proteínas), los espectros sean más complejos y difíciles de interpretar. En este aspecto, la RMN monodimensional nos brinda muy poca información porque resulta imposible poder asignar las señales observadas. Hoy en día es posible dispersar los experimentos en una segunda o tercera dimensión, lo que abre una puerta a oportunidades para reducir la complejidad de los espectros y poder asignar las señales.

La obtención de un espectro de dos dimensiones es básicamente la adquisición de varios espectros en una dimensión en donde se introduce un segundo periodo de tiempo variable (tiempo de evolución t_1) mediante incrementos constantes. Para cada incremento t_1 se detecta un FID (t_2). La señal que se obtiene está ahora en función de dos variables de tiempo (t_1 y t_2). Al igual que en el caso de los experimentos en una dimensión, se debe



pasar la información al dominio de frecuencias aplicando una transformada de Fourier sobre t_1 y una segunda en t_2 . En la figura 1.6 se observa las etapas básicas de un experimento en dos dimensiones

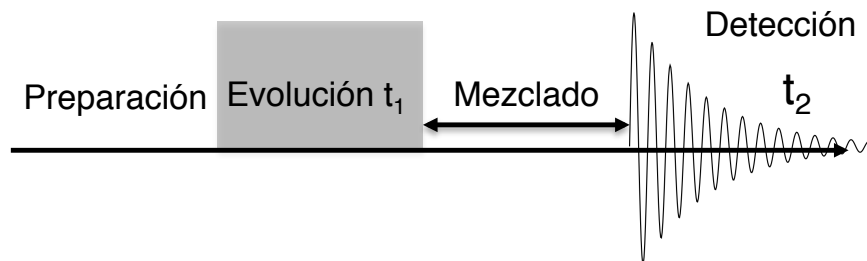


Figura 1.6. Etapas de un experimento de dos dimensiones. El tiempo de preparación permite que los núcleos excitados regresen a su estado de equilibrio y se puede introducir uno o varios pulsos que permiten que la magnetización, total o selectiva, se encuentre en el plano transversal. Durante el tiempo de evolución el sistema de espín evoluciona bajo el efecto de diferentes factores como el desplazamiento químico o el acoplamiento. En el tiempo de mezclado se lleva a cabo la transferencia de magnetización que permitirá obtener la información deseada dependiendo el experimento. Finalmente en el periodo de detección se adquiere el FID que contiene la información del sistema.

A continuación se describirán los tres experimentos bidimensionales básicos para asignación y determinación estructural de una proteína.

1.2.6.1. COSY

El experimento COSY, es un experimento ^1H bidimensional en donde podemos observar correlaciones entre núcleos a dos o tres enlaces de distancia (Keeler, 2005). En un polipéptido deberíamos observar señales provocadas por el acoplamiento entre los protones de amida (NH) con los protones alfa ($\text{H}\alpha$), los alfa ($\text{H}\alpha$) con los beta ($\text{H}\beta$), correlaciones entre cadenas alifáticas y correlaciones de sistemas aromáticos de cada aminoácido (Cavanagh, 1996). Por ejemplo, el aminoácido alanina (Figura 1.7) tiene tres tipos de protones no equivalentes, la señal del NH, del $\text{H}\alpha$ y las del CH_3 . En un experimento COSY veríamos señales que relacionan el NH con el $\text{H}\alpha$ por su acoplamiento ya que están a tres enlaces de distancia; también, veremos señales que relacionen al $\text{H}\alpha$ con los protones del CH_3 . En el caso de la treonina veríamos la relación entre el protón NH con el $\text{H}\alpha$, el $\text{H}\alpha$ con el $\text{H}\beta$ y el $\text{H}\beta$ con los CH_3 (Figura 1.3).



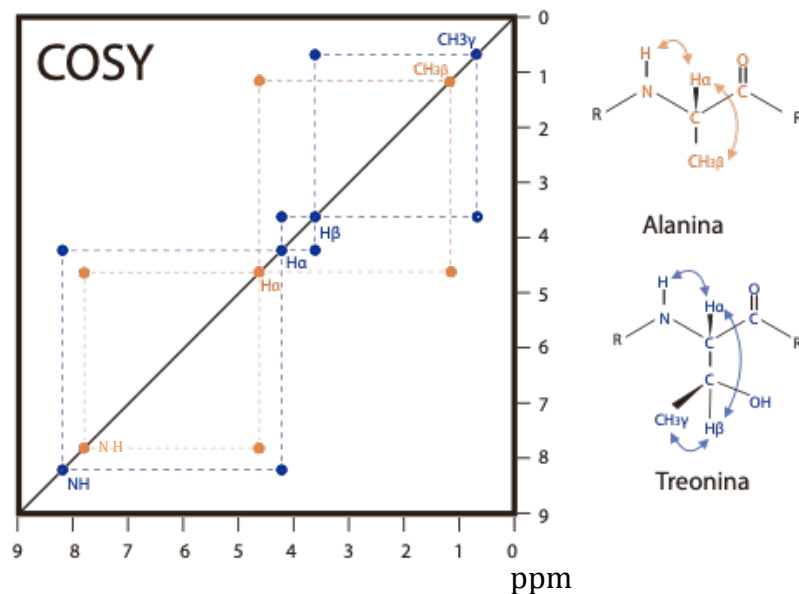


Figura 1.7. Sistema COSY de la alanina y treonina. En el experimento COSY se observan los acoplamientos a dos y tres enlaces de distancia. En la estructura de los aminoácidos (alanina y treonina) se resaltan con flechas los protones que interactúan.

1.2.6.2. TOCSY

El experimento TOCSY es otro experimento bidimensional (^1H) donde se observa la correlación en un sistema.

Imaginemos un núcleo A el cual está acoplado a uno B, el B a un núcleo C y el C a uno D todos ellos a tres enlaces de distancia entre sí (Figura 1.8, A), en un experimento TOCSY observaríamos una correlación entre A con C dado el acoplamiento de A-B y B-C. Así mismo, también se observaría correlación de B con D y A con D. En el tiempo de mezclado de este experimento se lleva a cabo la transferencia de magnetización en donde se transfiere la información de los núcleos acoplados generando correlaciones entre dos núcleos que están acoplados a un tercero en común. A este grupo de núcleos que correlacionan se les conoce como sistema de espín. En el contexto de las proteínas, cada aminoácido representa un sistema de espín. Por lo tanto, al analizarlas con un experimento TOCSY, deberíamos encontrar tantos sistemas de espín como aminoácidos tenga dicha proteína (Cavanagh, 1996).



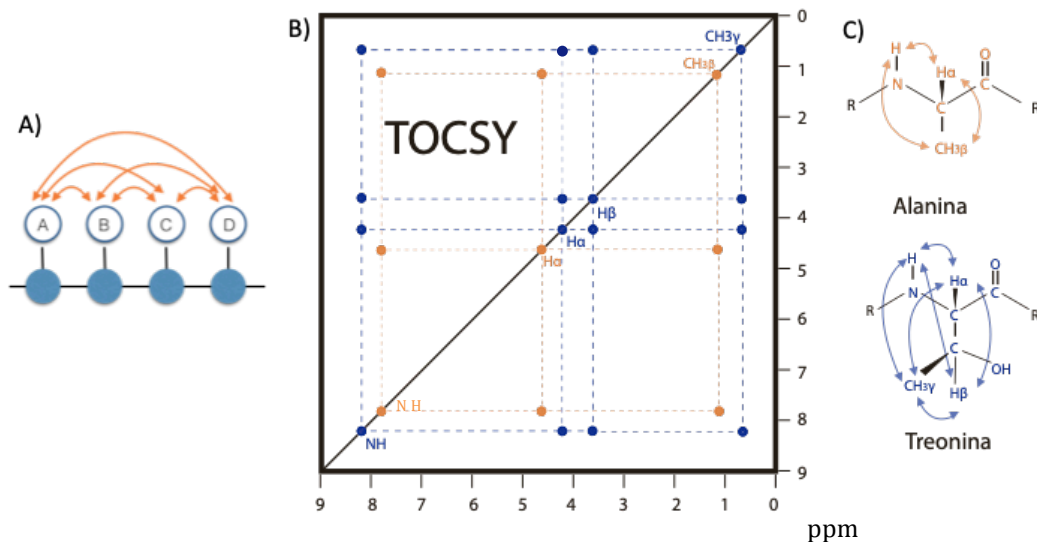


Figura 1.8. Sistema TOCSY de la alanina y treonina. A) Representación de un sistema de espín ABCD. B) Representación de espectro TOCSY de alanina y treonina. C) En la estructura de los aminoácidos se resaltan con flechas los protones que interactúan y forman el sistema de espín.

Regresando al ejemplo anterior, la representación del espectro TOCSY de la alanina se muestra en la figura 1.8 B. En esta se observa que ya no solo se ven interacciones del NH/H α y H α /CH $_3$, sino también, se observan las interacciones de NH con CH $_3$, el igual que en el caso de la treonina, en donde se generan todas la interacciones que dan lugar al sistema de espín (Figura 1.8, B).

1.2.6.3. NOESY

Uno de los experimentos más importantes para la determinación de la estructura tridimensional de una proteína es el NOESY. En este experimento se registra la interacción dipolar entre núcleos espacialmente cercanos. La distancia entre esos núcleos determinará la magnitud de la señal generada. Si los núcleos se encuentran a una distancia menor o igual a 5 Å, se observará correlación NOE. Con este experimento podremos encontrar la secuencia de los sistemas de espín (aminoácidos) previamente identificados y las interacciones a larga distancia fundamentales para determinar el plegamiento global de la proteína (Cavanagh, 1996).



1.2.7. Asignación de una proteína con experimentos bidimensionales

Existe una gran variedad de experimentos con los cuales se puede llevar a cabo la asignación de los aminoácidos de una proteína. A continuación describiré la metodología propuesta por el profesor Wüthrich en el 2003 utilizando los tres experimentos mencionados anteriormente. Cabe señalar que es fundamental contar con la secuencia de aminoácidos para llevar a cabo esta metodología.

1.2.7.1. Asignación secuencial de aminoácidos

El primer paso para obtener la estructura de una proteína por RMN es la asignación de todas las señales de los espectros de RMN (en este caso protones) de los aminoácidos que constituyen a la proteína. Imaginemos que tenemos un péptido o proteína de 60 aminoácidos, estamos hablando de aproximadamente 460 protones que debemos encontrar. Teóricamente, en el experimento COSY o TOCSY aparecen las señales de todos los protones de la proteína. La tarea es determinar que señal pertenece a cada uno de los protones.

Como se mencionó anteriormente, cada aminoácido es considerado un sistema de espín; el punto de partida es identificar uno para establecer a que aminoácido pertenece. Por su composición química y estructura, los aminoácidos tienen patrones particulares que los diferencian. La leucina por ejemplo (Figura 1.9), tiene señales del protón de amida (HN, 6.5-9 ppm), del protón alfa ($H\alpha$, 4-5 ppm), de los protones beta ($H\beta$, 1.5-2 ppm), del gamma ($H\gamma$, 1.3-1.6 ppm) y de los protones delta ($CH_3\delta$, 0.8-1.2 ppm). Este aminoácido tiene un sistema de espín muy característico por su gran cantidad de señales en menores partes por millón, similar a la isoleucina, arginina y lisina. Un sistema inconfundible es el de la valina, tienen las señales del NH (6.5-9 ppm), $H\alpha$ ($H\alpha$, 4-5 ppm), $H\beta$ (2-2.5 ppm) y dos señales de los protones de los metilos gamma en menores ppm (0.7- 1.2 ppm). Éste, al igual que la serina, glicina, treonina y alanina son sistemas inconfundibles. El resto de aminoácidos tienen sistemas similares como es el caso de la cisteína y el ácido aspártico los cuales se componen de NH, $H\alpha$ y dos $H\beta$. En la figura 1.9 se encuentra una representación de los sistemas de espín a partir del $H\alpha$ de los 20 aminoácidos, donde se pueden observar a grandes rasgos los desplazamientos químicos de sus protones y los patrones de señales que generan.



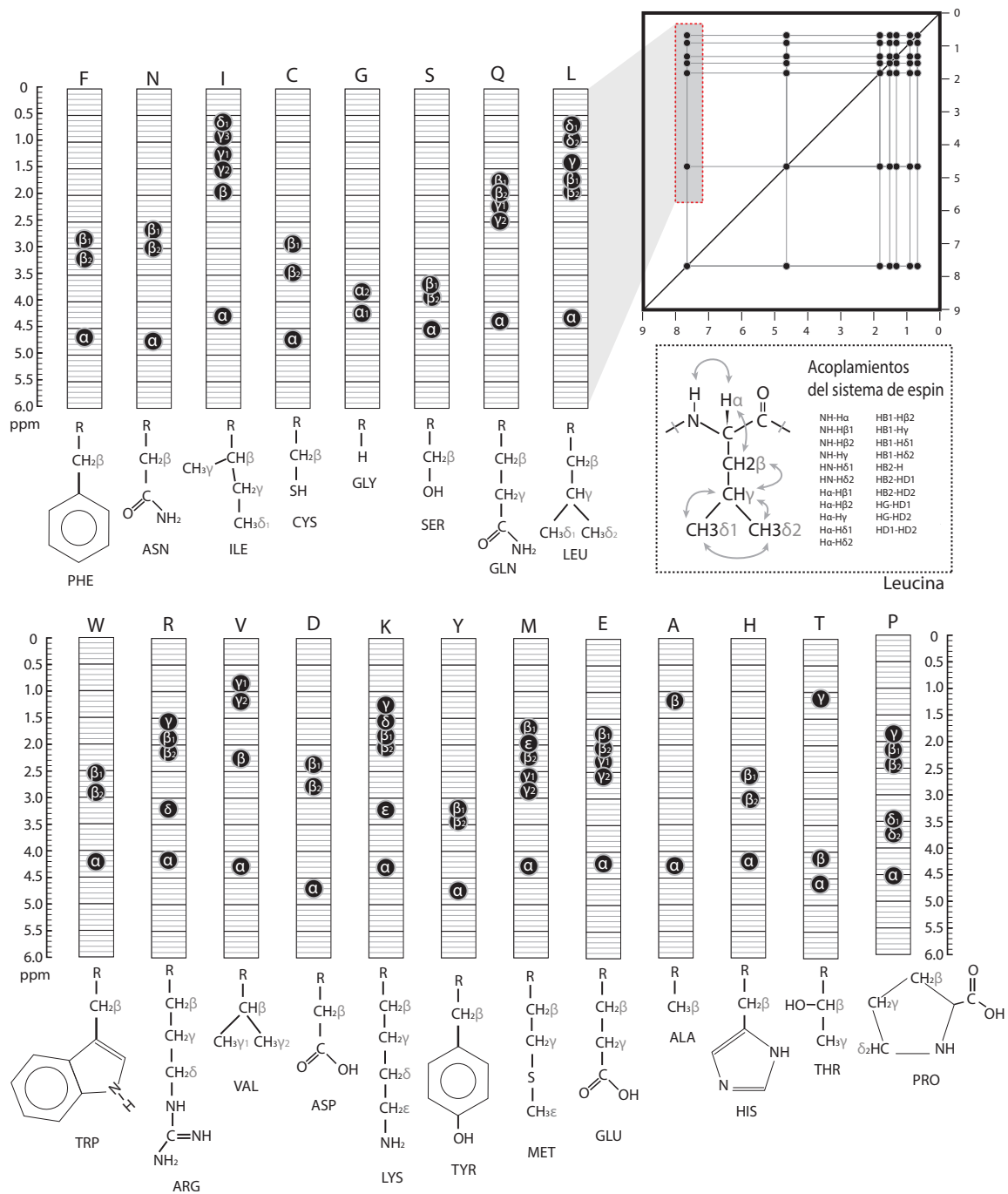


Figura 1.9. Sistemas de espín. Patrón de señales TOCSY de los 20 aminoácidos. En la figura se muestra un ejemplo de todas las interacciones TOCSY de la leucina, de ahí se toma el sistema de espín que se genera a partir del H α y se compara con los 20 aminoácidos.

Una vez que encontramos un sistema de espín en el experimento COSY y TOCSY debemos determinar a que aminoácido de nuestra secuencia pertenece. Suponiendo que



identificamos un sistema particular como el de una valina, podríamos estar seguros del aminoácido por el patrón de señales; sin embargo, ¿cómo identificar que valina es dentro de la secuencia, si no es la única en nuestra proteína? Para esto, nos ayudamos del experimento NOESY. Normalmente el NH de un aminoácido i tiene muy cerca a los protones del aminoácido $i-1$, no así los del aminoácido $i+1$, por ejemplo, en la figura 1.10 se puede observar la secuencia leucina (azul), cisteína (verde) y ácido aspártico (naranja), en la imagen se aprecia que el NH de la cisteína tiene cerca a los protones de la leucina (flechas azules) y el NH del aspartato tiene cerca los protones de la cisteína (flechas verdes). Los NOE's del aminoácido anterior son los que utilizamos para determinar la conectividad de los aminoácidos. Habiendo encontrado un sistema de espín, marcamos los desplazamientos químicos de sus protones y buscamos en que otros desplazamientos aparecen. Como se observa en la figura 1.10, las señales del sistema de espín de la leucina (azul), aparecen en el desplazamiento del NH de la cisteína (figura 1.10, línea punteada verde), esto indica que probablemente la leucina es el aminoácido anterior a la cisteína. Recordemos que los NOE's solo nos indica la proximidad de los átomos, y esto se debe a la proximidad por ser aminoácidos consecutivos, secuenciales o por el plegamiento global de la proteína, diferenciarlos es parte del análisis al momento de realizar la asignación.



A) Representación de interacciones NOESY a corta y larga distancia

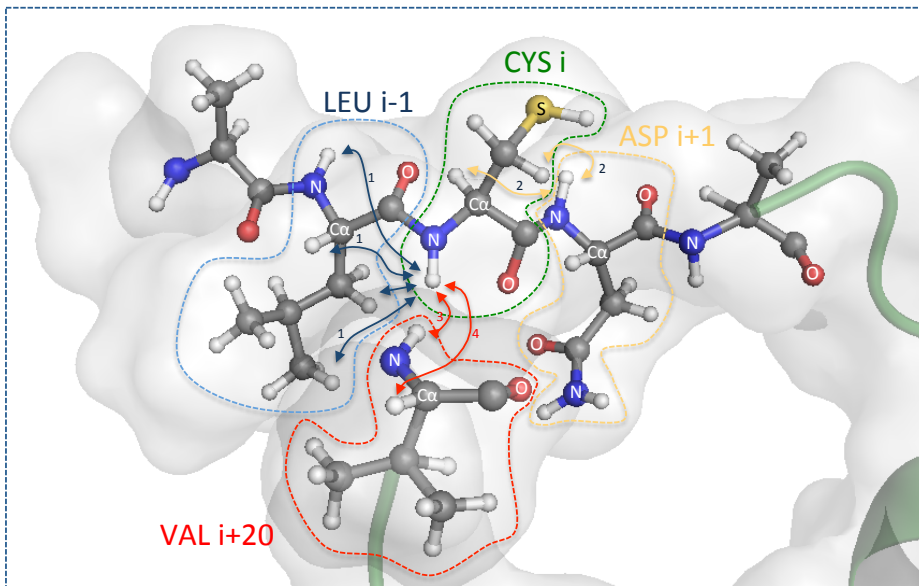
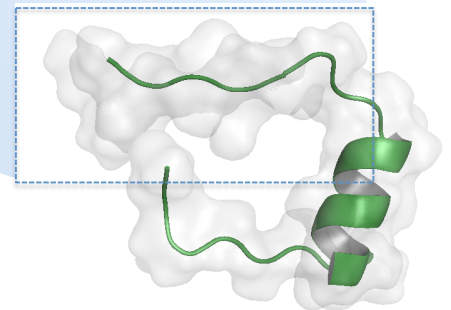
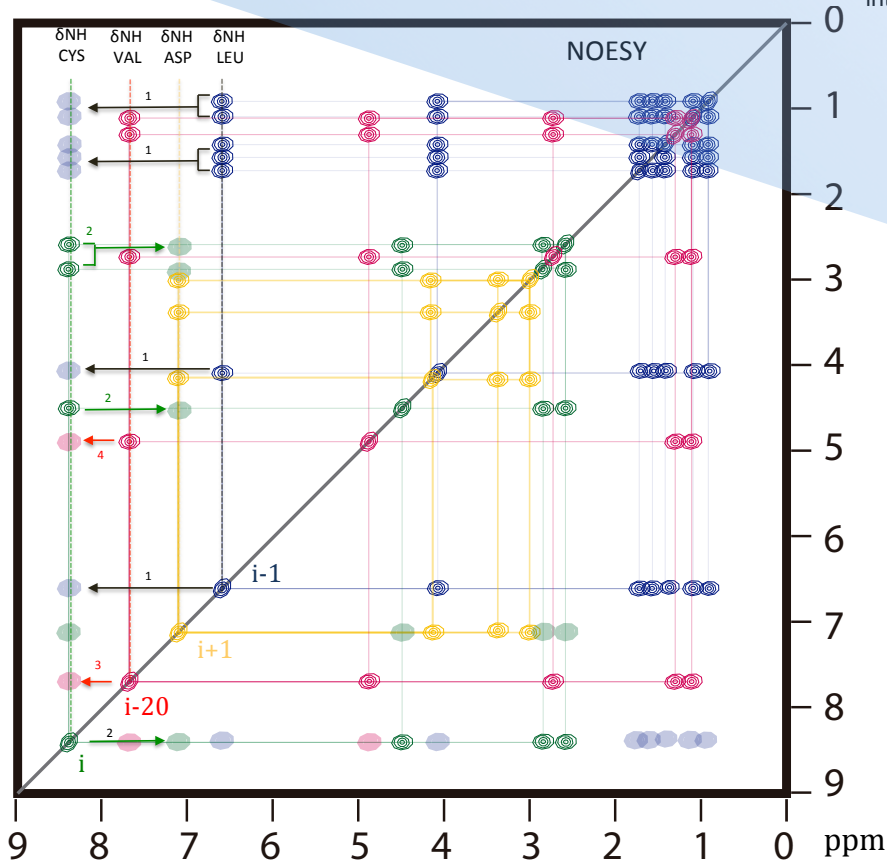


Figura 1.10. Asignación de NOE's secuenciales y larga distancia. El número 1 muestra interacciones secuenciales de los protones de la leucina con el NH de la cisteína A) y B). El sistema de espín de la leucina (señales azules) se observa en el δ del NH de la cisteína, lo que indica que la leucina es el aminoácido anterior a la cisteína. De igual forma las señales de la cisteína (verde) se ven reflejadas en el δ del NH de la asparagina, esta interacción se representa con el # 2. Los números 3 y 4 representan interacciones a larga distancia de los protones NH y H α de la valina con el NH de la cisteína A) y B). Las señales del NH y H α de la valina (rojo) se observan en el δ del NH de la cisteína lo que indica interacción a menos de 5 Å de distancia.

B) Representación de un espectro NOESY de los a.a. CYS, VAL, ASP y LEU



Este es el mecanismo de asignación de todas las señales o sistemas de espín que se observan en los espectros hasta completar todos o la mayor cantidad de aminoácidos que constituyen a la proteína.

1.2.7.2. NOE's a larga distancia

Una vez que terminamos con la asignación secuencial de todas las señales en el espectro TOCSY ya contamos en principio con las coordenadas de todos los protones, o la mayoría de los protones que constituyen a la proteína, el siguiente paso es asignar todas la señales NOE's no secuenciales.

En este punto ya se encuentran asignadas todas las señales de protones, que por su cercanía en la secuencia de aminoácidos generan interacciones NOE's. Todos los NOE's que faltan por asignar serán los generados por el plegamiento de la proteína, por lo tanto, al obtener estas señales tendremos la información necesaria para calcular una estructura tridimensional. En la figura 1.6 se observa un ejemplo claro de dos átomos secuencialmente separados que están espacialmente cercanos y generar una interacción NOE (NH y H α de la valina i+20 con el NH de la cisteína i).

Asignadas todas las señales NOE's, el siguiente paso es obtener las restricciones de distancia para poder calcular la estructura de la proteína. Como se mencionó anteriormente, la intensidad del NOE esta relacionada con la distancia entre los núcleos que interaccionan, si ahora nosotros integramos las señales del espectro NOESY tendremos las distancias necesarias para generar la estructura tridimensional de la proteína.

1.2.7.3 Cálculo de la estructura tridimensional

Como se mencionó anteriormente, los datos conformacionales clave para determinar la estructura tridimensional de una proteína son las restricciones de distancia derivadas del efecto nuclear Overhauser (Macura y Ernst, 1980). Para obtener estas restricciones, se deben asignar la mayor cantidad de picos fuera de la diagonal en el experimento NOESY, basándonos en la información de las señales previamente asignadas.



Uno de los programas más utilizados para generar estructuras de proteínas a partir de datos de RMN es CYANA (**c**ombined **a**ssignment and **d**ynamics **a**lgorithm for **N**MR **a**pplications). Este programa, genera la estructura calculando el valor de menor energía de una función objetivo que mide la relación entre la estructura y el conjunto de restricciones dadas. La función se define de modo que sea cero si y solo si se cumplen todas las restricciones de distancia experimentales, restricciones de ángulo de torsión y si todos los pares de átomos no unidos covalentemente, no presentan superposición estérica.

El algoritmo de minimización en CYANA se basa en un calentamiento simulado mediante dinámica molecular de ángulos de torsión, utilizando éstos como grados de libertad (Güntert et al. 1997). Debido que el panorama energético de una proteína es complejo y tiene muchos mínimos locales, el programa de temperatura es importante para el éxito y la eficiencia del cálculo. Este proceso ayuda a superar las barreras energéticas que se encuentra en la trayectoria de dinámica molecular. El protocolo estándar en el programa CYANA parte de una conformación con todos los ángulos de torsión tratados como variables aleatorias independientes, distribuidas uniformemente y consta de cinco etapas.

Primero se lleva a cabo una minimización inicial para reducir las interacciones de alta energía que pudieran sesgar el algoritmo de dinámica del ángulo de torsión. Después, comienza el cálculo de dinámica de ángulo de torsión a alta temperatura (10,000 K). Terminando este proceso comienza otra dinámica de enfriamiento lento cercano a la temperatura cero K. Después de este paso se incorporan todos los átomos de hidrógeno que inicialmente no se consideraban para una comprobación de superposición estérica para finalmente realizar una minimización final.

Debido a la precisión limitada de los valores de desplazamiento químico provocado por la sobreposición de señales, muchos picos no se pueden atribuir a un único átomo. Obtener un conjunto completo de restricciones de distancia de un espectro NOESY no es sencillo, y se convierte en un proceso iterativo en el que las estructuras preliminares, calculadas a partir de un número limitado de restricciones, sirven para reducir la ambigüedad de dichas asignaciones.



Otra de las características de CYANA es el proceso de asignación automática de los picos NOESY, el cual se puede llevar a cabo antes de la dinámica de ángulo de torsión. El módulo consiste en siete ciclos en los cuales, la información se transfiere exclusivamente a través de las estructuras intermedias, en el sentido de que la estructura molecular obtenida en un ciclo dado, se utiliza para guiar la asignación de NOEs en el siguiente ciclo. Los picos se interpretan en forma de un límite de distancia superior derivado del volumen de éstos. De este modo, se obtiene una restricción de distancia convencional para picos con una única asignación. De lo contrario, se genera una restricción de distancia ambigua que incorpora varias asignaciones. Los picos con una puntuación baja se descartan temporalmente para reducir los efectos negativos en la estructura resultante (Güntert, 2004).

En el primer ciclo de asignación automática se lleva a cabo la verificación de autoconsistencia de las señales NOESY, sin considerar una estructura. El segundo ciclo y los posteriores difieren del primero por el uso de criterios de selección adicionales para los picos y las asignaciones de NOE relacionadas con la estructura generada en el ciclo anterior. Debido a que la precisión de la determinación normalmente mejora con cada ciclo posterior, los criterios para aceptar asignaciones y restricciones de distancia se endurecen en los ciclos más avanzados del cálculo.

El cálculo utilizando el módulo de asignación automática se lleva a cabo de la siguiente manera:

1. El programa registra los datos de entrada experimentales: Secuencia de aminoácidos (.seq), lista de desplazamiento químico (.prot), lista de picos y volúmenes NOESY (.peaks), opcionalmente también puede leer restricciones conformacionales de otras fuentes (ángulos diedro .aco, puentes de hidrogeno .upl, etc).
2. Se crea una lista de asignación inicial. Para cada pico NOESY, se determinan una o varias asignaciones iniciales en función de los desplazamientos químicos dentro de un rango de tolerancia definido por el usuario.
3. Para cada pico NOESY individual, las asignaciones iniciales se ponderan con respecto a varios criterios, y las asignaciones iniciales con puntaje general bajo se



descartan. Los criterios incluyen la consistencia entre los valores de desplazamiento químico y la posición de los picos, la autoconsistencia dentro de toda la red de NOEs (Figura 1.11) y, si está disponible la compatibilidad con la estructura generada del ciclo anterior. El requisito de que cada asignación de NOE debe integrarse en la red de todas las demás asignaciones, hace que el anclaje de la red sea un enfoque sensible para detectar restricciones erróneas y "solitarias" que podrían restringir artificialmente partes no estructuradas de la proteína.

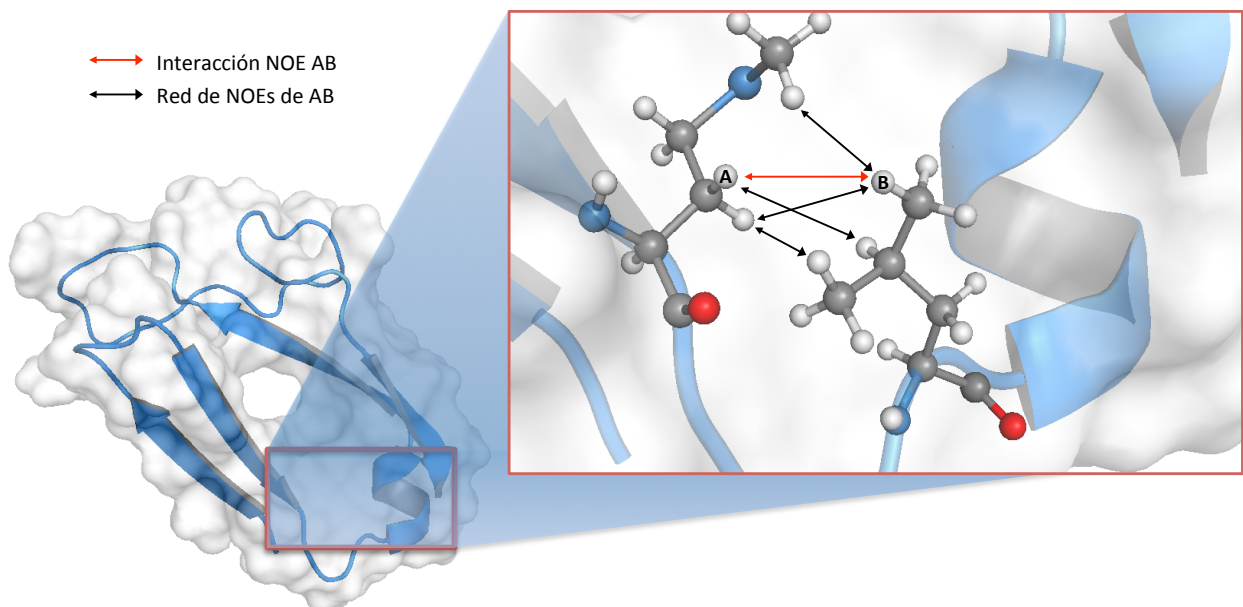


Figura 1.11. Red de NOEs. La red de NOE's es el conjunto de restricciones generadas que definen zonas de interacción en la proteína. La flecha roja muestra una interacción NOE entre el átomo A y B, las flechas negras son la red de NOEs que validan la asignación de la interacción A con B. La red de NOEs, resalta que las restricciones asignadas correctamente forman un subconjunto autoconsistente en una red de restricciones de distancia.

4. Se calibran las restricciones de distancia con los volúmenes o intensidades de los picos NOESY.
5. Se eliminan los picos NOESY espurios. Sólo se conservan los picos que tienen al menos una asignación con una puntuación de anclaje de red por encima de un



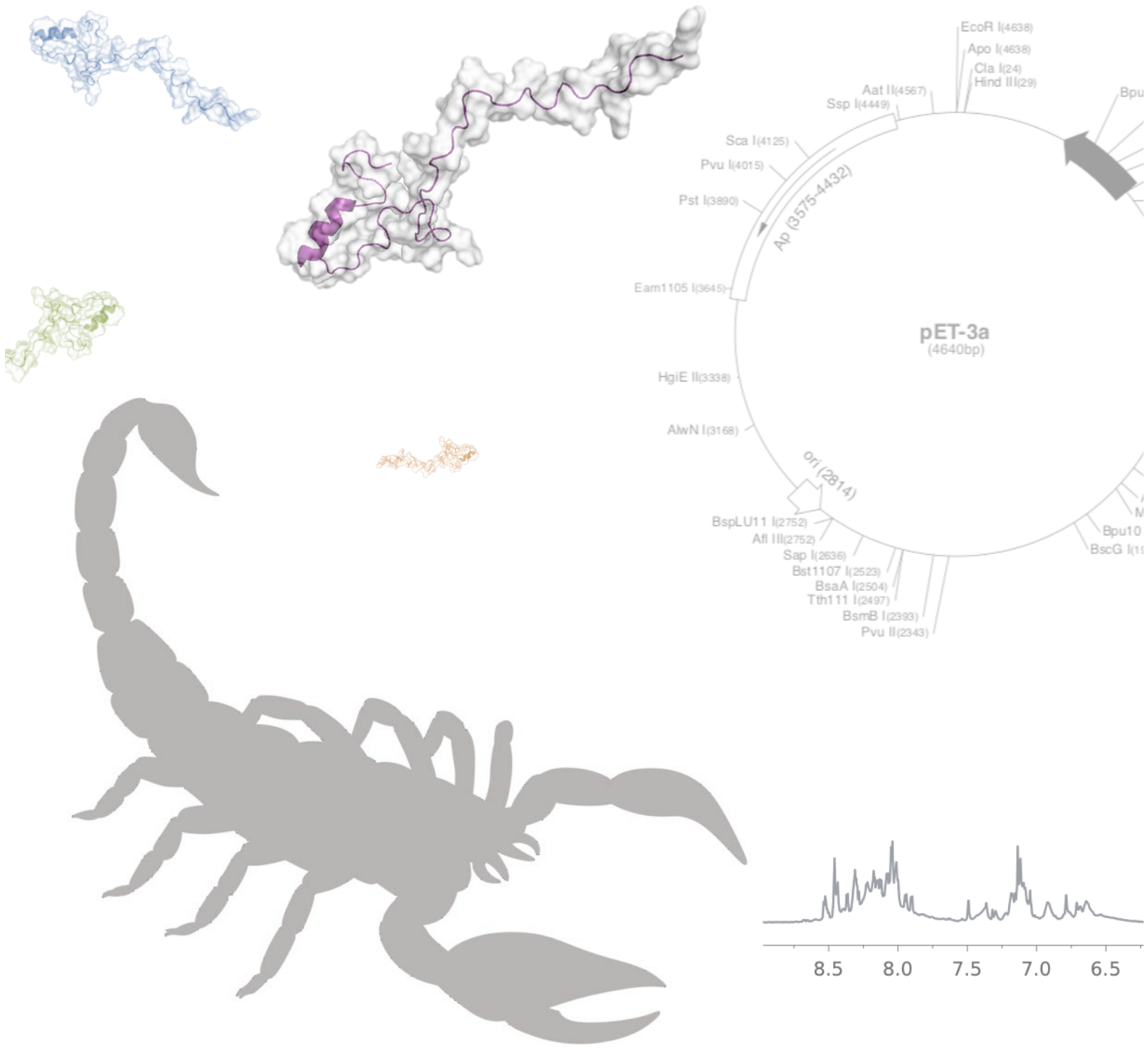
umbral definido por el usuario, y que son compatibles con la estructura intermedia generada en el ciclo anterior (ciclos 2, 3, ...).

6. Combinación de restricciones. En los ciclos 1 y 2 las restricciones de distancia de largo alcance no relacionadas a priori se combinan en nuevas restricciones de distancia virtual. La mala interpretación del ruido y los artefactos espectrales pueden generar malas interpretaciones que en el comienzo del cálculo pueden ser críticas, recordemos que en el ciclo inicial no contamos con una estructura preliminar disponible para el filtrado basado en la estructura. La combinación de restricciones tiene como objetivo, minimizar el impacto de tales imperfecciones en la estructura resultante a expensas de una pérdida temporal de información. Independientemente de la presencia de asignaciones erróneas adicionales, este mecanismo implica que tales restricciones combinadas tienen una probabilidad menor de ser erróneas que las restricciones originales correspondientes.
7. Cálculo de estructura. Después de 7 ciclos con las características anteriores, finalmente se lleva a cabo la dinámica de ángulo de torsión. En este último paso, las restricciones de distancia de los NOEs con asignaciones múltiples y las que resultan de la combinación de restricciones se introducen como restricciones ambiguas y son consideradas en el cálculo estructural (Güntert, 2004).

Al finalizar cada ciclo se genera una lista de asignaciones de picos NOESY, una lista de comentarios sobre decisiones de asignación individuales que pueden ayudar a reconocer posibles artefactos en los datos de entrada y un conjunto de estructuras. En el ciclo final, un paso de filtrado adicional asegura que todos los NOEs tengan asignaciones únicas a un solo par de átomos de hidrógeno, o se eliminen de la entrada para el cálculo de la estructura. Esto permite el uso directo de las restricciones de distancia NOE en programas de afinamiento y análisis posteriores que no manejan restricciones de distancia ambiguas.



TsTx-K β : Un péptido disfuncional



2.1 KTx- β y TsTx-K β

En 1991 Rogowski y colaboradores aislaron del veneno del alacrán *Tityus serrulatus* una serie de toxinas con actividad bloqueadora sobre sinaptomas en neuronas de rata. En estudios posteriores, llamó la atención la actividad que presentaban dos de estas toxinas, TsTx K2 y TsTx K4, las proteínas mostraban unas concentraciones de inhibición sobre canales de potasio de 8 y 30 nM respectivamente. Al determinar sus secuencias de aminoácidos resaltó que TsTx K4 contenía 60 aminoácidos, discrepando con las observaciones de toxinas conocidas hasta ese momento (Rogowski et al. 1994).

Con el paso del tiempo, se descubrieron toxinas como AaTx-K β con una gran homología con TsTx K4, surgiendo la necesidad de generar una nueva subfamilia de toxinas con actividad sobre canales de potasio conocidas como β -Ktx (Legros et al. 1998). TsTx K4 se renombró a TsTx-K β y fue la representante de este tipo de nuevas proteínas conocidas comúnmente como toxinas de cadena larga o β -toxinas.

Los estudios continuaron y se encontró el gen que codifica para la toxina TsTx-K β , el cual, produce un polipéptido de 87 a.a. constituido de un péptido señal en la región 1-19, un propéptido en la región 20-27 y el péptido maduro a partir del aminoácido 28 (Figura 2.1).

MERKLALLLILGMVTLASCGLREKHHVQKLVALIPNDQLRSILKAVVHKVAKTQFGCPAYEGYCNHHCNDIERKDGEGCHGFKCKCAKD
 -----Péptido señal-----Pro-péptido-----Péptido maduro-----

Figura 2.1. Secuencia de aminoácidos de la proteína TsTx-K β

En el 2004, Zhu y Tytgat realizaron un estudio donde compararon las secuencias de aminoácidos de las β KTx existentes hasta ese momento; la comparación fue con péptidos tipo escorpión famosos por su actividad antimicrobiana. Ellos encontraron una estrecha relación al observar la homología que existían entre estas proteínas principalmente por su tamaño. En el estudio, llamó la atención como las cisteínas no se encuentran distribuidas a lo largo de la cadena polipeptídica como ocurre normalmente con este tipo de toxinas. Como se observa en la figura 2.2, las cisteínas están concentradas de la mitad de la



secuencia hacia el carboxilo terminal, y sus posiciones son típicas del motivo estructural α/β observado ya en las α -KTx (Zhu et al., 2004). Por otro lado, una predicción de estructura revela que el extremo amino terminal puede ser capaz de adoptar una estructura de dos hélices α interrumpidas por un giro (Figura 2.2).

```

-----10-----20-----30-----40-----50-----60-----70-----
TsTXK-beta      G-LREKHVQKLVA-LIPNDQLRSILKAVVHKVAKTQFGCPA---YEGYCNDCNDIE-RKDGECHGFKCKCAKD-----
AaTXK-beta      G-LREKHVQKLVKYAVPVGTLRLTIQTAVHKLKGTQFGCPA---YQGYCDDHCQDIK-KEEGFCHGFKCKCGIPMGF--
BmTXK-beta-2    G-LREKHFQKLVKYAVPEGLRTIIQTAVHKLKGTQFGCPA---YQGYCDDHCQDIK-KEEGFCHGFKCKCGIPMGF--
MeuTXK-beta-1  G-FREKHFQRFVKYAVPESTLRTVLTAVHKLKGTQFGCPA---YQGYCDDHCQDIE-KKEGFCHGFKCKCGIPMGF--
beta-Ktx 2      G-KGKEILGKIKEKIIE--AKDKMKAGWERLTSQSEYACPA---IDKFCEDHCA--AKKAVGKCDDFKCNCKIKL----
BmTXK-beta      G-RGKEIMKNIKEKLTE--VKDKMKHSWNKLTSMSEYACPV---IEKWCEDHCA--AKKAIGKCEDTECKCLKLRK---
Hge-scorpine    GWMSEKKVQGILDKKLPEGIIRNAAKAIVHKMAKNQFGCFANVDVKGDCRKRCKAED-KE-GICHGTKCKCGVPIISYL-
Ev37            GLINEKKVQQYLDEKLPNGVVKGALKSLVHKAANKQNLCAFNVDTVGMCDADCKRQG-KAKGVCHGTKCKCDVELSYKK
Piscorpine-1    KWFNEKSIONKIDEKIGKNFLGGMKAVVHKLAKNEFMCVANVDMTKSCDTHCQKAS-GEKGYCHGTKCKCGVPLSY--
Scorpine        GWINEEKIQKKIDERMGNTVLGGMAKAIIVHKMAKNEFQCMANMDMLGNCEKHCQ-TSG-EKGYCHGTKCKCGTPLSY--
-aaaaaaaaaa---aaaaaaaaaaaaa-----aaaaaaaaa-----BBBBBBB-BBBBBBB-----

```

Figura 2.2. Alineamiento de las diferentes toxinas con similitudes a la TsTx-K β

El análisis de Zhu y Tytgat concluye que tanto las β KTx como las escorpinas presentan dos dominios estructurales y cada uno con una función independiente, una actividad antimicrobiana o citolítica para el extremo amino (comprobada en las escorpinas) y para el extremo carbonilo una actividad bloqueadora de canales de potasio (comprobada en TsTx-K β).

Años más tarde García y colaboradores en el 2007 proponen agrupar a las β KTx y escorpinas, generando una nueva subclasificación basados principalmente en la similitud de sus secuencias de aminoácidos. El primer grupo engloba a todas las toxinas homologas a TsTx-K β como son AaTx-K β , BmTx-K β 2 y TcoTx-K β . El segundo grupo son toxinas con similitud a BmTx-K β , la cual es una toxina que tiene un tamaño similar a TsTx-K β y las escorpinas, pero con homología baja. En el tercer grupo se encuentran las escorpinas y similares.

A pesar de que día a día aumenta el número de reportes de toxinas encontradas pertenecientes a las β KTx, se cuenta con muy pocos trabajos que las estudien a fondo, sobretodo a un nivel estructura terciaria. En un estudio realizado en el 2008, digieren la escorpina HgeScplp1 y evalúan la actividad de sus fragmentos, encontrando que el dominio carboxilo terminal presenta actividad bloqueadora de canales iónicos y su contraparte el amino terminal una actividad citolítica (García et al., 2008). Sin embargo, en



otro estudio en donde obtienen de forma recombinante la escorpina Ev37 y los dos dominios estructurales por separado (N-terminal y C-terminal), reportan una inexistente actividad citolítica o antimicrobiana, y el efecto sobre los canales de potasio sólo lo observan cuando evalúan la proteína completa. En este estudio concluyen que tanto el dominio amino terminal como el carboxilo terminal están implicados en la interacción de Ev37 sobre el canal Kv1.3 y que además esta interacción podrían contribuir a su selectividad, contradiciendo las propuestas anteriores en las que se proponen actividades independientes para cada dominio (Feng et al., 20013).

Otro ejemplo es el caso de la proteína MeuTx-K β (72 a.a.). Esta β KTx se produjo recombinantemente y por un corte inespecífico de la proteasa usada; se obtuvieron los péptidos correspondientes al fragmentos N(1-21 a.a.) y C(22-72 a.a.). Evaluaron sus actividades y reportan una débil actividad frente al canal Kv y un claro efecto hemolítico en el dominio del extremo carboxilo así como en el péptido completo, indicando que esta capacidad citolítica no se observa en el N-terminal (Zhu et al. 2010).

Quedan muchas preguntas sin responder acerca de las β KTx, por ejemplo, aún no queda clara la función que sus dominios desempeñan, si estos actúan de forma conjunta o no, si existe una relación directa entre el mecanismo de acción de las β KTx y las escorpinas, pero sobre todo no hay estudios estructurales que nos permitan comprender mejor su naturaleza. De los tres grupos que propone García y colaboradores, no cabe duda que el menos estudiado tanto biológicamente como estructuralmente es el grupo 1. Curiosamente, en este grupo se encuentra la proteína más representativa de las β KTx, la primera en ser aislada y por la cual se inicio la clasificaron de toxina con cadena larga y actividad sobre canales de potasio, la TsTx-K β . En este estudio proponemos generar un sistema de expresión para obtener esta proteína de forma recombinante, y llevar a cabo su caracterización estructural y funcional.



2.2. Hipótesis

Sí se optimizan las condiciones en las que TsTx-K β es producida de forma recombinante, podremos obtenerla para su caracterización estructural. Su estructura estará compuesta por un dominio desestructurado N-terminal y uno C-terminal con el típico motivo α/β observado en la mayoría de toxinas con actividad sobre canales de potasio.

2.3.1 Objetivo general

Encontrar el sistema de expresión adecuado para obtener de forma recombinante la toxina TsTx-K β y determinar su estructura tridimensional por resonancia magnética nuclear.

2.3.2. Objetivos particulares

- Construir el sistema de expresión para la obtención de la proteína recombinante.
- Determinar el sistema donde se obtenga la proteína de manera plegada y con la formación correcta de los puentes disulfuro.
- Determinar de manera experimental la conectividad de los puentes disulfuro.
- Obtener los experimentos de RMN monodimensionales y bidimensionales.
- Procesar y asignar los espectros obtenidos.
- Determinar la estructura de la toxina TsTx-K β .
- Realizar pruebas de actividad biológica de TsTx-K β .



2.4.1. Diseño de los sistemas de expresión

Por la complejidad de la proteína, en el transcurso del proyecto se utilizaron cuatro sistemas de expresión diferentes; en todos los casos el diseño del gen se realizó con base en la secuencia de aminoácidos reportada en UniProt (P69940). El gen fue ensamblado y adquirido a GenScript ®. El primer sistema se realizó con el objetivo de obtener la proteína de manera soluble y se utilizó el pEt-32a (rTsTx-K β), el gen se colocó entre los sitios BamH1 y KpnI generando una proteína de fusión que cuenta con un dominio de tiorredoxina, un His-tag y el sitio de corte para la proteasa trombina (figura 19). En la literatura hay reportes de proteínas expresadas solo con la etiqueta de histidinas para su purificación; esta construcción aumenta el rendimiento al evitar pasos como el corte enzimático. Con esto en mente, en el segundo sistema (rHTsTx-K β) el gen se colocó en el pET-3a entre los sitios NdeI y BamH1. La proteína se diseñó con el His-tag en el extremo amino terminal de la secuencia de TsTx-K β separados por una metionina para poder remover, si se desea, la etiqueta de histidinas mediante la reacción con bromuro de cianógeno. Dado que no se conocen las rutas de plegamiento de estas proteínas, queríamos asegurarnos que el His-tag en el amino terminal no interfería en el plegamiento de TsTx-K β . Para esto, diseñamos un tercer sistema (rTsHTx-K β) idéntico al segundo, pero con el His-tag unido al extremo carboxilo de la TsTx-K β . Finalmente un cuarto sistema fue construido (rTs37Tx-K β) con el objetivo de obtener la parte más estructurada de la toxina y donde se encuentra el motivo estructura α/β , para obtenerlo se utilizó el vector de clonación pET-32a insertando el fragmento de TsTx-K β del aminoácido 24 al 60 entre los sitios BamH1 y KpnI. Esta construcción al igual que rTsTx-K β cuenta con un dominio de tiorredoxina, un His-tag, y se cambió el sitio de corte de trombina por la proteasa TEV. En la figura 2.3 se observa un esquema de las proteínas de fusión y las secuencias de aminoácidos que generan las cuatro construcciones.



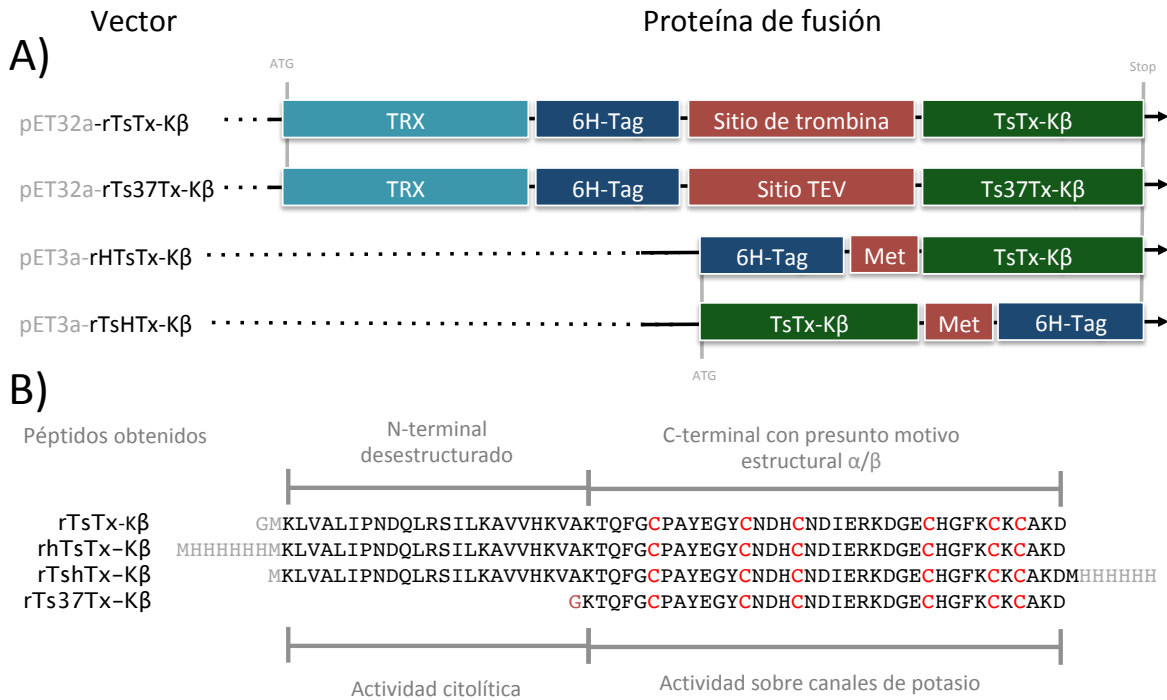


Figura 2.3. Construcciones para la obtención TsTx-K β . A) Vectores de clonación B) Secuencias producidas.

Los plásmidos de todas las construcciones fueron transformados en diferentes cepas de *E. coli* con la siguiente metodología:

- I. Se mezcló 1.5 μ L (60 ng/mL) de la disolución donde se encuentra el plásmido de trabajo a 50 μ L de una suspensión de células competentes de las cepas seleccionadas.
- II. Esta mezcla se incubó por un periodo de 30 min en un baño de hielo.
- III. Transcurrido este tiempo, se colocó en un baño térmico a 42 $^{\circ}$ C por 45 s y se dejó en hielo de 2 a 3 min.
- IV. Se le agregó 300 μ L de medio de cultivo LB y se deja incubar a 37 $^{\circ}$ C por una hora.
- V. Se tomó 50 μ L de la suspensión y se siembran en medio LB agar que contiene el antibiótico al cual es resistente la cepa y el plásmido dejándolo incubar de 14 a 16 horas a 37 $^{\circ}$ C.

Se eligió la cepa donde se veía una mayor sobreexpresión. También se monitorearon condiciones de temperatura y concentraciones de IPTG. En todos los caso el crecimiento



de las células se realizó a partir de una colonia aislada obtenida de la transformación. Se picó la colonia con una punta estéril y se resuspendió en 100 mL de medio LB con el antibiótico correspondiente y se dejó incubar 16 horas a 37 °C. 50 mL de este inóculo se transfirió a 1500 mL de medio LB con antibiótico y se incubó a 37 °C monitoreando su crecimiento hasta alcanzar una densidad óptica (DO_{600}) entre 0.6 y 0.8 unidades de absorbancia. En la siguiente sección se describirán las condiciones ya optimizadas para la obtención de cada proteína.

2.4.2. Expresión y purificación de rTsTx-K β .

rTsTx-K β fue producida en la cepa BL21 trx de *E. coli*. Alcanzada la densidad óptica, se indujo con 0.25 mM de IPTG y se dejó por 18 horas a 18 °C. Transcurrido el tiempo se colocó el medio en frascos los cuales fueron centrifugados a 6,500 rpm por 15 min y las células se resuspendieron en 50 mL de solución amortiguadora de lisis (50 mM de tris pH 8.0., 300 mM NaCl).

Las células resuspendidas se sometieron un programa de choques sónicos por periodos de 20 s a 45 W con descansos de 40 s por un tiempo efectivo de 10 min a 4 °C. El lisado celular se centrifugó por 45 min a 16,500 rpm conservando el sobrenadante. La purificación de la proteína de fusión (fracción soluble) se llevó a cabo mediante CAMQ en una columna HipTrap®. La muestra se filtró con una membrana de 0.45 μ m y se pasó por la columna a un flujo de 5 mL/min cargada previamente con Ni²⁺. Se realizó un lavado con cinco volúmenes de columna (25 mL) con solución amortiguadora de lisis (50 mM de tris pH 8.0, 300 mM NaCl) y posteriormente con disoluciones que contenían una mezcla de solución amortiguadora de lisis más imidazol a 5, 25, 50, 100, 200 y 300 mM.

La fracción con la proteína de fusión se dializó para remover el imidazol en una membrana de diálisis de 10 kDa contra 500 mL (tris 50 mM, pH 9.0), con dos cambios de dos horas cada uno. Ya en estas condiciones se realizó el corte enzimático por 3 horas utilizando por cada 2 mg de proteína de fusión 5 μ L de suspensión de agarosa donde se encuentra fija la trombina (SIGMA®). Transcurrido el tiempo se separó la trombina por centrifugación y se llevó a una concentración de 300 mM de NaCl. Inmediatamente después, se pasó por la columna HipTrap® con un flujo de 5 mL/min y se lavó con 50 mL



de solución amortiguadora de lisis (tris 50 mM pH 9.0, 300 mM de NaCl y 300 mM de imidazol).

Para concentrar la muestra se pasó por un cartucho de fase sólida Strata C18-T. Posteriormente, se eluyó con diferentes fracciones de acetonitrilo/agua 10-100% ACN. Todas las fracciones se liofilizaron y prepararon para la purificación final en una columna de fase reversa C5 adaptada a un sistema de cromatografía de líquidos de alta eficiencia. Se utilizó una fase móvil en gradiente lineal de 15 a 95% de agua-acetonitrilo al 0.05% de TFA, todo esto acoplado a un sistema de detección UV registrando a 230 y 280nm.

2.4.3. Expresión, purificación y replegamiento de rHTsTx-K β y rTsHTx-K β

Tanto rHTsTx-K β como rTsHTx-K β fueron expresadas en cepas de *E. coli* C41, alcanzada la densidad óptica se indujeron las proteínas con 0.5 mM de IPTG y la producción se llevó a cabo por 6 horas a 30 °C. Los paquetes celulares se obtuvieron por centrifugación, como se describió anteriormente. Las células se resuspendieron en solución amortiguadora (50 mM de tris pH 8.0, 300 mM NaCl) con 1 Molar de urea. Se aplicó el mismo programa de sonicación y se separó la fracción insoluble de la soluble (centrifugación 15 min a 10000 g). Las dos proteínas se extrajeron de la fracción insoluble por lo que esta fracción se lavó por triplicado con 30 mL de tris 50 mM pH 8.0 y finalmente se resuspendió con una solución de 1% SDS tris 50 mM pH 8.0 y 300 mM de NaCl. La fracción resuspendida se sonó por segunda vez con el mismo programa y se pasó por la columna de níquel con el procedimiento descrito en la sección anterior. La fracción donde se encontraba la proteína (300 mM de imidazol) se diluyó 10 veces y se volvió a pasar por la columna de níquel con el objetivo de remover en su totalidad el SDS.

Para el replegamiento de las dos proteínas se les adicionó lo suficiente para alcanzar una concentración de 10 mM de DTT y 5 M de GndCl, dejándose en agitación por 3 horas a temperatura ambiente. Transcurrido el tiempo se dializaron contra 500 mL de Tris 50 mM pH 9.0 y 500 mM de GndCl, con dos cambios por 2 horas cada uno en una membrana de diálisis de 3 kDa. Esta fracción se adicionó por goteo a una solución de replegamiento que contenía tris 50 mM pH 9.0, 500 mM de GndCl, glutatión oxidado 1 mM y glutatión reducido 3 mM y se dejó por 4 días a temperatura ambiente y agitación. Para concentrar



la proteína se aprovechó que contenían la etiqueta de histidinas y se pasaron por la columna de níquel eluyéndolas con una solución que contenía tris 50 mM pH 9.0 y 300 mM de imidazol.

Finalmente las proteínas se purificaron con una columna C5 como se describe en la sección anterior.

2.4.4. Expresión y purificación de rTs37Tx-K β .

rTs37Tx-K β fue expresada en *E. coli* C41. Alcanzada la densidad óptica se indujo con 0.5 mM de IPTG y se dejó por 6 horas a 30 °C. La obtención del paquete celular así como la lisis y la purificación por la columna de níquel se llevó a cabo como se describe en la sección 2.4.2. La fracción con la proteína (300 mM de imidazol) se diluyó 10 veces con una solución que contenía 50 mM de tris pH 8.0, 300 mM NaCl, 5 mM MgCl₂, 0.25 mM DTT y 5 % glicerol y se le adicionó la enzima TEV con una proporción de 1 mg de enzima por 50 mg de proteína de fusión; esta reacción se llevó a cabo por 16 horas a temperatura ambiente. Transcurrido el tiempo se pasó por la columna de níquel y se conservó la fracción sin afinidad. La proteína fue concentrada en un centrifugón de 3,000 Da y se purificó finalmente por una columna C5 con el procedimiento descrito anteriormente en la sección 2.4.2.

2.4.5. Reacción de bromuro de cianógeno

Para eliminar la etiqueta de histidinas en las construcciones rHTsTx-K β como rTsHTx-K β y el GM residual del corte con trombina de rTsTx-K β se llevó a cabo una reacción con bromuro de cianógeno. Las proteínas previamente liofilizadas se resuspendieron en una mezcla de agua/TFA en proporción 1:1 con el volumen necesario para alcanzar una concentración de 10 mg/mL. La solución se cubrió de la luz y se hizo pasar un flujo de nitrógeno por 15 min, transcurrido el tiempo se adicionó el CNBr suficiente para alcanzar una relación molar 1:100 por cada metionina que se desea cortar (en este caso es sólo una). Las reacciones se dejaron por 16 horas a temperatura ambiente con atmósfera de nitrógeno y se pararon con tres volúmenes de agua. Después se aumentó el pH de la mezcla de reacción con una solución saturada de tris base hasta alcanzar un pH neutro.



El exceso de tris fue eliminado por diálisis (1 kDa) contra agua y posteriormente purificado por cromatografía de líquidos de alta resolución con el método descrito anteriormente.

2.4.6. Caracterización de los péptidos tipo TsTx-K β

Todas las muestras obtenidas se analizaron por espectrometría de masas MALDI-TOF utilizando como matriz una solución saturada de ácido α -ciano-4-hidroxicinámico en relaciones con la muestra de 1:5, 1:10 y 1:20 al 1% de TFA. Esta medición se realizó en un equipo MALDI-TOF Bruker Daltonics Microflex en el laboratorio de servicios analíticos del Instituto de Química de la UNAM, el modo de adquisición fue lineal, en polaridad positiva utilizando 120 pulsos con una intensidad del 20% y una ventana espectral de 1,500 a 13,000 Da.

Todas las muestras analizadas por RMN fueron liofilizadas y resuspendidas en 200 μ L de agua con 5% de D₂O. Los experimentos se realizaron en un equipo Bruker ASCEND 700 MHz equipado con una criosonda de triple resonancia (¹H, ¹³C y ¹⁵N) a 298 K. El experimento TOCSY fue obtenido con un tiempo de mezclado de 80 ms mientras que los espectros NOESY se adquirieron usando tiempos de mezclado de 150 y 300 ms. Se usó la supresión de agua por pre-saturación y DPGSE durante la obtención de los diferentes experimentos. Una vez adquiridos los FID's de cada experimento, se procesaron los datos en el programa nmrPipe y nmrDraw (Delaglio et al. 1995).

También se realizaron pruebas de dicroísmo circular, los espectros de los péptidos se obtuvieron en un espectropolarímetro Jasco J-1500 en concentraciones entre 200 y 300 mg/mL disueltos en agua desionizada. Los espectros se adquirieron a 20°C y se midieron longitudes de onda de 240–190 nm utilizando una celda de cuarzo de paso óptico de 0.1 cm de espesor. Todos los datos se registraron en intervalos de 1 nm con una velocidad de barrido de 20 nm/min realizándolos por triplicado.



2.5.1. Purificación de rTsTx-K β

En trabajos previos ya se había observado la dificultad de trabajar con TsTx-K β . En el 2014 se obtuvo de manera soluble la proteína de fusión (Reynoso, 2014); sin embargo, al retirarle la tiorredoxina, la proteína precipitaba y no se logró obtener soluble para su estudio. En dicho trabajo se concluyó que la proteína tenía un mal plegamiento. Con esto en mente, el principal objetivo de esta investigación fue encontrar un sistema con el cual se lograra obtener la toxina de manera soluble y así determinar su estructura por RMN. La primera acción fue optimizar las condiciones de la producción para que la proteína se obtuviera soluble desde la producción en la bacteria y posiblemente con el plegamiento adecuado.

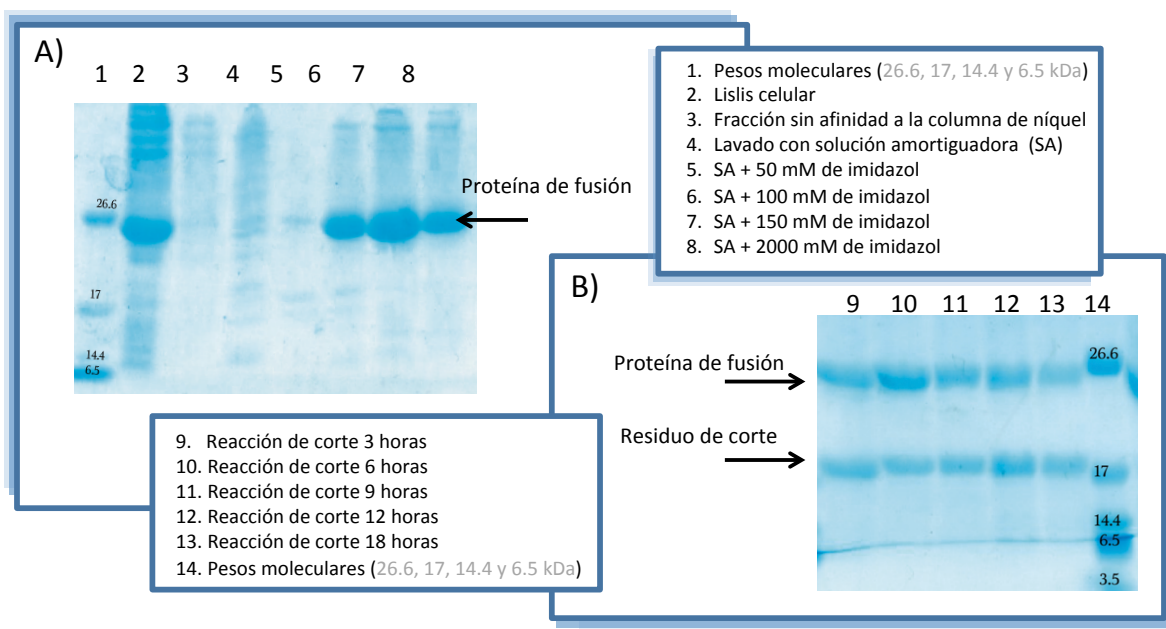


Figura 2.4. Purificación de rGMTsTx-K β . Geles de poliacrilamida al 14% donde se observa la purificación de rGMTsTx-K β (A) y el corte enzimático con trombina (B).

Con la construcción rTsTx-K β los mejores resultados se consiguieron al reducir tanto la concentración del inductor (IPTG, 0.250 mM) como la temperatura de producción (18 °C). En la figura 2.4 se encuentra el gel de poliacrilamida (15 %) donde se observa la producción y purificación de la toxina rTsTx-K β . En el carril 2 se encuentra la lisis celular donde se observa la sobrexpresión de una proteína con la masa esperada para la construcción de fusión (23,953 Da), en el carril 6, 7 y 8 se muestra como la proteína



empieza a eluir con las diferentes concentraciones de imidazol. El corte enzimático para separar la toxina de interés de la tiorredoxina se puede observar del carril 9 al 13 de la misma figura. Se observa que la reacción no se lleva a cabo en un 100% y con el paso del tiempo se empieza a formar un poco de precipitado en la muestra. Se decidió entonces trabajar con tiempos de 3 horas donde no existe precipitado y hay proteína cortada.

Después de corte con trombina la mezcla de corte se pasó por la columna de níquel y TsTx-K β no tuvo ninguna afinidad por ésta. Se concentró con ayuda de una columna de fase sólida C-18 y la proteína eluyó con 30 % de ACN. La fracción se liofilizó y de disolvió con agua para su inyección a la columna. De esta inyección se obtuvo un pico mayoritario con un tiempo de retención de 18.3 min (Figura 2.5) el cual tenía la masa esperada para la TsTx-K β . La masa teórica de esta proteína es de 6,904.9 Da en estado oxidado y nosotros obtuvimos 6,905.6 Da corroborando así que teníamos la proteína de interés y además contaba con la formación de los seis enlaces disulfuro.

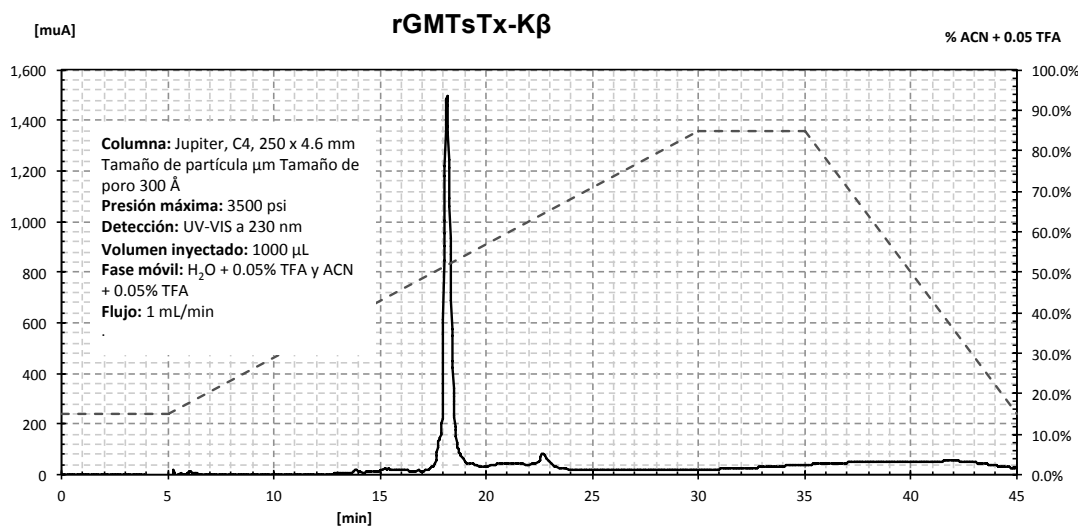


Figura 2.5. Cromatograma de rGMTsTx-K β . En la línea punteada se muestra el porcentaje de acetonitrilo durante la corrida. La detección se realizó a una longitud de onda de 230 nm.



2.5.2. Purificación de rHTsTx-K β y rTsHTx-K β

En el caso de rHTsTx-K β y rTsHTx-K β todas las pruebas mostraron que la proteína no se obtenía soluble. A manera de ejemplo, en la figura 2.6 se muestran algunas expresiones de rHTsTx-K β en diferentes cepas de *E. coli* monitoreadas en un gel de poliacrilamida al 15%. Se muestra el tiempo en que inicia la expresión (t0, antes de inducción) y la lisis celular separada en la fracción soluble (Tf) e insoluble (TfP). En todos lo casos se observa que la proteína se encuentra en la fracción insoluble, y hay una mayor cantidad de esta en la cepa C41.

Dados estos resultados, se optimizó un método para extraer las proteínas de la fracción insoluble. Entre las diferentes metodologías probadas, se encontró que hacer una lisis inicial con 1 molar de urea previa a la solubilización de la proteína de interés, provocaba la extracción de esta proteína sin tantos contaminantes. Después de los lavados de la fracción insoluble, se llevó a cabo una segunda sesión de choques sónicos en presencia de 1% de SDS lo que mejoraba mucho la solubilidad de la proteína. En la figura 2.6 de observa las diferentes etapas de la purificación de rHTsTx-K β , en esta fase se obtiene alrededor de 120 mg de proteína por litro de cultivo.

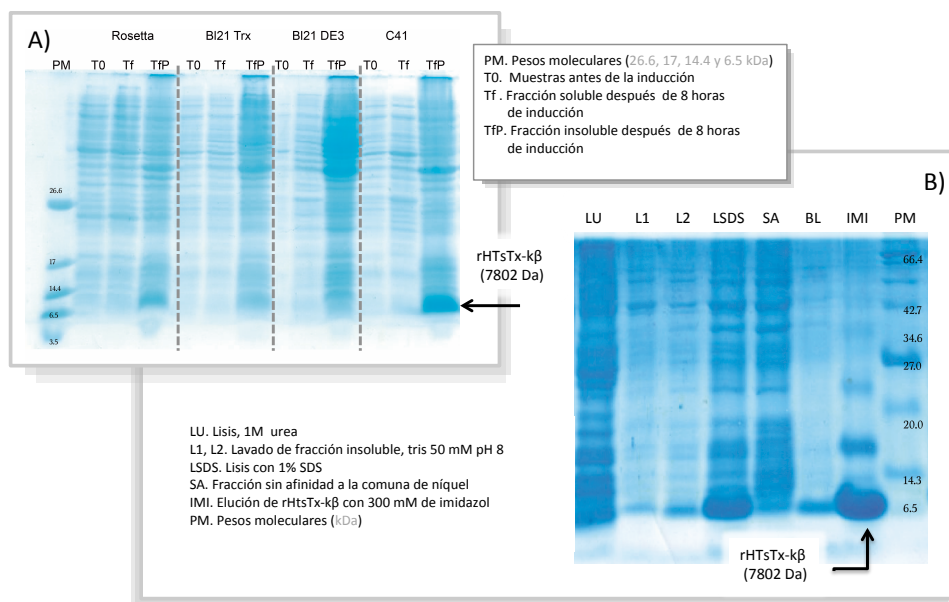


Figura 2.6. Expresión y purificación de rHTsTx-K β . Geles de poliacrilamida al 14 % donde se muestra la expresión de rHTsTx-K β en diferentes cepas de *E. coli*. A) y la purificación de la toxina de la fracción insoluble B).



A pesar de obtener una gran cantidad de proteína con la metodología utilizada, la cromatografía de fase reversa reveló que las proteínas no contenían un único plegamiento. Tanto rHTsTx-K β como rTsHTx-K β mostraban diferentes conformaciones obteniendo diferentes picos con diferentes tiempos de retención y los cuales contaban con la misma masa molecular (Figura 2.6). Por este motivo se llevó a cabo el replegamiento de las dos proteínas. La metodología descrita en la sección 2.4.3 fue la que mostró los mejores resultados. En la figura 2.7 se pueden comparar los cromatogramas de las toxinas antes y después del proceso de replegamiento. De tener varias conformaciones se obtiene una única especie con la masa esperada para la proteína con los tres enlaces disulfuro formados, 7,803.1 para rHTsTx-K β y 7,803.3 para rTsHTx-K β .

Para las dos proteínas, se realizó con éxito la eliminación de His-tag por medio de la reacción con bromuro de cianógeno. rhTsTx-K β y rTshTx-K β representan las toxinas sin la secuencia de histidinas.

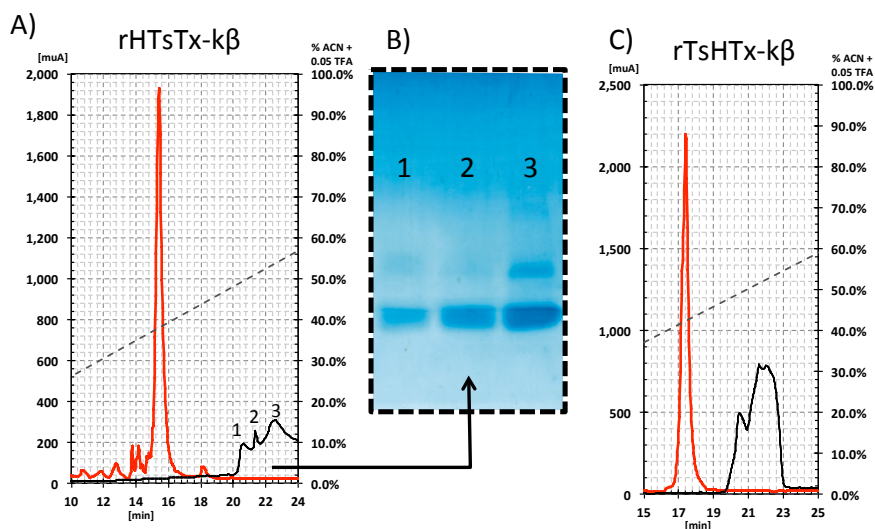


Figura 2.7. Cromatograma de rHTsTx-K β (A) y rTsHTx-K β (C). En ambos caso la línea negra representa el cromatograma de la proteína antes del replegamiento y la línea roja después de replegamiento. La línea punteada muestra la porcentaje de acetonitrilo durante la corrida. B) Gel de poliácridamida al 14 % donde se analizan los tres picos obtenidos de rHTsTx-K β antes del replegamiento



2.5.3. Expresión y purificación de rTs37Tx-K β

Dadas las características de TsTx-K β se diseñó una construcción para obtener también el dominio carboxilo terminal, al cual se le atribuye la actividad bloqueadora de canales. La expresión se realizó en la cepa C41 y la proteína se purificó de forma soluble. En la figura 2.8 se observa la fracción eluída de la columna de níquel con 300 mM de imidazol (carril 2); posteriormente, se observa el corte enzimático con la proteasa TEV (carril 3-5). Finalmente, en el gel se observa la fracción sin afinidad a la columna de níquel (carril 6) después del corte enzimático en donde se encuentra la rTs37Tx-K β . Esta fracción se concentró y se purificó por una columna de fase reversa C5 mostrando un pico mayoritario con un tiempo de retención de 14 min (Figura 24), esta muestra tiene la masa correspondiente de la toxina esperada con los tres enlaces disulfuro formados (4,265.5 Da).

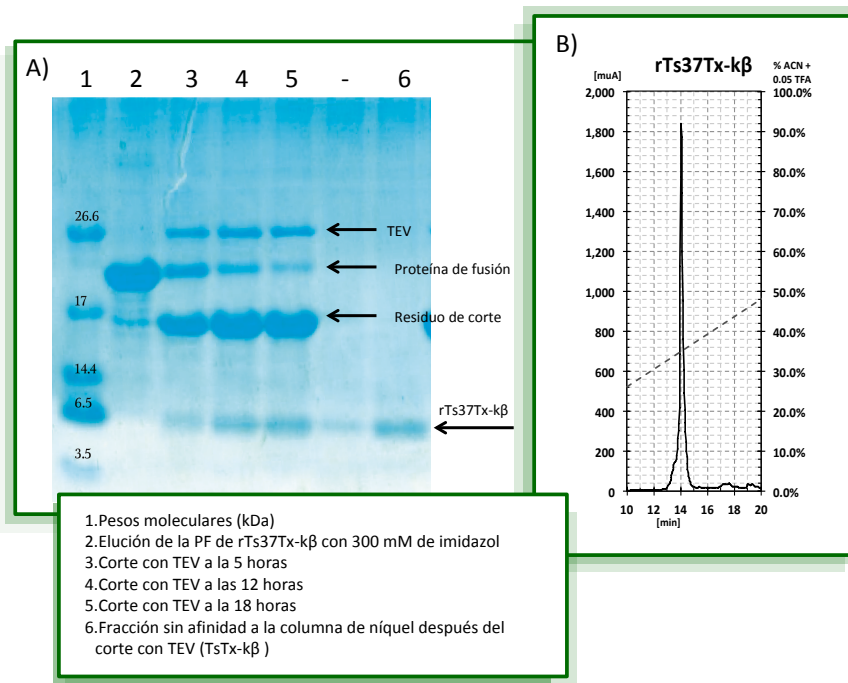


Figura 2.8. Producción y purificación de rTs37Tx-K β . A) Gel de poliacrilamida al 15% donde se sigue la purificación y la proteólisis enzimática de rTs37Tx-K β con TEV. B) Cromatograma de la purificación de rTs37Tx-K β . La línea punteada muestra la porcentaje de acetonitrilo durante la corrida.



2.5.4. Caracterización de las construcciones de TsTx-K β .

Los picos colectados en la purificación por cromatografía fueron colectados y liofilizados. Estas muestras se resuspendieron en una mezcla H₂O:D₂O 95:5 para su análisis por RMN. Todas las muestras resultaron muy solubles y se analizaron en concentraciones de 0.5 – 1 mM.

La primera muestra que se obtuvo y caracterizó fue rTsTx-K β . Su espectro monodimensional de protón se muestra en la figura 2.9 A). El espectro no es el convencional de una proteína plegada. Se observa una baja dispersión en las señales de los protones de amida en la región de 6 a 9 ppm así como la de los metilos (0.5 a 2 ppm). La baja dispersión resulta más evidente si la comparamos con el espectro de la proteína EhMybS3v3, la cual contiene el mismo número de aminoácidos y un motivo estructural bien definido (Figura 2.9, B). Al observar el espectro TOCSY de rTsTx-K β no se encuentran los 60 sistemas de espín esperados y se corrobora la baja dispersión.

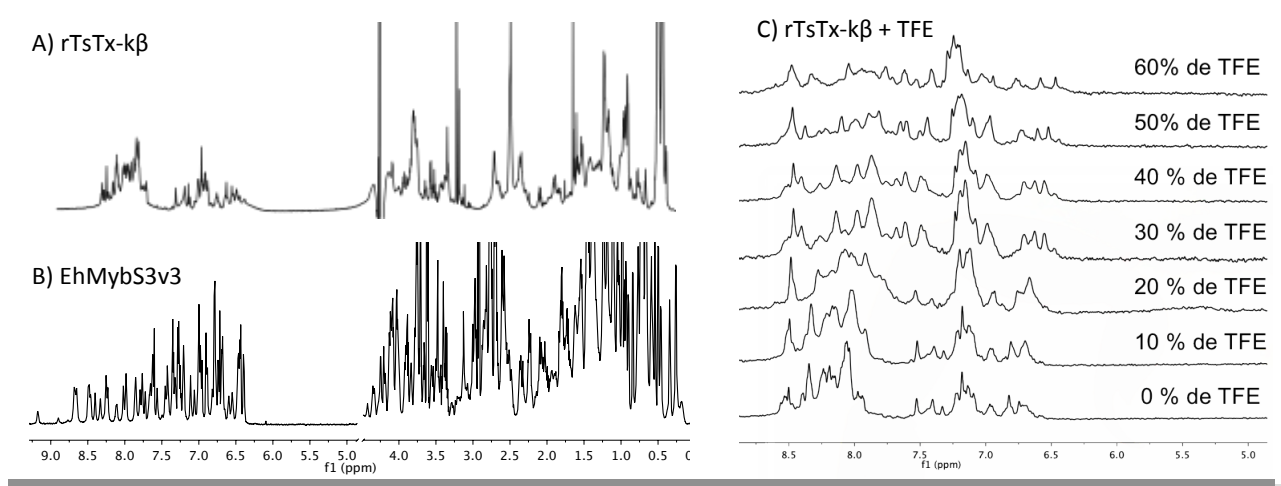


Figura 2.9. Espectros de RMN de rTsTx-K β y EhMybS3v3. A) y B) RMN monodimensional de protón, Muestras disueltas en 95% de H₂O y 5% D₂O y adquiridos en un equipo Bruker 700 MHz a 25 °C. C) Titulación de rTsTx-K β con TFE.

En el experimento NOESY se observan muy pocas señales lo que imposibilita la asignación de los sistemas de espín encontrados. En el caso de rhTsTx-K β y rTshTx-K β



los espectros obtenidos muestran exactamente los mismos resultados que los mostrados para rTsTx-K β .

Se probaron diferentes condiciones en la realización de los experimentos de RMN con el fin de dispersar las señales y generar espectros de mejor calidad. Lo primero que se probó fue adicionarle a la muestra diferentes concentraciones de trifluoroetanol (figura 2.9 C). Este compuesto se utiliza normalmente para simular un ambiente hidrofóbico y que las zonas desestructuradas que teóricamente interaccionan con la membrana adopten un plegamiento definido (Zhu et al. 2010). Sin embargo, aunque la dispersión mejoró siguen sin aparecer todas las señales y el espectro NOESY no cambió. También se probaron varias temperaturas (35, 25, 15, 5 y en el caso de la muestra con TFE -5°C), se modificó la fuerza iónica (0, 50 y 100 mM de NaCl), se ajustaron diferentes condiciones de pH (2, 5 y 7) y a pesar que en la mayoría de los casos se observó movimiento de las señales, las condiciones no fueron lo suficientemente favorable para realizar la asignación. Cabe señalar que aunque en el experimento TOCSY se observan y son identificables algunos sistemas de espín, sin el NOESY que lo complementa, solo son aminoácidos aislados que no podemos relacionar y por lo tanto asignar.

Recordemos que la proteína cuenta con dos dominios estructurales, un amino terminal sin aparente estructura y un carboxilo terminal donde se encuentran las cisteínas y presuntamente forman el motivo estructural α/β . Una posible explicación que surgió después de ver los experimentos de RMN fue que precisamente esta zona del amino terminal al ser desestructurada tiene una gran movilidad. Estos átomos al no tener ambientes químicos definidos generan señales anchas que dificultan la visualización de la parte carboxilo terminal que teóricamente es la zona estructurada de la proteína. Por este motivo se diseñó una cuarta construcción (rTs37Tx-K β) eliminando la parte desestructurada de la secuencia, conservando el dominio carboxilo terminal con el cual podríamos corroborar la existencia del motivo estructural α/β y no fue posible asignar en las construcciones anteriores.

La espectroscopia de esta nueva construcción (rTs37Tx-K β) mostró los mismos resultados observados por cuarta vez, una baja dispersión de las señales y una falta de



NOE's tanto secuenciales como a la larga distancia para llevar a cabo la asignación de la proteína.

En la figura 2.10 se pueden comparar los espectros de las cuatro especies de TsTx-K β obtenidas en este trabajo. Se observa que las tres construcciones del péptido completo es decir rTsTx-K β , rhTsTx-K β y rTshTx-K β muestra un mismo patrón de señales. Esto nos confirma que con las tres metodologías estamos llegando a la misma especie. La proteína obtenida de manera soluble es la misma que la que se obtiene después de un proceso de repliegamiento y la secuencia de histidinas no influye en la conformación final que ésta adopta.

En cuanto al péptido rTs37Tx-K β su espectro es muy similar al de las otras tres construcciones. Sorprendentemente, los sistemas TOCSY's observados en él son los mismos que en el caso de los péptidos completos; por lo tanto, el extremo amino terminal nunca se ha visto y los pocos sistemas observados son los del fragmento rico en cisteínas, el cual tiene una conformación particular que adoptan las cuatro construcciones.



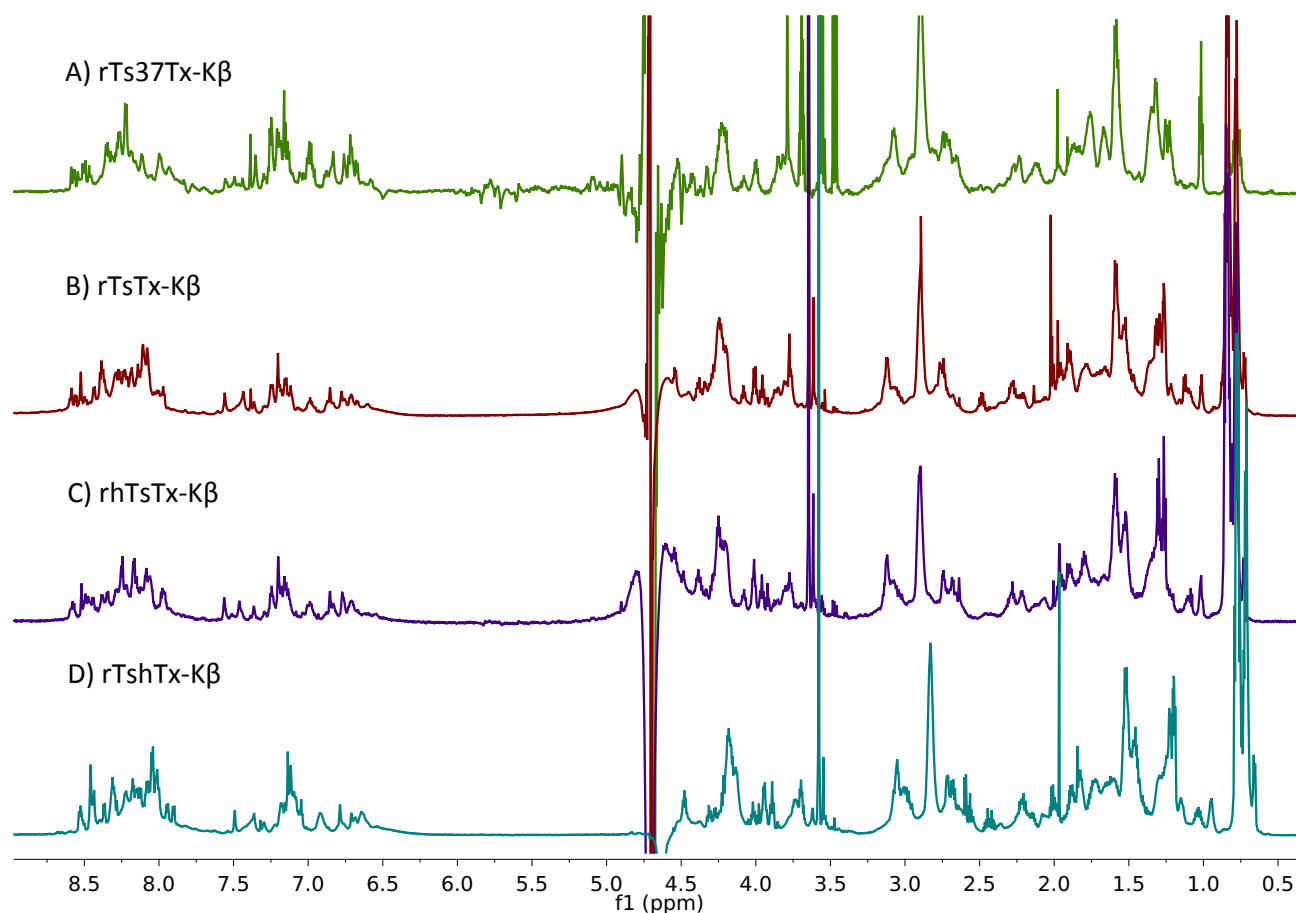


Figura 2.10. Comparación de los espectros monodimensionales de protón de rTs37Tx-K β , rTsTx-K β , rhTsTx-K β y rTshTx-K β . De 0.5 a 1 mg de toxina fue disuelta en 95% de H₂O y 5% D₂O. Bruker 700 MHz, 25° C.

Un hecho importante es la formación de los tres enlaces disulfuro, a pesar que la especie en la que convergen las cuatro construcciones tiene los enlaces formados, aparentemente no genera una conformación rígida y globular, de lo contrario, se deberían observar NOE's a corta y larga distancia. Por otro lado, la falta de señales en la parte del amino terminal sugiere una zona completamente desestructurada con mucha flexibilidad por la falta de NOE's y que, además, es muy susceptible al intercambio de los protones de amida por la falta de TOCSY's. Ambos procesos ensanchan las señales y no son observables.



A manera de ejemplo en la Figura 2.11 se muestra el espectro de dicroísmo circular de la proteína rTs37Tx-K β , en esta figura se observa que la proteína no tienen la apariencia típica de una proteína con un plegamiento alfa/beta. Debido a que en este tipo de espectros observamos la suma de las contribuciones de estructura secundaria que tiene la proteína, es difícil determinar con exactitud su contenido. Utilizamos el servidor DichroWeb para analizar los datos experimentales, el programa sugiere que rTs37Tx-K β tiene el porcentaje de estructura secundaria observado en la Figura 2.11 C), mostrando que la proteína presenta un alto porcentaje de zonas desestructuradas (43%), corroborando los datos de RMN.

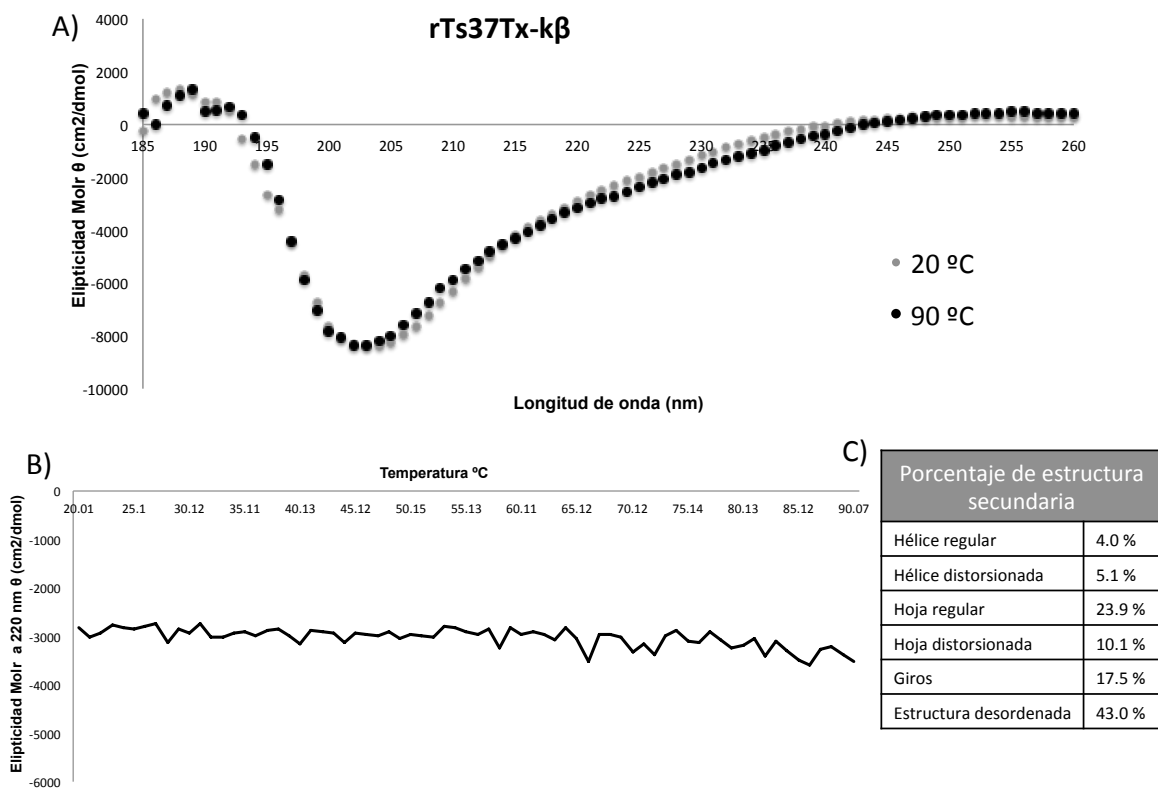


Figura 2.11. Dicroísmo circular de rTs37Tx-K β . A) Espectro de dicroísmo circular de rTs37Tx-K β a 20 °C (gris) y 90 °C (negro). B) Curva de desplegamiento de rTs37Tx-K β mediante calentamiento (de 20 a 90 °C) registrando su señal a 220 nm. C) Porcentajes de estructura secundaria empleando el método SELCON3 en el servidor DichroWeb.



También se realizó una curva de calentamiento para evaluar como se comportaba rTs37Tx-K β a altas temperaturas, la rampa de calentamiento alcanzó una temperatura de 90 °C registrando a una longitud de onda de 220 nm. En la Figura 2.11 B) se observa esta rampa de calentamiento sin un cambio de absorción importante. Al realizar un barrido completo a la temperatura máxima, se observa un espectro casi idéntico al tomado a la temperatura inicial (20 °C, Figura 2.11 A), indicando que este calentamiento no generó ningún cambio conformacional, debido posiblemente a la formación de los enlaces disulfuro.

Al probar la actividad de los péptidos obtenidos no mostraron una actividad relevante (inhibición de 6% del canal a 3 μ M), confirmando así que no estamos obteniendo la conformación nativa de la toxina TsTx-K β .

Los resultados sugieren que a pesar de tener una sola conformación y estar totalmente oxidada, la proteína tiene un plegamiento con cierta movilidad que no permite llevar a cabo una asignación y mucho menos la obtención de su estructura tridimensional.

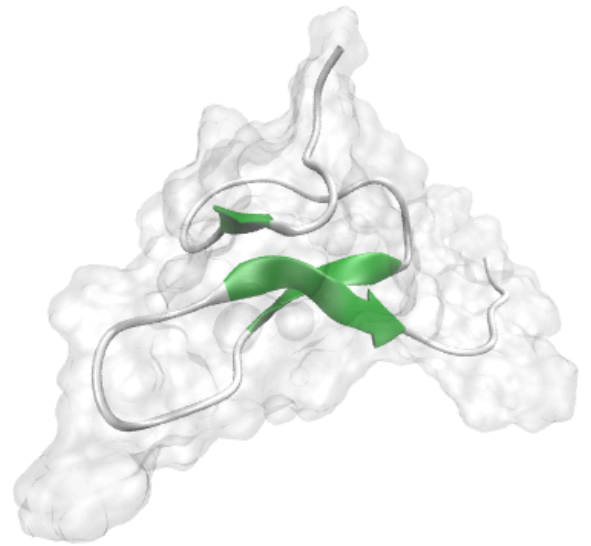


Se logró la obtención del péptido con la secuencia de TsTx-K β de forma soluble. Este proceso se realizó por tres metodologías diferentes, las cuales llegaron a la misma especie conformacional visto por resonancia magnética nuclear. La especie obtenida es soluble y estable; sin embargo, el péptido no posee aparentemente la conformación nativa por su falta de actividad frente al canal de potasio Kv 1.3, en el cual ya se ha probado su actividad en estudios anteriores con la toxina nativa. Los resultados obtenidos sugieren que las condiciones necesarias para alcanzar el plegamiento nativo de TsTx-K β no dependen solamente de la secuencia de aminoácidos. Claramente en el proceso de replegamiento así como en la bacteria, se obtuvo una especie de baja energía, la cual es incapaz de pasar a la conformación nativa. Probablemente, para que esta toxina alcance su plegamiento nativo al sintetizarse en su entorno natural debe existir una proteína chaperona o la existencia de condiciones específicas en su entorno que favorezcan dicho plegamiento.



- Para continuar con el estudio de TsTx-K β probablemente debemos concentrarnos más en imitar las condiciones de sistemas eucariontes en las que esta se produce. Hasta ahora hemos probado el plegamiento *in situ* y una producción en un sistema bacteriano; claramente eso se aleja mucho del sistema eucarionte en el que se produce naturalmente. Otra opción es probar la biosíntesis de esta proteína en alguna levadura o células de insecto que imiten mejor las condiciones de síntesis nativas.
- Continuar con las pruebas de actividad en diferentes canales de potasio y otros sistemas biológicos como parásitos y/o líneas cancerosas.
- Desde el punto de vista estructural, se pueden continuar con los estudios de RMN ya que es posible el análisis de proteínas desestructuradas al enriquecer con isotopos como ^{15}N y ^{13}C . La asignación al dispersarse en una tercera dimensión resulta mas fácil e inclusive se pueden realizar estudios de movilidad de los distintos fragmentos de la proteína.

3. MAGI 3: UNA PROTEÍNA RICA EN PUENTES DISULFURO



3.1 Macrothele Gigas y Magi3

Macrothele gigas es una especie de araña de la familia *macrothelidae* originaria del sureste de Asia. En el 2003 su veneno fue estudiado por el Dr. Corzo y colaboradores del cual purificaron 6 toxinas con actividad sobre canales de sodio.



Macrothele gigas

Toxina	secuencia	#cisteínas
Magi1	CMGYDIHC T DRLP CC FGLE C VKTSGYWYKTTY CRR KS	(6)
Magi2	CMGYDIE C NENLP CC KHRKLE C VETSGYWYKRKY CR PIK	(6)
Magi3	GG C IKWNH S C Q T T TL K CC G K C V V C Y CHTPWGT N C R CD R TRL F CT E D	(10)
Magi4	CGSKRAW C KEK KD CC C GY N CVYAWYNQ Q SS C ERKWKYL F T G EC	(8)
Magi5	CKLTFWK C KN K KE CC GW N ACAL G IC M PR	(6)
Magi6	K C V D GS C DPYSS D AP R CC G S Q IC Q C I FF V PC Y CK Y R	(8)

Figura 3.1. Secuencias primarias de Magi1-6

En la figura 3.1 se muestran las secuencias primarias de las 6 toxinas las cuales se denominaron Magi1-6 por el nombre de la araña *Macrothele gigas* y el número correspondiente al orden en el que fueron purificadas, en la figura 3.1 podemos observar que Magi3 además de ser la toxina más grande (46 aminoácidos), también es la que presenta el mayor número de cisteínas en sus secuencia de aminoácidos, generando la formación de 5 enlaces disulfuro. Aunque en la naturaleza ya se han observado algunas toxinas con un gran número de enlaces disulfuro (Klint et al., 2012) hay pocos estudios y no se conocen sus estructuras así como las conectividades de sus cisteínas al formar los enlaces disulfuro. Es por esto que nos enfocamos en el estudio de Magi3. Una toxina que hasta la fecha resulta intrigante, su secuencia primaria no tiene grandes similitudes con sus hermanas Magis u otros péptidos con actividad sobre canales de sodio. Su estructura no se conoce y no se tienen indicios de que plegamiento puede adoptar, la poca homología que presenta esta proteína (menor al 45%) se relaciona con toxinas afines a canales de calcio por lo que se le atribuye una actividad promiscua (Corzo et al. 2003).

Antes de iniciar esta parte del proyecto ya se contaba con estudios de expresión de esta proteína (Titau, 2012). En estos estudios se habían determinado la complejidad del sistema para obtener la toxina de manera recombinante, finalmente se obtuvo una



metodología con la cual fue posible obtener la muestra. Sin embargo, la cantidad de toxina obtenida no era suficiente para estudios estructurales y no se contaba con resultados que demostraran que el plegamiento obtenido era el nativo. En este proyecto se retomó dicha investigación.



3.2.1. Objetivo general

Retomar la investigación referente a la toxina Magi3, obtenerla de forma recombinante, obtener su estructura tridimensional en disolución por resonancia magnética nuclear y determinar su actividad en un sistema biológico.

3.2.2. Objetivos particulares.

- Expresar de forma heteróloga en *E. coli* la proteína de fusión que contenga a la toxina Magi3.
- Purificar la proteína de fusión por métodos cromatográficos y obtener la toxina Magi3 aplicando metodologías de replegamiento.
- Caracterización de la proteína de interés.
- Obtener e interpretar los experimentos de RMN (COSY, TOCSY y NOESY)
- Calcular la estructura tridimensional considerando las restricciones obtenidas por los experimentos de RMN.
- Realizar pruebas de actividad.



3.3.1. Expresión de la proteína de fusión.

El diseño del gen codificante de la Magi3 se realizó con base en la secuencia de aminoácidos reportados previamente (UniProt, P83559). Se sintetizaron dos oligonucleótidos con secuencias complementarias y se insertaron al plásmido pET32(a) entre los sitios de restricción Bam H1 y KpnI. Se escogió este plásmido debido a que genera una proteína de fusión como se muestra en la figura 3.2, la cual está constituida por un dominio de tiorredoxina que nos ayudará a fomentar la formación de enlaces disulfuro y solubilizar la proteína de fusión, un sitio de reconocimiento a la proteasa trombina para liberar la proteína de fusión de la toxina de interés y finalmente la Magi3; además de un fragmento SAAA entre el sitio de corte y la toxina. Estudios anteriores mostraron que este fragmento mejora el rendimiento de la proteólisis.

Vector pET-32a



Figura 3.2. Composición de la proteína de fusión que contiene a la Magi3.

La proteína de fusión producida por el plásmido de trabajo fue sobreexpresada en células de *E. coli* cepa Rosseta-Gami donde previamente ya se conocía que se expresaba de manera soluble. La transformación de éstas células con el plásmido se realizó con la metodología mencionada anteriormente en la sección 2.3.1.

El crecimiento de las células se realizó a partir de una colonia aislada obtenida de la transformación. Se seleccionó una colonia con una punta estéril y se resuspendió en 100 mL de medio LB con el antibiótico correspondiente dejándolo incubar 16 horas a 37 °C. 50 mL de este inóculo se transfirió a 1,500 mL de medio LB con antibiótico y se incubó a 37 °C monitoreando su crecimiento. Cuando el cultivo alcanzó una densidad óptica (DO_{600})



entre 0.6 y 0.8 unidades de absorbancia, la expresión de la proteína de fusión fue inducida con IPTG a una concentración de 0.5 mM. La temperatura se llevó a 30 °C y se dejó por 8 horas. Transcurrido el tiempo, se colocó el medio en frascos los cuales fueron centrifugados a 6,500 rpm por 15 min y se resuspendieron en 50 mL de solución amortiguadora de lisis (50 mM de tris pH 8.0, 300 mM NaCl).

3.3.2. Purificación de la proteína de fisión.

Después de la resuspensión, las células se sometieron a un programa de choques sónicos por periodos de 20 s a 45 W con descansos de 40 s por un tiempo efectivo de 10 min dejándolo en un baño de hielo durante todo el proceso. El lisado celular se centrifugó por 45 min a 16,500 rpm conservando el sobrenadante. El siguiente paso fue separar la proteína de fusión del resto de material soluble que se encuentra en el lisado celular, la purificación fue realizada por CAMQ en una columna HipTrap®. La muestra se filtró con una membrana de 0.45 µm y se pasó por la columna a un flujo de 5 mL/min cargada previamente con Ni²⁺. La PF se quedó retenida en la columna debido a una secuencia de 6 histidinas las cuales se coordinan con el níquel. Ya con la PF retenida, se realizó un lavado con cinco volúmenes de columna (25 mL) con solución amortiguadora (50 mM de tris pH 8.0 , 300 mM NaCl) y consecutivamente con disoluciones que contenían una mezcla de solución amortiguadora de lisis más imidazol 5, 25, 50,100, 200 y 300 mM para eliminar otras proteínas y, finalmente, eluir la PF.

3.3.3. Replegamiento de la proteína de fusión.

El replegamiento de la proteína de fusión se llevó a cabo desplegando y reduciendo la PF previamente purificada. Se cambió la solución amortiguadora (tris 50 mM pH 9.0.) mediante diálisis en una membrana de 10 kDa, con dos recambios de 500 mL de una hora aproximadamente cada uno. A esta solución se le agregó como agente reductor DTT a una concentración 10 mM y cloruro de guanidinio 5 M como desnaturizante. La reacción se realizó durante cuatro horas a temperatura ambiente. Con una membrana de diálisis, (10 kDa) se eliminó el DTT y se redujo la concentración de cloruro de guanidinio a 1 M. La muestra reducida se adicionó de manera muy lenta y con agitación baja a la solución amortiguadora de replegamiento. El par redox fue glutatión oxidado y reducido en



relación 1:5. La concentración final de la PF fue de 60-80 µg/mL y se dejó por 4 días a temperatura ambiente. La PF fue concentrada por ultrafiltración en un sistema *Amicon* con una membrana de 10 kDa a 4 °C y posteriormente se volvió a pasar por la columna de níquel de la misma forma que se mencionó anteriormente

3.3.4. Corte enzimático con trombina.

Una vez replegada la PF se llevó a cabo el corte con trombina de la siguiente manera. La reacción se realizó por 3 horas en solución amortiguadora (50 mM de Tris/HCl pH 9.0 y 2 mM de CaCl₂) temperatura ambiente y una concentración de proteínas de 3 mg/mL. Para 1 mL de muestra se utilizó 50 µL de suspensión de agarosa en la cual se encuentra fija la trombina (SIGMA®). Transcurrido el tiempo se separó la trombina por centrifugación y se llevó a una concentración de 300 mM de NaCl. Inmediatamente después, se pasó por la columna HipTrap® con un flujo de 1 mL/min y se lavó con 50 mL de solución amortiguadora de lisis (tris 50 mM pH 9.0 y 300 mM de NaCl). Todo se concentró por ultrafiltración con una membrana de 3 kDa a 4 °C.

Todas las muestras se analizaron mediante electroforesis desnaturizante en geles de poliacrilamida (SDS-PAGE) al 14%, los cuales se corrieron en una cámara de electroforesis con solución amortiguadora de tanque a un voltaje inicial de 80 V por 18 min y después se aumentó el voltaje a 120 V por 60-100 min. Los geles fueron teñidos con “silver blue”.

Finalmente, la toxina eluída se sometió a un último paso de purificación por cromatografía de líquidos de alta eficiencia, utilizando como fase estacionaria una columna de fase reversa C5 (Phenomenex® C5) y como fase móvil un gradiente lineal del 15 al 50% de agua-acetonitrilo al 0.05% de TFA; todo esto acoplado a un sistema de detección UV registrando a 230 y 280nm.

3.3.5. Obtención y caracterización de la Magi3.

Se determinó la masa molecular de la muestra obtenida por espectroscopia de masas MALDI-TOF como se describe en la sección 2.3.5.



La Magi 3 se obtuvo con siete aminoácidos extra en el N-terminal (SAAAAGM) como consecuencia de haber elegido la trombina para la realización del corte enzimático y la modificación para mejorarlo; este fragmento se removió con la reacción de bromuro de cianógeno como se describió en la sección 2.3.4.

3.3.6. Proteólisis enzimática para la determinación de enlaces disulfuro

Con el objetivo de tener más información acerca de la conectividad de enlaces disulfuro que contenía la proteína, se realizó una digestión enzimática con la proteasa tripsina. Se colocó 5 µg de Magi3 pura y 250 ng de la enzima en 15 µL de solución amortiguadora (tris 50 mM pH 8.0.0) la mezcla de reacción se colocó a 37 °C por 16 horas y fue analizada directamente por MALDI-TOF en las condiciones previamente descritas (sección 2.3.5).

3.3.7. Obtención y análisis de los experimentos de RMN-¹H.

Los espectros de RMN de Magi3 fueron obtenidos y procesados como se describe en la sección 2.3.5.

Para la asignación de los espectros bidimensionales, se utilizó el programa CARRA 1.5 y la herramienta XEASY (Keller, 2004). Con las restricciones obtenidas del espectro NOESY, se realizaron cálculos de estructura en CYANA 2.1 (López-Méndez y Güntert, 2006). Con las 20 estructuras de menor energía, se llevó a cabo un afinamiento de la estructura por dinámica molecular con el programa AMBER (Salomon-Ferrer et al. 20013). La dinámica se realizó para la 20 estructuras de forma individual en vacío y con disolvente explícito, utilizando el campo de fuerza ff99SB.

3.3.8. Pruebas de actividad de Magi3 en grillos y ratones.

Grillos (*Gryllus bimaculatus*) fueron inyectados intratorácicamente entre el segundo y tercer par de patas con diferentes dosis de la toxina disuelta en agua destilada. Un control fue inyectado solo con agua destilada. Los grillos fueron observados por 2 horas. Se



determinó “parálisis” cuando los grillos perdían toda la movilidad durante las 2 horas y regresaba su movilidad. El efecto letal se observó cuando los grillos morían después de la inyección.

Ratones (CD-1 20 g) machos fueron utilizados para las pruebas siguiendo el protocolo de ensayos de la actividad de péptidos *in vivo* de las normas del Instituto de Biotecnología de la UNAM. La prueba se realizó vía intracraneal inyectando 10 μ L con 1 y 5 μ g de Magi3 disuelta en una disolución 0.25 mg/mL de BSA. Un blanco negativo se utilizó con solo BSA y uno positivo con el péptido de escorpión CsslI purificado y prestado por el grupo del Dr. G. Corzo (Hernández-Salgado et al., 2009). Los ratones fueron observados por 24 horas para determinar la toxicidad del péptido.



3.4.1. Expresión de la proteína de fusión con Magi3.

Se observó una buena expresión de la proteína de fusión en la cepa de *E. coli* Rosseta-gami. El paquete celular obtenido después del tiempo de inducción es de aproximadamente 2.8 g por litro de cultivo. La figura 3.3 A) muestra la elución (300 mM imidazol) en la columna de afinidad, la proteína purificada presenta una masa molecular mayor a los 17 kDa pero por debajo de los 26 kDa (masa esperada 22,463 Da).

3.4.2. Replegamiento de la proteína de fusión y corte enzimático con trombina.

En términos de recuperación de muestra, el rendimiento del replegamiento fue de un 50%. La proteólisis enzimática se llevó a cabo dejando la reacción por tres horas observando una eficiencia del 50% aproximadamente (Figura 3.3, B). En este tiempo, aunque no se cortaba la proteína en su totalidad, tampoco se observaba precipitado y la trombina se podía lavar fácilmente para su reutilización.

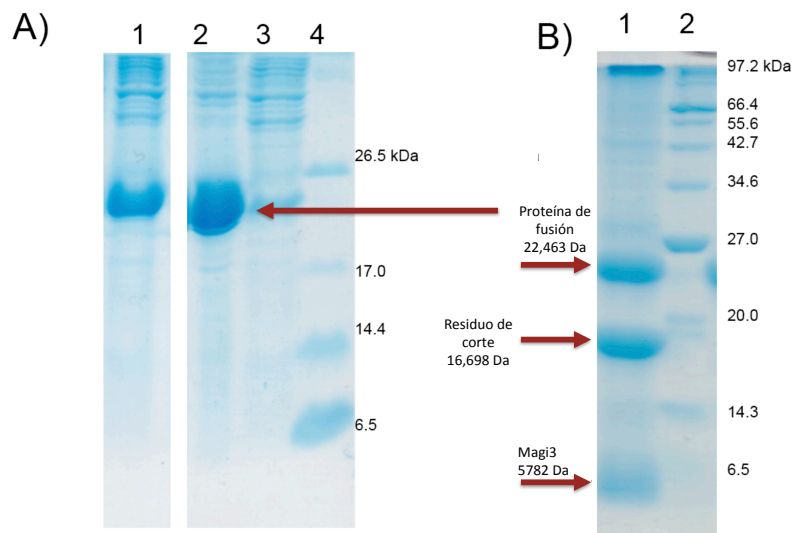


Figura 3.3. Purificación y corte enzimático de la proteína de fusión de Magi3. Gel de poliacrilamida al 15% donde se sigue la purificación y la proteólisis enzimática de Pf-Magi3 con trombina. A) 1. Lisis bacteriana, 2. Elución de Pf-Magi3 con 300 mM de imidazol. 3. Lavado de pf-Magi3 con 20 mM de imidazol, 4. Marcadores de masas moleculares. B) 1. Corte enzimático de Pf-Magi3 con trombina después de 3 horas. 2. Marcadores de masas moleculares.



3.4.3. Purificación y caracterización de la Magi3.

A continuación se muestra el cromatograma de la purificación de la fracción sin afinidad a la columna Hip-Trap® (Figura 9). En él, se puede observar un pico con un tiempo de retención en 16.8 min. Al analizar el espectro de masas de la muestra, se determina una masa molecular de 5,782.21 Da. Esta masa es la correspondiente a la Magi3 más los 7 aminoácidos GSAAAAM, producto de la reacción de corte. Esta masa molecular también nos indica que la proteína se encuentra en un estado oxidado formando 5 puentes disulfuro. En esta figura también podemos observar claramente que no existe una gran cantidad de especies en el cromatograma, lo esperado al replegar una proteína con tantas cisteínas. Sorprendentemente en este replegamiento, con las condiciones propuestas sólo estamos obteniendo una especie y está completamente oxidada. El rendimiento final fue de 0.43 mg de toxina por L de medio.

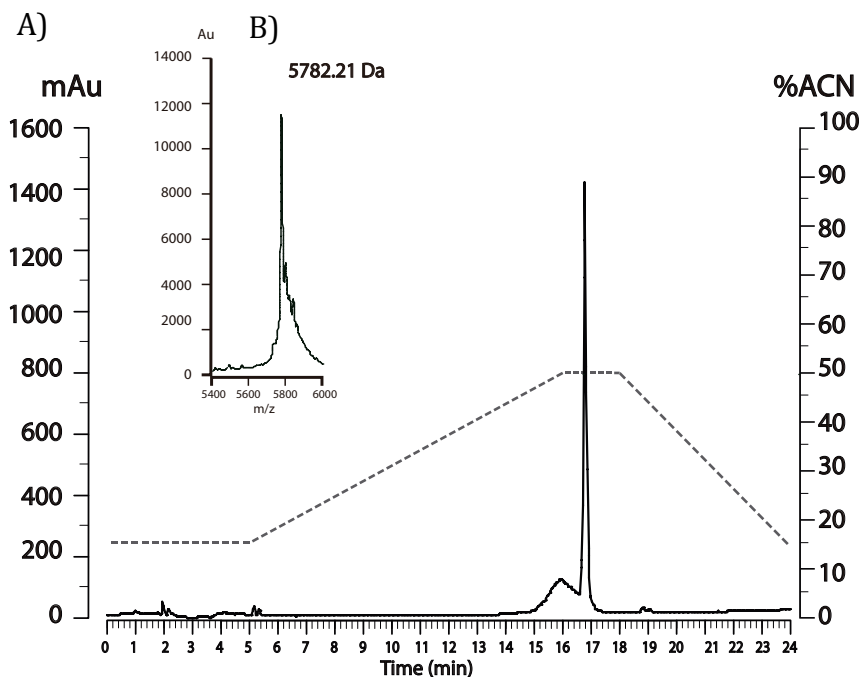


Figura 3.4. Purificación por cromatografía fase reversa de GSAAAAM-Magi3 (A), columna jupiter C4. Masa molecular de GSAAAAM-Magi3 obtenido por espectrometría de masas MALDI-TOF (B).



El corte con bromuro de cianógeno se efectuó casi al 100 % y al purificarlo por el método cromatográfico establecido generó un cromatograma similar al mostrado en la figura 3.4. La masa determinada para esta muestra fue de 5,223.77 Da (Figura 3.5 inciso A), la cual, coincide con la reportada en la literatura de la Magi3 en estado oxidado.

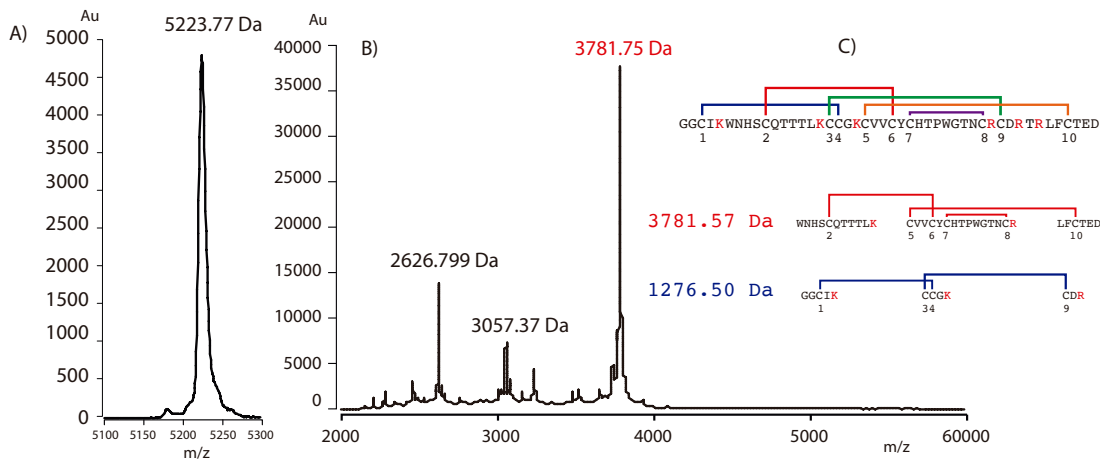


Figura 3.5. Espectro de masas MALDI-TOF de Magi3 despues del corte con bromuro de cianogeno A). Proteólisis enzimática para determinar enlaces disulfuro; B) y C) Fragmentos de Magi3 que se deducen por los datos encontrados en espectrometria de masas.

Después de la proteólisis enzimática de Magi3 con tripsina, se logró la identificación de un fragmento con masa molecular de 3,781.75 Da. En la tabla 3.1 se pueden ver todos los fragmentos teóricos que se pueden obtener de la reacción, se consideran todas las combinaciones de fragmentos cortados en los sitios carboxilo terminal de lisinas y argininas. Es posible que la enzima no necesariamente cortara en todos los sitios. El fragmento encontrado corresponde a la suma de los fragmentos 21, 18 y 23 de la tabla mencionada, y esto nos indica que existen enlaces disulfuro entre las cisteínas 2, 5, 6, 7, 8 y 10. Analizando estos fragmentos nos damos cuenta que el fragmento 18 contiene las cisteínas 5, 6, 7, y 8, por lo que dos de estas cuatro cisteínas forman enlaces con la 2 y 10, y existe otro enlace entre ellas para completar los 3 puentes. Por otro lado, de las cisteínas faltantes, el fragmento 26 contiene dos cisteínas consecutivas por lo tanto estas formarían enlaces con los fragmentos 25 y 27.



Tabla 3.1. Fragmentos posibles de la proteólisis de Magi3 con Tripsina.

Masa M.	Secuencia	Fragmento
5222.54	GGCIKWNHSCQTTTLKCCGKCVVYCHTPWGTNCRCDRTRLFCTED	1
4772.01	WNHSCQTTTLKCCGKCVVYCHTPWGTNCRCDRTRLFCTED	2
4521.96	GGCIKWNHSCQTTTLKCCGKCVVYCHTPWGTNCRCDRTR	3
4264.81	GGCIKWNHSCQTTTLKCCGKCVVYCHTPWGTNCRCDR	4
4063.73	WNHSCQTTTLKCCGKCVVYCHTPWGTNCRCDRTR	5
3890.68	GGCIKWNHSCQTTTLKCCGKCVVYCHTPWGTNCR	6
3806.58	WNHSCQTTTLKCCGKCVVYCHTPWGTNCRCDR	7
3472.41	CCGKCVVYCHTPWGTNCRCDRTRLFCTED	8
3432.44	WNHSCQTTTLKCCGKCVVYCHTPWGTNCR	9
3081.27	CVVYCHTPWGTNCRCDRTRLFCTED	10
2764.13	CCGKCVVYCHTPWGTNCRCDRTR	11
2506.98	CCGKCVVYCHTPWGTNCRCDR	12
2372.10	CVVYCHTPWGTNCRCDRTR	13
2167.99	GGCIKWNHSCQTTTLKCCGK	14
2132.84	CCGKCVVYCHTPWGTNCR	15
2115.84	CVVYCHTPWGTNCRCDR	16
1776.85	GGCIKWNHSCQTTTLK	17
1741.71	CVVYCHTPWGTNCR	18
1709.76	WNHSCQTTTLKCCGK	19
1358.58	CDRTRLFCTED	20
1318.62	WNHSCQTTTLK	21
984.44	TRLFCTED	22
727.30	LFCTED	23
650.30	CDRTR	24
477.25	GGCIK	25
410.15	CCGK	26
393.15	CDR	27
276.1666	TR	28

3.4.4. Caracterización estructural de Magi3

En la figura 3.6 se muestran los espectros de protón unidimensional de la Magi3, en éste se observa que la muestra analizada presenta un plegamiento definido debido a la dispersión de señales entre 6 y 9 ppm correspondiente de los protones de amida. Esto, es debido a que cuando una proteína se pliega muchos de los protones de amida normalmente se encuentran formando puentes de hidrógeno, lo que ocasiona que éstos tengan un ambiente químico definido y no aparezcan en un desplazamiento típico de protón de amida. Dos señales características son las observadas en 10.13 y 10.07 ppm; éstas pertenecen a la señal del HN del grupo indol perteneciente al triptófano. Esto último nos confirma la presencia de dos triptófanos, lo cual es congruente con la secuencia de la Magi3. Con los espectros bidimensionales TOCSY y NOESY se logró la asignación de un total de 1,161 señales pertenecientes a 41 de los 46 sistemas de espín, NOEs secuenciales y NOEs a larga distancia.



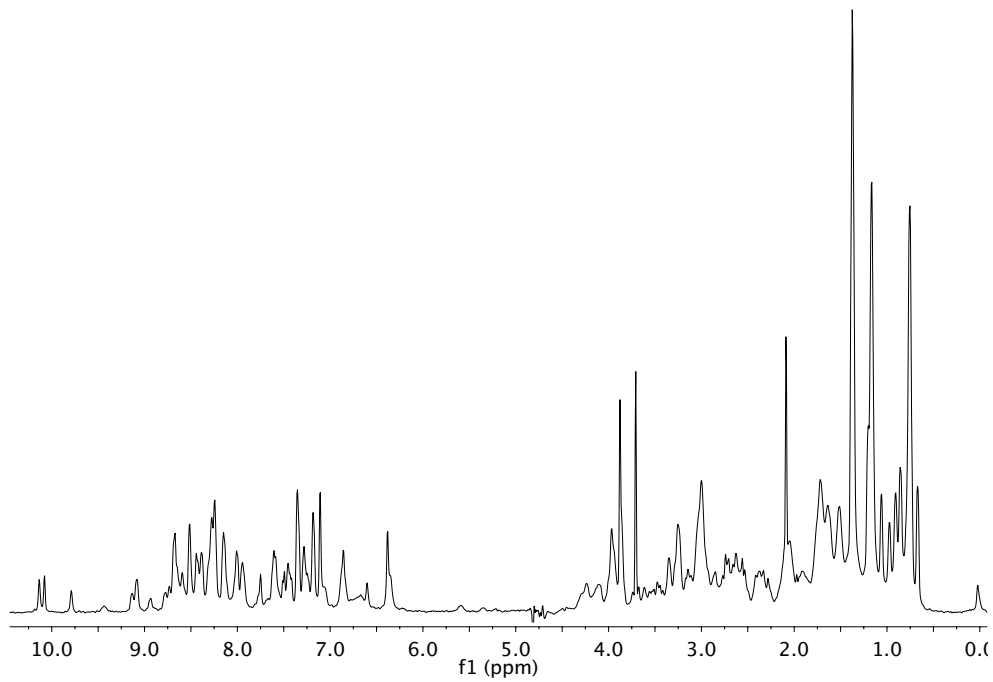


Figura 3.6. Espectro monodimensional de hidrógeno de Magi3. 1.3 mg de la toxina fue disuelta en 95% de H₂O y 5% D₂O. Bruker 700 MHz, 25° C.

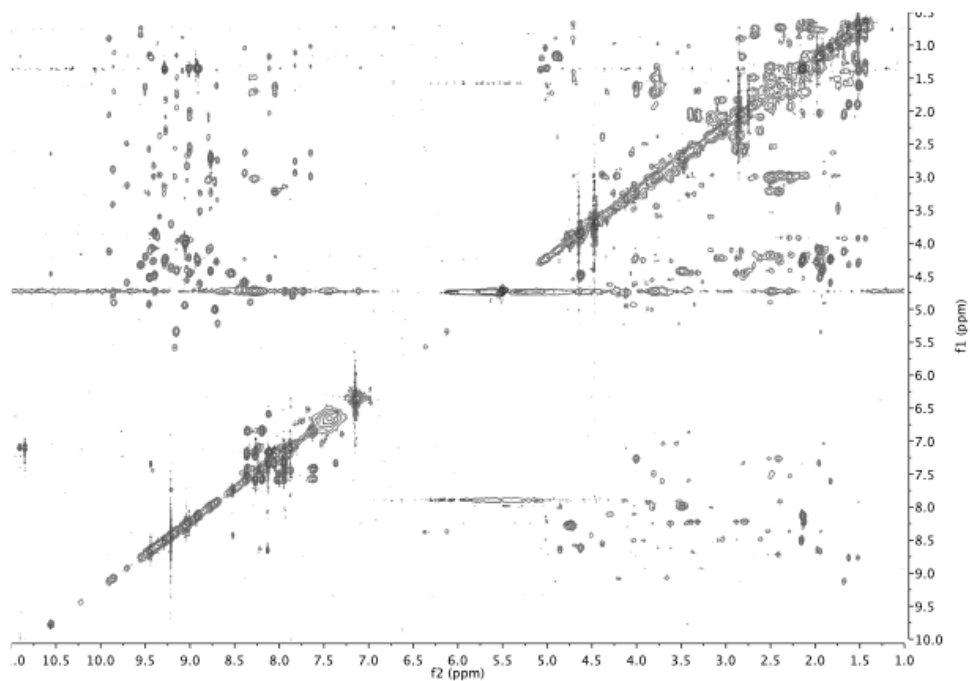


Figura 3.7. Espectro bidimensional TOCSY de Magi3. 1.3 mg de la toxina fue disuelta en 95% de H₂O y 5% D₂O. Bruker 700 MHz, tiempo de mezcla 80 ms, 25 °C.



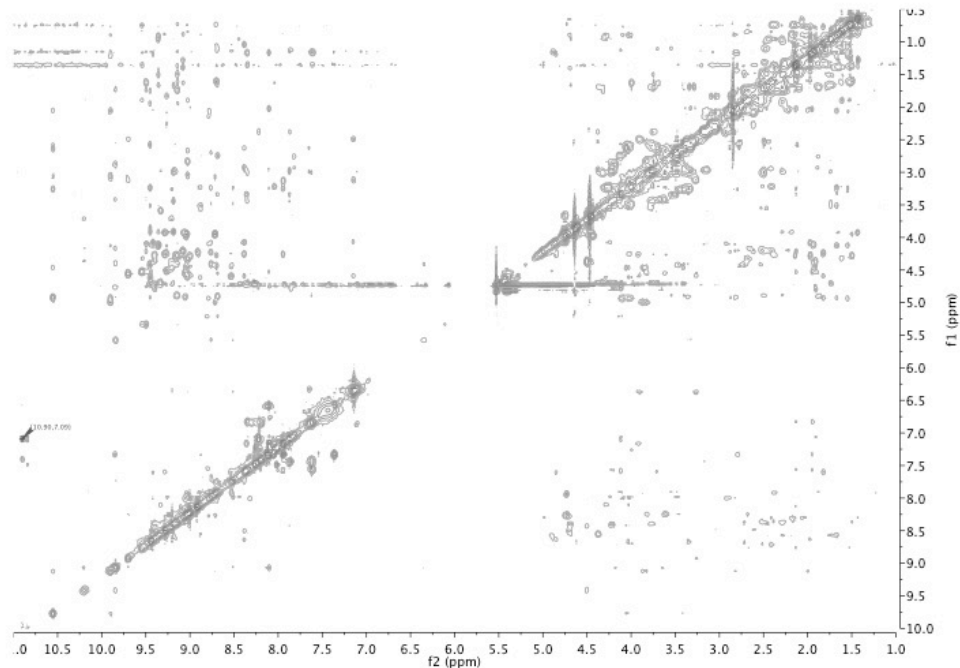


Figura 3.8. Espectro bidimensional NOESY de Magi3. 1.3 mg de la toxina fue disuelta en 95% de H₂O y 5% D₂O. Bruker 700 MHz, tiempo de mezclado 300 ms, 25°C.

Con el programa Cyana se llevaron a cabo tres cálculos estructurales. En el primero (cálculo 1) sólo se utilizaron los NOE's como restricciones de distancia para obtener la estructura tridimensional. En el segundo (cálculo 2), se utilizaron además de los NOE's los 5 enlaces disulfuro que reportaba uniprot (C3–C18, C10–C24, C17–C36, C21–C34 y C26–C43), los cuales eran propuestos por homología con otras proteínas. Haciendo un análisis visual con la estructura obtenida del primer cálculo, resultaba clara la cercanía de las cisteínas las cuales formaban los enlaces disulfuro. Con estas observaciones se generó el tercer cálculo (cálculo 3) asignando los 5 enlaces (C3–C18, C10–C24, C17–C36, C21–C43 y C26–C34) que se observaban, tanto por la cercanía de las cisteínas en la estructura sin puentes, como por algunos NOE's particulares. Entre estos NOE's destacan: el protón alfa de la cisteínas 26 con el protón alfa de la cisteína 34, el protón beta de la cisteínas 3 con el protón alfa de las cisteína 18, el protón beta de la cisteína 24 con el de la 10 y el protón beta de la cisteína 21 con el protón alfa de la arginina 40 entre otros. Estos datos son congruentes con la información obtenida por el análisis de la proteólisis enzimática y espectrometría de masas (Figura 3.5).



Tabla 3.2. Estadísticas de RMN de Magi3.

	Calc 1	Calc 2	Calc 3
Estadísticas de NMR			
Total de NOE's	421	401	419
Corta distancia $ i-j \leq 1$	306	306	306
Mediana distancia $1 < i-j < 5$	33	33	29
Larga distancia $ i-j \geq 5$	82	62	84
RMSD	1.47	2.6	1.2
Ramachandran			
Regiones más favorecidas	58.5%	49.2%	55.6%
Regiones permitidas	41.0%	49.5%	44.4%
Regiones poco permitidas	0.0%	1.0%	0.0%
Regiones prohibidas	0.2%	0.2%	0.0%

Otro dato importante es observar como al utilizar los enlaces propuestos por *uniprot* para el cálculo 1, el RMSD es mayor (2.6 Å) que el obtenido en la estructura sin puentes disulfuro (1.47 Å) y cuando utilizamos los puentes que observamos por la cercanía de las cisteínas (calculo 3) el RMSD baja aún más (1.2 Å). Este dato nos indica que los enlaces que se proponen en *uniprot* no corresponden con las interacciones NOEs que se observan en la proteína y esto hace que las estructuras generadas están mas forzadas a cumplir, tanto con las restricciones NOE's que se observan en RMN, como con los enlaces que nosotros le imponemos, por lo tanto las estructuras no convergen.

En la figura 3.9 inciso A) se puede observar la sobreposición de las 20 estructuras del cálculo 3, después de la dinámica molecular con Amber. En los incisos B) y C) se representa la estructura de menor energía en listones donde se aprecia una hoja beta antiparalela entre los aminoácidos Val₂₂-Cys₂₆ y Cys₃₄-Arg₃₈; incluso se llega a formar el inicio de una tercera hebra con los residuos His₈-Ser₉. Esta estructura fue validada y depositada en el PDB (Protein data bank) con el identificador 6AX2.



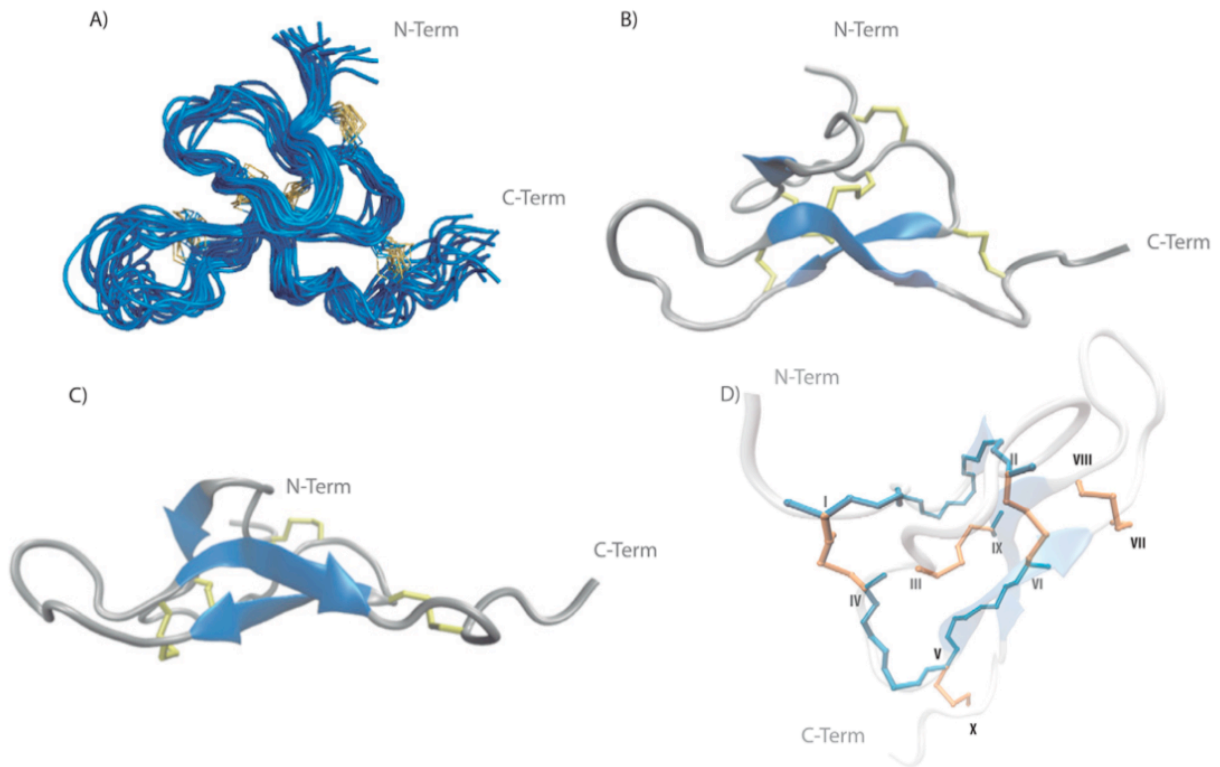


Figura 3.9. Estructura tridimensional de Magi3. A) Sobreposición de las 20 estructuras de menor energía. B) Representación de la estructura de Magi3 en listones. C) Motivo estructural “nudo de cisteínas” presente en la Magi3.

Analizando a fondo la estructura obtenida podemos observar la presencia del motivo estructural ICK, estas siglas vienen del inglés *“inhibitory cystine-knot”* el cual es un nudo de cisteínas generado por tres enlaces disulfuro, dos enlaces forman un anillo y el tercero los cruza (Norton y Pallaghy, 1998). El ICK es un motivo que genera una gran estabilidad y es uno de los motivos más comunes en las toxinas de araña con actividad sobre canales iónicos aunque también se ha observado en otras toxinas como las de conos (Craik et al., 2001).

Aunque todavía no existe una asociación directa entre los péptidos neurotóxicos que tienen el motivo ICK y una especificidad dada hacia los receptores celulares, Magi3 es un ejemplo de una toxina de araña con actividad insecticida que presenta cinco enlaces disulfuro de los cuales tres de ellos forman dicho motivo.



3.4.5. Actividad biológica de Magi3

Se probó la actividad de Magi3 inyectando la toxina a grillos y en el cerebro de ratones. Los resultados obtenidos se muestran en la tabla 3.3. Magi3 presentó un efecto letal en grillos a concentraciones altas, y se pudo observar un efecto paralizante hasta una concentración de 0.93 μg de toxina por gramo de insecto. Cuando se inyectó en el cerebro de ratón no se observó ningún efecto, con 1 y 5 μg , esto nos indica la alta selectividad que tiene la toxina por los canales de insecto.

Tabla 3.3. Pruebas de actividad de Magi3 en grillos y ratones.

Toxicidad en insectos		Toxicidad en ratones	
μg toxina /g de insecto	Resultados	Conc.	Resultados
119.00	Parálisis	1 μg por ratón	Sin efecto
79.80	Parálisis		
28.66	Parálisis	5 μg por ratón	Sin efecto
19.23	Parálisis		
3.63	Parálisis		
0.93	Parálisis		
0.71	Sin efecto		

Dada su actividad Magi3 es clasificada como una toxina con afinidad a canales de sodio (Corzo et al. 2003). Sin embargo, no se encuentra en ninguna familia de las toxinas con dicha actividad. En el 2012 el Dr. Klint y colaboradores realizaron una clasificación general de toxinas de araña con actividad sobre canales de sodio; dicha clasificación generó 12 familias basadas en el sitio de acción, el tamaño de la toxina, así como el número y la distribución de cisteínas que contenían estas toxinas (Klint et al., 2012).

En la figura 3.10 se pueden observar las diferentes familias de toxinas de araña con actividad sobre canales de sodio. Estas van desde los 29 hasta los 79 aminoácidos y contienen de 6 a 14 cisteínas.



Clasificar a Magi3 dentro de una de estas familias no es tarea fácil. Por el número de cisteínas el único grupo en el que encaja es en la familia seis; sin embargo, cuando analizamos sus secuencias y la posición, de las cisteínas observamos que la cisteína número 3 no se encuentran en la misma posición por lo que no correspondería a esta familia (Figura 3.11). De todas estas familias, hay muy poca información estructural, comparando las estructuras que se encuentran depositadas en el PDB encontramos grandes similitudes con las toxinas de la familia cuatro. En la figura 3.11 podemos observar dos estructuras de toxinas pertenecientes a esta familia; es claro ver su gran similitud con Magi3. Las tres comparten el motivo estructural ICK y forman la hoja beta en la misma posición; sin embargo, esta familia esta compuesta por toxinas con 4 enlaces disulfuro y aunque las conectividades son similares, sin tomar en cuenta el quinto enlace de la Magi3, no se podría clasificar aquí.

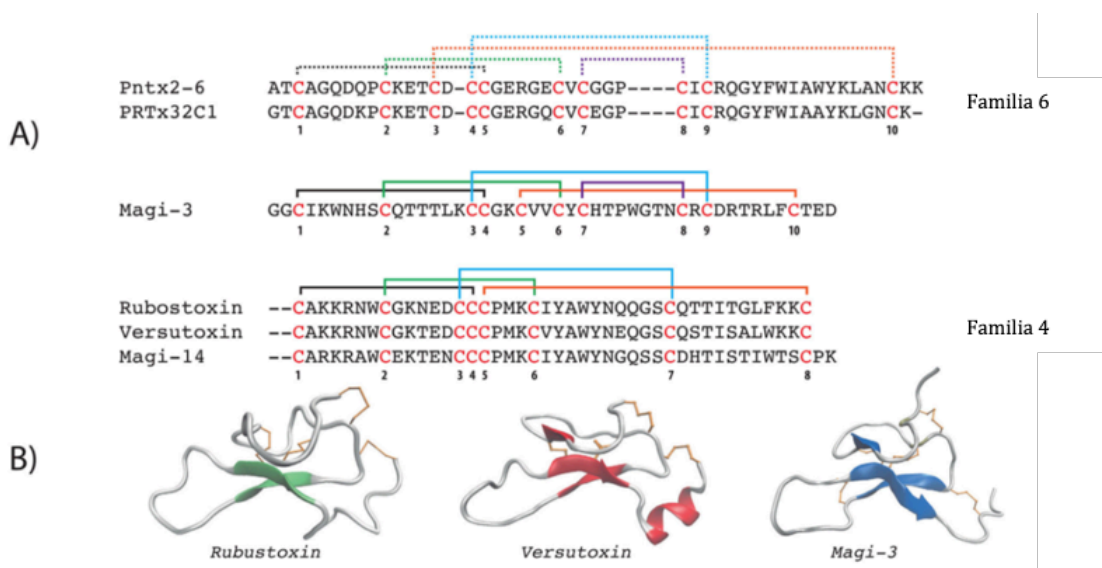


Figura 3.11. Comparación secuencial y estructural de Magi3. A) Comparación de la secuencia de Magi3 con la familia 4 y 6 de toxinas con actividad sobre canales de sodio. B) Estructuras tridimensionales de Magi3, Rubustoxina y Versutoxin.

Dado que la Magi3 es considerada una toxina con actividad sobre canales de sodio y visto que no es posible incluirla en ninguna de las doce familias existente, Magi3 sería la primera toxina de una nueva familia de toxinas con estas características.



Se logró la caracterización estructural de Magi3, la cual, fue obtenida de manera recombinante después de un proceso de re-plegamiento en donde con las condiciones propuestas generan una única conformación. La Magi3 presenta un motivo estructural ICK característico en las toxinas de araña. También se observó que en su estructura tridimensional se forma una hoja beta antiparalela y se logró determinar la conectividad de los 5 enlaces disulfuro (C3–C18, C10–C24, C17–C36, C21–C43 y C26–C34). La toxina obtenida muestra una actividad selectiva sobre insectos. Este resultado junto con la obtención de una única conformación demostrando que obtuvimos la toxina con su plegamiento nativo. Dada las características de esta toxina proponemos que Magi3 pertenece a una nueva familia de toxinas con actividad sobre canales de sodio.



- Para mejorar la caracterización de Magi3 se podrían realizar ensayos de actividad y determinar su IC_{50} en canales Nav de insecto.
- Evaluar si Magi3 tiene actividad sobre canales iónicos de calcio y comprobar la presunta actividad promiscua que sugiere las homologías con otras proteínas.
- Para determinar los aminoácidos clave de Magi3 que interaccionan con el canal se podrían realizar mutaciones puntuales y monitorear como se ve afectada su actividad.



- Alexander, S.P.H., Benson, H.E., Faccenda, E., Pawson, A.J., Sharman, J.L., Catterall, W.A., Spedding, M., Peters, J.A., Harmar, A.J. (2013) The Concise Guide to pharmacology 2013/14: ion channels. *Br. J. Pharmacol.* 170, 1607-1651.
- Burgess-Brown, N.A., Burgess-Brown N. A., Sharma, S., Sobott, F., Loenarz,C., Oppermann, U., Gileadi, O. (2008) Codon optimization can improve expression of human genes in *Escherichia coli*: a multi-gene study. *Protein Expr. Purif.* 59, 94–102.
- Cavanagh, J. (1996) *Protein NMR Spectroscopy*. Academic Press. EUA.
- Cheung, R.C., Wong, J. H., Ng, T. B. (2012) Immobilized metal ion affinity chromatography: a review on its applications. *Appl. Microbiol. Biotechnol.* 96, 1411–1420.
- Collet, J.F., Bardwell, J.C. (2002) Oxidative protein folding in bacteria. *Mol. Microbiol.* 44, 1-8
- Corzo G, Gilles N, Satake H, Villegas E, Dai L, Nakajima T, Haupt J. (2003) Distinct primary structures of the major peptide toxins from the venom of the spider *Macrothele gigas* that bind to sites 3 and 4 in the sodium channel. *FEBS Lett.* 547, 43–50.
- Craik DJ, Daly NL, Waine C. (2001) The cystine knot motif in toxins and implications for drug design. *Toxicon.* 39, 43–60.
- Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR.* 6, 277–293.
- Diego-García, E., Abdel-Mottaleb, Y., Schwartz, E. F., de la Vega, R. C., Tytgat, J., Possani, L. D. (2008) Cytolytic and K⁺ channel blocking activities of β -KTx and scorpine-like peptides purified from scorpion venoms. *Cell. Mol. Life Sci.* 65, 187–200
- Diego-García, E., Schwartz, E. F., D’Suze, G., González, S. A., Batista, C. V., García, B. L., de la Vega, R. C., Possani, L. D. (2007) Wide phylogenetic distribution of Scorpine and long-chain β -KTx-like peptides in scorpion venoms: Identification “orphan” components. *Peptides* 28, 31–37
- El-Din, T., Lenaeus, M., Zheng, N., Catterall, W. (2018) Fenestrations control resting-state block of a voltage-gated sodium channel. *PNAS* 115 (51), 13111–13116
- Escoubas, P. (2006) Molecular diversification in spider venoms: a web of combinatorial peptide libraries. *Mol. Divers.* 10, 545-554.
- Feng, J., Yu, C., Wang, M., Li, Z., Wu, Y., Cao, Z., Li, W., He, X., Han, S. (2013) Expression and characterization of a novel scorpine-like peptide Ev37, from the scorpion *Euscorpions validus*. *Protein Expr. Purif.* 88, 127–133.
- Fischer, B., Sumner, I., Goodenough, P. (1993) Isolation, renaturation and formation of disulfide bonds of eukaryotic proteins expressed in *E. coli* as inclusion bodies. *Biotechnol. Bioeng.* 41, 3–13.
- Güntert, P., Mumenthaler, C., and Wüthrich, K. (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.* 273, 283–298.
- Güntert, Peter. (2004). Automated NMR structure calculation with CYANA. *Methods Mol. Biol.* 278, 353-378. *Methods in molecular biology* (Clifton, N.J.). 278. 353-78.
- Hernández-Salgado K, Estrada G, Olvera A, Coronas FI, Possani LD, Corzo G (2009) Heterologous expressed toxic and non-toxic peptide variants of toxin Cssl are capable to produce neutralizing antibodies against the venom of the scorpion *Centruroides suffusus suffusus*. *Immunol Lett.* 125, 93–



99.

- Israel, M. R., Morgan, M., Tay, B., Deuis, J. R., (2018) Toxins as tools: Fingerprinting neuronal pharmacology. *Neuroscience. Letters.* 679, 4-14.
- Israel, M.R., Tay, B., Deuis, J.R., Vetter, I. (2017) Sodium channels and venom peptide pharmacology, *Adv. Pharmacol.* 79, 67–116.
- Jiménez, M.A., Curso avanzado de resonancia magnética nuclear, Estructura de péptidos y proteínas en disolución. Escuela de verano universidad de Zaragoza, Jaca. 15-20 de junio 1997.
- Keeler, J. (2005) *Understanding NMR spectroscopy.* John Wiley & Sons. England
- Keller R (2004) *The computer aided resonance assignment tutorial.* Goldau, Switzerland: Cantina Verlag, pp. 1–81.
- King, G., Hardy, M., (2013) Spider-Venom Peptides: Structure, Pharmacology, and Potential for Control of Insect Pests. *Annu. Rev. Entomol.* 58, 475–96.
- Klint, J.K., Senff, S., Rupasinghe, D., Er, S. Y., Herzig, V., Nicholson, G. M., King, G. F. (2012) Spider-venom peptides that target voltage-gated sodium channels: Pharmacological tools and potential therapeutic leads. *Toxicon.* 60, 478–491
- Kozlov, S. (2018) Animal toxins for channelopathy treatment. *Neuropharmacology.* 132, 83-97
- Laws, D.D., Dedios, A.C., Oldfield, E. (1993) NMR Chemical Shifts and Structure Refinement in Proteins. *J Biomol NMR.* 3, 607–612
- Lecomte, C., Sabatier, J.M., Van Rietschoten, J., Rochat, H. (1998) Synthetic peptides as tools to investigate the structure and pharmacology of potassium channel-acting short-chain scorpion toxins. *Biochimie.* 80, 151–154.
- Legros, C., Céard, B., Bougis, P. E., Martin-Eauclaire, M. F. (1998) Evidence for a new class of scorpion toxins active against K⁺ channels. *FEBS Lett.* 431, 375–380
- Long, C., McAnally, J.R., Shelton, J.M., Mireault, A.A., Bassel-Duby, R., Olson, E.N. (2014). Prevention of muscular dystrophy in mice by CRISPR/Cas9-mediated editing of germline DNA. *Science* 345, 1184-1188.
- López-Méndez B, Güntert P (2006) Automated protein structure determination from NMR spectra. *J Am Chem Soc,* 128:13112–13122.
- Mompean, M., Li, W., Li, J., Laage, S., Siemer, A.B., Bozkurt, G., Wu, H., McDermott, A.E. (2018) The structure of the necrosome RIPK1-RIPK3 core, a human hetero-amyloid signaling complex, *Cell* 173, 5 1244–1253.
- MacKinnon, R. (2003) Potassium channels. *FEBS Letters.* 555, 62-65
- Macura, S. and Ernst, R. R. (1980) Elucidation of cross relaxation in liquids by 2D NMR spectroscopy. *Mol. Phys.* 41, 95–117.
- Maeder, M.L., Gersbach, C.A. (2016) Genome-editing technologies for gene and cell therapy. *Mol. Ther.* 24, 430-446.
- Norton RS, Pallaghy PK. (1998). The cystine knot structure of ion channel toxins and related polypeptides. *Toxicon* 36:1573–1583.



- Overton, T. W. 2014. Recombinant protein production in bacterial hosts. *Drug Discovery Today* 19, 590–601.
- Quintero-Hernández, V., Ortiz, E., Rendón-Anaya, M., Schwartz, E. F., Becerril, B., Corzo, G., Possani, L. D. (2011). Scorpion and spider venom peptides: gene cloning and peptide expression. *Toxicon*. 58, 644–663.
- Reddy, K.R.C., Lilie, H.; Rudolph, R.; Lange, C. (2005). L-Arginine increases the solubility of unfolded species of hen egg white lysozyme. *Protein Sci.* 14, 929-935.
- Reynoso, C, N. (2014). Biosíntesis, Purificación y caracterización de la toxina de veneno de alacrán TsTXK-b, Tesis-UNAM
- Reznikoff, W.S. (1992). The lactose operon-controlling elements: a complex paradigm. *Mol. Microbiol.* 6, 2419-2422
- Rogowski, R. S., Krueger, B. K., Collins, J. H., Blaustein, M. P. (1994). Tityustoxin K alpha blocks voltage-gated non-inactivating K⁺ channels and unblocks inactivating K⁺ channels blocked by alpha-dendrotoxin in synaptosomes. *Proc. Natl. Acad. Sci. USA* 91, 1475 – 1479.
- Zhu, S., Li, W., Zeng, X., Jiang, D. Mao, X., Liu, H. (1999) Molecular cloning and sequencing of two 'short chain' and two 'long chain' K⁺ channel-blocking peptides from the Chinese scorpion *Buthus martensii* Karsch, *FEBS Lett.* 457, 509–514.
- Salceda, E., Ortega, A. (2009) Neurotoxinas: significado biológico y mecanismos de acción. *Elementos*. 74, 29-37.
- Salomon-Ferrer, R., Case, D. A., Walker, R. C. (2013) An overview of the Amber biomolecular simulation package. *Wiley Interdiscip Rev Comput Mol Sci* 3, 198– 210.
- Sansom, M., Shrivastav, I., Brighta, J., Tatec, J., Capenera C., Biggin, P., 2002. Potassium channels: structures, models, simulations. *Biochimica et Biophysica Acta* 1565, 294–307
- Singh, S. M., Panda, A. K. (2005) Solubilization and refolding of bacterial inclusion body proteins. *J. Biosci. Bioeng.* 99, 303–310.
- Studier, F.W. (1991). Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. *J. Mol. Biol.* 219, 37-44
- Tamayo, L. (2007) Bloqueo de los canales de potasio en el shock séptico, ¿otra esperanza perdida?, *med intensiva* 31 (5), 252-257.
- Tytgat, J., Chandy, K.G., Garcia, M.L., Gutman, G.A., Martin-Eauclaire, M.F., van der Walt, J.J., Possani, L.D. (1999) A unified nomenclature for short-chain peptides isolated from scorpion venoms: alpha-KTx molecular subfamilies, *Trends Pharmacol. Sci.* 20 (11) 444-447.
- Vallejo, L. F., Rinas, U. (2004) Strategy for recovery of active protein through refolding of bacterial inclusion body proteins. *Microb. Cell Fact.*, 3, 2–12.
- Wanger, G., Wüthrich, K., (1982) Sequential Resonance Assignments in Protein H-1 Nuclear Magnetic-Resonance Spectra – Basic Pancreatic Trypsin-Inhibitor. *J. Mol. Biol.* 155, 347-366.
- Wickman, K., Clapham, D. (1995) Ion channel regulation by G proteins. *Physiol. Rev.* 75, 865-885.
- Wüthrich, K. (2003) NMR studies of structures and function of biological macromolecules. Nobel lecture, Vol. 42, 3340-3363



- Zhu, S. & Tytgat, J. (2004) The scorpine family of defensins: gene structure, alternative polyadenylation and fold recognition. *Cell. Mol. Life Sci.* 61, 1751–63
- Zhu, S., Gao, B., Aumelas, A., del Carmen Rodríguez, M., Lanz-Mendoza, H., Peigneur, S., Diego-García, E., Martin-Eauclaire, M.-F., Tytgat, J., Possani, L.D., (2010) MeuTXK β 1, a scorpion venom-derived two-domain potassium channel toxin-like peptide with cytolytic activity. *Biochimica Biophysica Acta (BBA) e Proteins Proteomics* 1804, 872-883.





Successful refolding and NMR structure of rMagi3: A disulfide-rich insecticidal spider toxin

Gustavo Titaux-Delgado,¹ Elisa Carrillo,² Angeles Mendoza,³
Marlen Mayorga-Flores,¹ Fátima C. Escobedo-González,¹
Patricia Cano-Sánchez,¹ Estuardo López-Vera,² Gerardo Corzo,⁴ and
Federico del Rio-Portilla^{1*}

¹Departamento de Química de Biomacromoléculas, Instituto de Química, Universidad Nacional Autónoma de México, CU, Ciudad de México, 04510, México

²Laboratorio de Toxinología Marina, Unidad Académica de Ecología y Biodiversidad Acuática, Instituto de Ciencias del Mar y Limnología, Universidad Nacional Autónoma de México, Ciudad Universitaria, Ciudad de México, 04510, México

³División de Ciencias e Ingenierías, Campus León, Universidad de Guanajuato, Loma del Bosque 103, Col. Lomas del Campestre León, Gto, 37150, México

⁴Departamento de Medicina Molecular y Bioprocesos, Instituto de Biotecnología, Universidad Nacional Autónoma de México, UNAM, Apartado Postal 510-3, Cuernavaca, Morelos, 61500, México

Received 8 November 2017; Accepted 14 December 2017

DOI: 10.1002/pro.3363

Published online 16 December 2017 proteinscience.org

Abstract: The need for molecules with high specificity against noxious insects leads the search towards spider venoms that have evolved highly selective toxins for insect preys. In this respect, spiders as a highly diversified group of almost exclusive insect predators appear to possess infinite potential for the discovery of novel insect-selective toxins. In 2003, a group of toxins was isolated from the spider *Macrothele gigas* and the amino acid sequence was reported. We obtained, by molecular biology techniques in a heterologous system, one of these toxins. Purification process was optimized by chromatographic methods to determine the three-dimensional structure by nuclear magnetic resonance in solution, and, finally, their biological activity was tested. rMagi3 resulted to be a specific insect toxin with no effect on mice.

Keywords: refolding disulfide-bonded toxins; NMR solution structure; spider toxin; insecticidal; *Macrothele gigas*

Abbreviations: CNBr, cyanogen bromide; TFA, trifluoroacetic acid; can, acetonitrile; DTT, 1,4-dithiothreitol; DPFGE, Double Pulsed Field Gradient Spin Echo; MALDI-TOF MS, matrix assisted laser desorption ionization-time of flight mass spectrometry.

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: Dirección General del Personal Académico, UNAM; Grant number: IN204116, IN702516.

*Correspondence to: [F. del-Rio Portilla, Química de Biomacromoléculas 1, Instituto de Química, Universidad Nacional Autónoma de México, CU, Ciudad de México, 04510, México]. E-mail address: federico.delrio@gmail.com or jfrp@unam.mx

Introduction

Animal venoms are one of the most intriguing natural products already studied. Among them, spider venoms are chemical cocktails consisting mainly of low-molecular weight polypeptide toxins and large enzymatic proteins with an enormous value of insecticidal leads,¹ which have great potential for being used as biopesticides and in biotechnological applications due to their high specificity on insect ion channels.² Therefore, studies on the relationship between toxin structures and the pharmacology of insecticidal ion channels are fundamental to understand and to develop more efficient and selective



insecticidal toxins.³ However, the amount of venom material usually obtained in spider venoms is low, which often had limited structural studies and applications of many insecticidal spider toxins to pursuit further research. Other difficulties observed in studying some spider neurotoxins is their modified N-termini, such as formation of pyroglutamate, or most commonly C-terminal amidation.³ Also, approximately 90% of spider venom peptides are compact globular proteins possessing several disulfide bridges, which usually range from two to seven. Furthermore, spider toxins often contain a “disulfide pseudoknot” which places them in a class of toxins and inhibitory polypeptides with an “inhibitory cystine-knot” (ICK) motif.⁴ This structural motif is currently exemplified as an antiparallel β -hairpin stabilized by a cystine-knot^{5,6} and this has been reported in the majority of spider toxin structures.

In this work, we solved the structure of the insecticidal spider toxin rMagi3, which was originally obtained from the venom of the hexathelidae spider *Macrothele gigas*. Magi 3 is a 46-residue polypeptide containing five disulfide bridges, and it affects the insect sodium channels.⁷ Magi 3 was expressed heterologously in *E. coli* and refolded *in vitro*. It had a low affinity to the Nav1.4 channel confirming its absence of toxicity to mice, but its effective insecticidal preference in paralyzing house crickets.

Results

Expression and purification

The trx-6His tagged rMagi3 was recombinantly expressed in *E. coli* Rosetta gami/pET32a. The protein was obtained soluble from the cellular lysate, and it was efficiently retained by a HiTrap Ni²⁺ column [Fig. 1(A)].

Since uncontrolled disulfide-bond oxidative scramble of cysteine-rich arachnid toxins could occur during bacterial expression,^{8–10} the trx-6His-rMagi3 fusion protein was submitted to a refolding process to improve the yield of a well-structured rMagi3. After that, the trx-6His-rMagi3 was enzymatically cleaved with thrombin, and the trx-His6-LVPR residue was satisfactorily removed from the reaction mixture by a second IMAC separation [Fig. 1(B)]. Purification of the non-retained fraction was performed by RP-HPLC. It was observed that there was only one major component at 16.8 min, the protein fraction of interest contained the molecular mass of 5782.2 Da corresponding to the theoretical molecular mass of the peptidic fragment SAAAAGM-Magi3 (Fig. 2). A broad protein fraction was observed at 16.0 min, which did not show any of the expected molecular masses that indicated the presence of a different recombinant isoform or conformer. The efficiency of the enzymatic proteolysis was 25% after

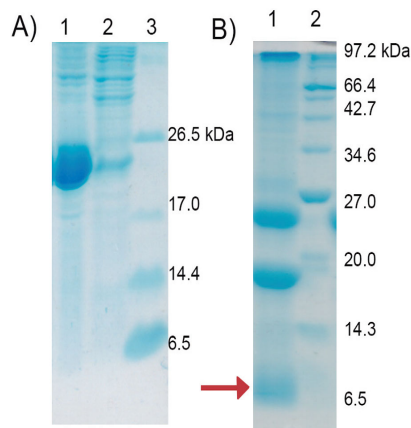


Figure 1. Expression, purification and proteolysis of fusion protein from rMagi3. (A) SDS-PAGE analysis of the expression and purification of the trx-6His-rMagi3. Lane 1, fusion protein after affinity chromatography; lane 2, flow through; lane 3, molecular mass markers. (B) Lane 1; trx-6His-rMagi3 was cleaved by thrombin and the released peptide GSAAAAM-rMagi3 is labeled by an arrow; lane 2, molecular mass markers.

the refolding process. Therefore, to remove the N-terminal peptide SAAAAGM to obtain the reported mature toxin of Magi3, SAAAAGM-Magi3 was chemically hydrolyzed using CNBr to achieve the mature toxin named rMagi3. After chemical hydrolysis, rMagi3 was purified by HPLC and its molecular mass value agreed with the expected theoretical one of 5223.77 Da containing five disulfide bonds (Fig. 3). The final yield in LB media was 0.3 mg/L.

rMagi3 disulfide bond connectivities and three dimensional structure

NMR 1D-1H and the 2D-TOCSY finger region spectra are shown in Figures S1 and S2. It was not identified other conformation structures rather than that observed in Figures S1 and S2. After that, rMagi3 was subjected to enzymatic digestion with trypsin, and the disulfide-bond connectivities were established based on the NMR constraints and the results of the enzymatic digestion. The major fragment from the enzymatic cleavage had a molecular mass of 3781.7 Da, which corresponds to the amino acid sequences that comprises the connectivities between the cysteines C₂, C₅, C₆, C₇, C₈ y C₁₀ (Fig. 3). Also, the position of the disulfide bridges was established by analyzing the NMR experimental results independently. A NOE signal between H β – H β from Cys 10 and Cys 24, which corresponds to C₂-C₆ cystine pair, was found. Also, there were observed NOE correlations between H α of Cys 34 with H α of Cys 26 and HN of His 27 and HN Try 25 with HN Arg 35;

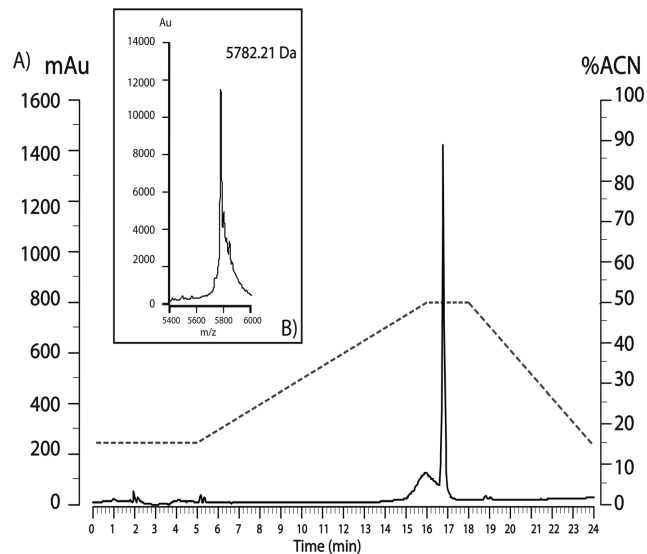


Figure 2. (A) RP-HPLC purification of GSAAAAM-rMagi3. (B) The inset shows the mass spectrum of GSAAAAM-rMagi3.

H α of Cys 21 with HN of Arg 40 and H β of Cys 21 with HN of Arg 40; H β of Cys 3 with H α of Cys 18 and H α of Cys 3 with HN of Cys 18; all of these signals are in agreement with the C₇-C₈, C₅-C₁₀ and C₁-C₄ disulfide pairs, respectively. The NOE connectivities are shown in Figure S3.

Most of the spin systems and proton resonance signals were unambiguously assigned (43 out of 46 amino acids). The ¹H proton spectrum showed dispersed signals, which are typical characteristics of a folded protein. A well dispersed signal at 9.8 ppm from NH of Cys 18 which is forming a hydrogen bond with the carbonyl of Ile 4 (data not shown).

Three CYANA calculations were performed with different input parameters to analyze the disulfide bond connectivities. The first calculation included data without disulfide bond restrictions. This result provided the NOE constraint contributions for the structure, which was used as a reference. The second calculation was done using disulfide bond restrictions as mentioned in Uniprot: P83559 (C₁-C₄, C₂-C₆, C₅-C₈, C₃-C₉ and C₇-C₁₀). Finally, the last set of structures were calculated considering the most proximate cysteines in the structure generated by the calculation without disulfide bond restrictions to form the restrictions of disulfide bridges (C₁-C₄,

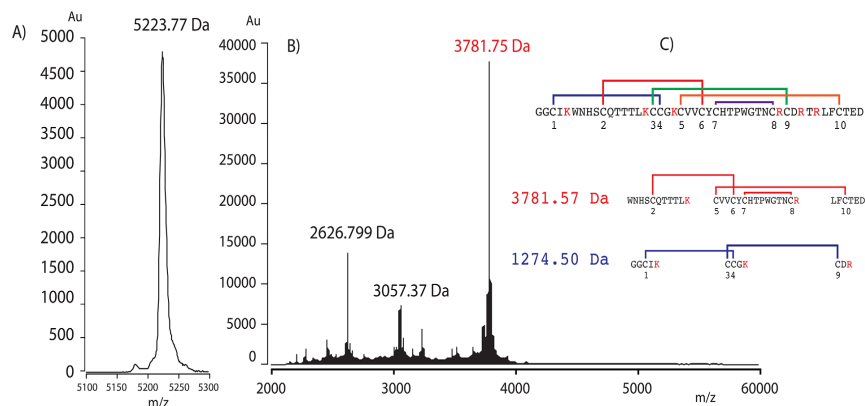


Figure 3. Mass spectra of rMagi3 after CNBr reaction and after proteolysis with trypsin. (A) Mass spectrum of rMagi3. (B) Fragments generated by the enzymatic digestion of rMagi3 with trypsin. (C) Proposed connectivities based on (B).

C₂-C₆, C₃-C₉, C₅-C₁₀ and C₇-C₈) and given that these bonds were consistent with the molecular mass of the fragments obtained after trypsin hydrolysis.

Table I shows the statistics and RMSD values of these three calculations. After the comparison of the backbone from residues 3 to 43 of the lowest energy structure for each set of calculations, the RMSD of the reference structure resulted of 2.6 Å for the disulfide bridges proposed by Uniprot, and of 1.2 Å for the disulfide bridges anticipated in this work. A total of 200 structures were generated with CYANA 2.1, and only 20 structures with the lowest target function are shown. We also performed molecular dynamic simulations in AMBER 9 to refine unrestricted regions of the structural models obtained in CYANA (Fig. 4). As a result of the molecular dynamic calculations using only NOE data, all proline residues were found to be in trans conformation. rMagi3 adopts an ICK motif commonly observed in neurotoxic arachnid polypeptides;^{11,12} that is, rMagi3 possesses a cysteine knot formed by three disulfide bonds paired as C₁-C₄, C₂-C₆, and C₃-C₉. The disulfide bond C₃-C₉ passes through a 16-residue ring formed by the other two disulfide bonds (C₁-C₄ y C₂-C₆) and the corresponding polypeptide backbone. The main structural characteristic of rMagi3 is a double-stranded antiparallel-sheet [Fig. 4(B)], which is formed by the residues Val22-Cys26 and Cys34-Arg38, and by a third strand that crosses over the residues His8-Ser9. At the N- and C-terminal regions, the rMagi3 structure is quite disordered, which is a consequence of the lack of long range NOEs constrains. The Magi3 structure is in good agreement with the chemical shift index¹³ (see Fig. S4).

Voltage-gated ion channels assay in oocytes

Previously, it was reported that Magi3 inhibits the voltage-gated insect sodium channel, to evaluate whether the effect of rMagi3 was specific for voltage-gated insect sodium channels, we explored the activity of rMagi3 in rat Nav1.4 channel. Where rMagi3 inhibited 23 ± 3% and 35 ± 2% of the Nav1.4 ion currents at concentrations of 50 and 100 μM, respectively (Fig. 5); without affecting the V_{1/2} (-25 ± 2.1, control, vs. 22.5 ± 1.8, rMagi3 100 μM). This data fit was obtained using a Boltzmann equation. Moreover, the time constant of inactivation at -10 mV were fit by a single exponential, with τ control of 10 ± 0.1 ms and τ rMagi3 of 5 ± 0.2ms. These values indicates that rMagi3 has not mammalian toxicity.

Insect and mice activity

The insecticidal activity of rMagi3 was tested using doses from 0.93 to 119 μg/g crickets. The minimal amount used caused neurotoxic effects immediately

Table I. Final Set of NOEs Included in the Structures Calculated by CYANA^a

	Calc 1	Calc 2	Calc 3
NMR statistics			
Total	421	401	419
Short-range, i-j ≤ 1	306	306	306
Medium-range, 1 < i-j < 5	33	33	29
Long-range, i-j ≥ 5	82	62	84
Average target function value	0.08	0.32	0.02
Average backbone RMSD to mean	1.47	2.6	1.2
Ramachandran statistics			
Most favored regions	58.5%	49.2%	55.6%
Additional allowed regions	41.0%	49.5%	44.4%
Generously allowed regions	0.0%	1.0%	0.0%
Disallowed regions	0.2%	0.2%	0.0%

^a The initial data set of NOE restraints for all calculations was the same. The first set was calculated using only NOE restraints and no disulfide bridge constrains (Calc 1). A second structure calculation set was done considering using as disulfide bond restrictions mentioned in Uniprot (Calc 2): C₁-C₄, C₂-C₆, C₅-C₈, C₃-C₉ and C₇-C₁₀. Last calculations were done including disulfide bond restrictions observed (Calc 3): C₁-C₄, C₂-C₆, C₃-C₉, C₅-C₁₀ and C₇-C₈.

after intra-thoracic injection. The knock/down effect was immediate; insects were unable to stand up, or to show leg movements. The largest dose caused the insect's death. On the other hand, mice were intracranially injected using doses of 1 and 5 μg per mouse (Table II), rMagi3 did not show toxic effects up to 5 μg/mouse (4.8 pmol/g). Mice were observed for a period of 24 h.

Discussion

rMagi3 was obtained as a single molecule as observed by reversed-phase chromatography, mass spectrometry and NMR spectra; so, considering that this is the first structure of a recombinant spider toxin with five disulfide bonds, it is a remarkable result. The NOE's restriction data suggest that the disulfide bonds are formed by C₁-C₄, C₂-C₆, C₃-C₉, C₅-C₁₀, and C₇-C₈, which are different from those suggested in the Uniprot data base that points out the disulfide bonds C₅-C₈ and C₇-C₁₀ instead of the C₅-C₁₀ and C₇-C₈ bonds. The ICK motif, which is a highly conserved characteristic in spider toxins, was present in rMagi3, and it could thermodynamically explain the only structure found after the refolding process. The conservation of the ICK in several neurotoxic peptides appears to be an evolutionary strategy to be used by venomous organisms to display chemical and biological diversity on an optimized and stable protein scaffold.¹⁴ Several neurotoxins that conserve the ICK are active on sodium and calcium insect ion channels.^{15,16} rMagi3 is also the first five disulfide bond protein reported to present the ICF motif. Although, there is still no direct association between neurotoxic ICK peptides and a given specificity towards cell receptors; rMagi3 is an



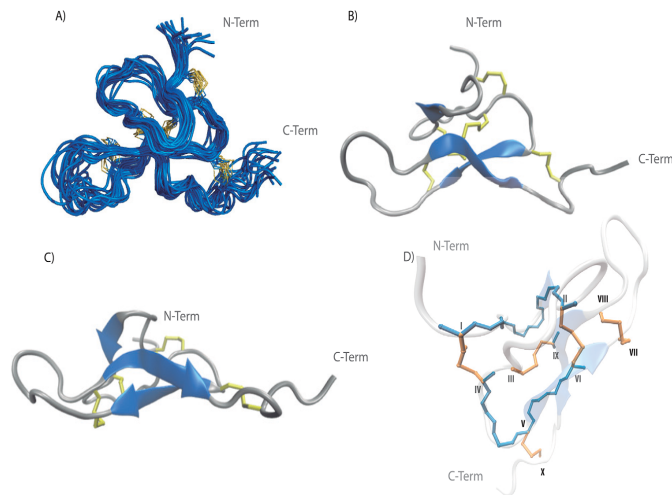


Figure 4. NMR structures of rMagi3. (A) Superimposition of the 20 lowest energy structures, only the backbone atoms are displayed. (B) and (C) Ribbon representations of two minimized structures, only the backbone trace and the disulfide bridges are shown. (D) Illustration of rMagi3 cystine knot. The disulfides bridges are shown as orange tubes.

example of a five disulfide bond insecticidal spider toxin with ICK motif. Since only one conformational structure in the recombinant protein with just one disulfide bond connectivity out of 945 possibilities (full oxidized peptide) was obtained, and the characteristic cystine knot motif for this type of toxins was observed, it is very likely that we obtained the Magi3 native structure.

Furthermore, rMagi 3 has a significant homology with a group of spider insecticidal toxins already described; for example, Ptlx-II, Ptlx-XII, Ptlx-V/VI, Ptlx-XI, and Ptlx-VIII isolated from venom of the North American spider *Plectreurys tristis*, and OxyTx1 from the venom of the Asian spider *Oxyopes sp.*^{8,17} All of them present 10 cysteines in the same

analog positions; however, the connectivity of their disulfide bonds and their three-dimensional structure have not been yet determined experimentally. Although, Magi 3 has been shown to bind to the insect's sodium channel, but Ptlx-II and OxyTx1 blocked the insect calcium channel,^{17,18} rMagi3 structure could be used to establish the disulfide bridges of such spider neurotoxins.

Although Magi3 inhibits the binding of the ¹²⁵I-labelled scorpion alpha-insect toxin Lqh-alpha-IT to the site-3 of the voltage-gated sodium channel in cockroach synaptosomes,⁷ Magi3 has not been assigned to any sodium channel spider toxin (NaSpTx) family according to Klint,¹⁹ who had classified spider neurotoxins into 12 different protein

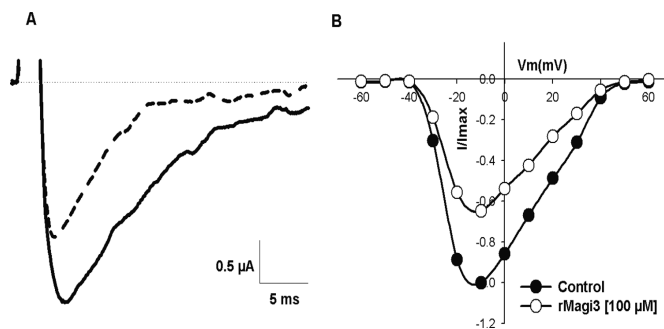


Figure 5. Inhibition of the skeletal muscle sodium channel by rMagi3. (A) Recordings of sodium Nav1.4 currents in control solution (solid line) and in the presence of 100 μ M rMagi3 (dash line). The membrane potential was set at -10 mV. (B) The I-V current-voltage relationships before (black circle) and after (white circle) the presence of rMagi3 (100 μ M). The oocytes were depolarized between -60 mV and $+60$ mV, in -10 mV increments from a holding potential of -100 mV.

Table II. Biological Effects of rMagi3 on Insects and Mice

Insect toxicity		Mouse toxicity	
μg toxin/g insect	Results	Conc.	Results
119.00	Paralysis	1 μg per mouse	No effects
79.80	Paralysis		
28.66	Paralysis	5 μg per mouse	No effects
19.23	Paralysis		
3.63	Paralysis		
0.93	Paralysis		
0.71	No effects		

families based on their primary structure and cysteine scaffold. If we want to include Magi 3 in Klint classification, we have to analyze the NaSpTx families 4 and 6. Family 6 is the only one that is presented by the 10 cysteine neurotoxins as Magi3. However, this peptide family contains primary structures between 47 and 54 residues, and the position of the cysteine C3 is different in rMagi3 [Fig. 6(A)]. It seems that the position of the cysteine of C5 of Magi3 is interchanged with C3 in such family 6; both of them form the bond with C10 [Fig. 6(A)]. The structural comparison of the rMagi3 and those found in the other families proves important similarities between the Magi3 with other well-studied spider neurotoxins such robustoxin²⁰ and versutoxin,²¹ which belong to family 4 [Fig. 6(B)]. The tertiary structure of Magi3 and family 4 toxins are also similar due to the fact that family 4 disulfide bonds correspond to those of Magi3 without the extra link C7-C8. The disulfide bond connectivity probably allows

these structural similarities; but they are not alike in their primary structure.

Because the activity of rMagi3 towards hNav1.4 was very poor, it was decided to test its toxicity directly on mice. rMagi3 did not show toxicity against mice; so we could assert that it is not toxic to mammals. The non-mammalian toxicity contrast with its notable insecticidal activity in house crickets. rMagi3 produced immediate effects causing paralysis and death of such insects. This results confirm our previous statement that rMagi3 disulfide bond connectivities are well established, since the correct formation of disulfide bond is indispensable for biological activity in this type of neurotoxins.²² Comparing its activity against crickets with other insecticidal neurotoxins from those belonging to NaSpTx family 4 and 6, and those which obtained from the same venom of *Macrothele gigas*, Magi3 has a very low LD₅₀ (191 nmol/g), see Table II. For example, versutoxin (from family 4) is active at 770 pmol/g of *Gryllus sp.* This straight comparison confirms that rMagi3 is four times more active than versutoxin. Also, curtatoxin II²³ (against *Acheta domesticus* with an LD₅₀ of 975 pmol/g) is five times less toxic than rMagi3. However, if we compare mammalian toxicity against mice of rMagi3 (LD₅₀ > 50 pmol/g) versus Magi4 and 5 (LD₅₀ 0.15 and 0.79 pmol/g, respectively),⁷ rMagi3 has none toxicity against mice. rMagi3 is one of several insect toxins purified from *Macrothele gigas*, but with high insecticidal potency.

Also, it is clear that the Magi3 is similar in the ICK motif among the NaSpTx families; however, the disulfide bridge formed by a large number of

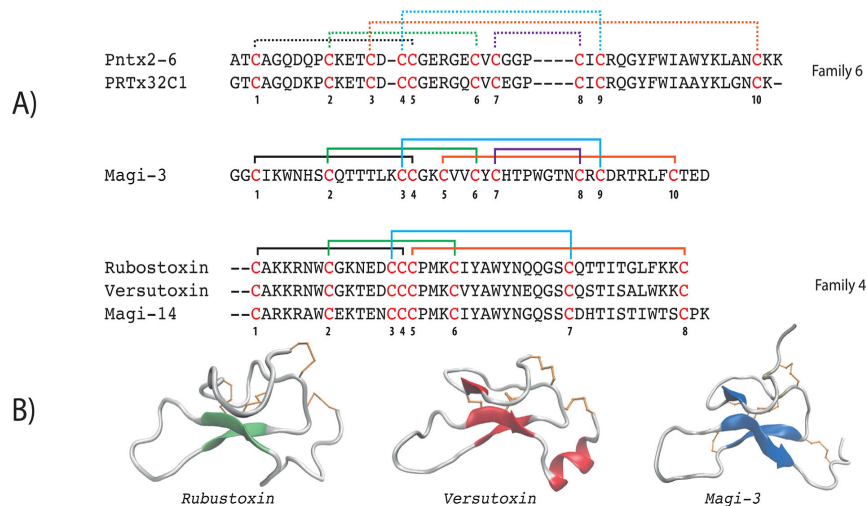


Figure 6. (A) Comparison of the amino-acid sequences of insect-selective spider toxins and (B) structural comparison of rMagi3 (PDB ID: 6AX2), robustoxin (PDB ID: 1QDP) and versutoxin (PDB ID: 1VTX). Continuous lines show disulfide link and dotted lines represent predicted disulfide bond connectivities which have not been experimentally determined.

cysteins, and its sequence let Magi3 to be an specific toxin for insects. The lack of identity with this kind of toxins strongly suggest that the ICK motif is not so relevant for its activity, but the distribution of aminoacids in the toxin surface it is.

Conclusions

We report for the first time the structure of the five disulfide-bonded rMagi3, an insecticidal toxin which shows notable activity on insects. The rMagi3 structure is an anti-parallel β -sheet with an additional cross β -strand. The disulfide bond connectivity of rMagi3 is a new motif found for these type of toxins (C₁-C₄, C₂-C₆, C₃-C₉, C₅-C₁₀, and C₇-C₈). The 3D structure shows similarities with that reported for the NaSpTx family 4, despite this family has only four disulfide bonded protein members, and they are not similar in their amino acid sequences. rMagi3 has not toxicity on mice considering several other spider toxins in such family 4; but, it shows a very good activity against house crickets. Our results strongly suggest that Magi3 belongs to a new insect toxin family of NaSpTx.

Materials and Methods

rMagi3 cloning

Magi3's gene was designed based upon the amino acid sequence reported in the UniProt entry P83559⁷ and codon-optimized for *E. coli* expression. The synthetic designed Magi3 gene was manufactured by Epoch Life Science (Missouri City, TX, USA) and cloned into the pET32a expression vector which includes a thrombin recognition site at the N-termini. The recombinant protein was expressed as a fusion thioredoxinA chaperon protein, a 6His-Tag for purification and a thrombin cleavage site for further elimination of the 6His-Tag. The genetic construction included also an N-termini GSAAAAM amino acid sequence just after the thrombin cleavage site; accordingly with our experience, it is necessary to have a considerable molecular distance between such cleavage site and the first cysteine residue, which form a disulfide bond to make the enzymatic cleavage more effective. The designed Magi3's gene was cloned into the pET32a vector, and the pET32a-rMagi3 constructed plasmid was confirmed by DNA sequencing from Laragen Inc. (Culver City, CA, USA).

Production, purification, and refolding of recombinant rMagi3

The recombinant peptide was expressed in *E. coli* Rosetta-gami strain as follows; bacterial cells already transformed with pET32a-rMagi3 were incubated in 1.5 L culture flasks containing LB media (Miller's modification) at 37°C and a rotation speed

of 220 rpm. When the optical density of the cell culture (600 nm) reached a value of 0.7 units, the expression of pET32a-rMagi3 was induced with 0.5 mM IPTG. Simultaneously, the incubation temperature was lowered to 30°C and maintained for 8 h more. The bacterial cells were collected by centrifugation at 6000g for 20 min, and subsequently, a gram of wet-weight cells were re-suspended in 7 mL of a buffer containing 50 mM Tris, 300 mM NaCl at pH 8.0. Cells were broken down by sonication (10 min pulses of 20 s with 55 W; 40 s delay between each pulse), and the soluble fraction was clarified by centrifugation at 37,500g for 45 min. The soluble proteins were fractionated by ion-metal affinity chromatography (IMAC; Hi-Trap chelating columns from GE-Healthcare, Waukesha, WI, USA), and the eluted protein fraction of interest was dialyzed overnight against Tris 50 mM at pH 8. The cysteines of the rMagi3 fusion protein were oxidized with buffer 1 (GndCl 5M, DTT 10 mM, Tris 50 mM pH = 8) for 3 h at 25°C. Later, the DTT was removed again by dialysis using Tris 50 mM and GndCl 1 M at pH 8, and a refolding buffer (50 mM Tris, 1 mM oxidized glutathione and 5 mM reduced glutathione at pH 8) at a flow rate of 1 mL per min until a final protein concentration of 80–120 μ g/mL was added. The refolding reaction was carried out for 4 days at 25°C, and after that, it was concentrated by ultrafiltration using an Amicon system (Millipore, Billerica, MA, USA) with a 10 kDa cut membrane at 4°C. The refolded protein was washed with 50 mM Tris pH 9 buffer, and the protein concentration, determined by Bradford, was adjusted to 1.5 mg/mL for further enzymatic digestion with thrombin. The enzymatic reaction was carried out at room temperature for 3 h in 50 mM Tris-HCl, pH 9. The cleaved protein products were purified again by IMAC to remove the 6His-tail of the rMagi3 fusion protein.

Reversed-phase chromatography

The final rMagi3 purification were performed with a Proteo C4 analytical reversed-phase column (300 Å pore size, 4.0 μ m particle size, and 4.6 \times 250 mm, Culver, CA, USA). The linear gradient was from 15 to 50% of solution B over 24 min at a flow of 1 mL/min, where solution A was 0.05% (v/v) aqueous TFA and solution B was 0.05% (v/v) TFA in ACN.

CNBr reaction

AAAAM-rMagi3 was directly dissolved in 50% (v/v) TFA to a final protein concentration of 1.7 mM. The protein solution was purged with nitrogen prior to the addition of an excess of 100-fold molar equivalents of cyanogen bromide (CNBr 5M; Sigma-Aldrich, St. Louis, MO, USA) for each methionine in the protein sequence. The reaction was protected from light for 24 h under stirring in N₂ atmosphere.



The reaction was quenched adding water with a proportion of four-fold reaction volume. The protein solution was diluted with Tris 50 mM at pH 8 buffer until the pH was higher than 1.5. Finally, the sample was recovered from the solution by reversed-phase HPLC purification as mentioned.

Mass spectrometry and disulfide bond connectivity determination

For disulfide bond connectivity determination, 5 μg of pure rMagi3 and 250 ng of trypsin were dissolved in 15 μL of reaction buffer (Tris 50 mM at pH 8.0) and then incubated at 37°C for 16 h. For MS analysis, 1 μL (ca 10 ng) of a given eluted fraction from HPLC was mixed with 1 μL of a supersaturated solution of alpha-cyano-4-hydroxy-cinnamic acid (acetonitrile-water 2:1 and 0.05% TFA). The samples were processed and analyzed by a Bruker Daltonics Microflex LT MALDI-TOF spectrometer.

NMR spectroscopy

rMagi3 was dissolved in deionized water and 5% (v/v) D_2O (Cambridge Isotopes Laboratories, Tewksbury, MA, USA). The pH was 3.5 and the protein concentration was 1 mM. The COSY-PS, 2D-TOCSY (80 ms of mixing time) and 2D-NOESY (150 and 300 ms of mixing time) experiments were acquired in a Unity Inova 500 (Varian, Palo Alto, CA, USA) and in a Bruker ASCEND 700 (Bruker BioSpin, Billerica, MA, USA) spectrometer. Water signal was successfully suppressed with Double Pulsed Field Gradient Spin Echo (DPFGSE).²⁴ The signals were weighted with sinusoidal function, zero filled to the double of number of points used in both dimensions.

Structure calculation

Experimental data sets were processed with NMRPIPE software.²⁵ Peak picking, spin systems identification, and NOE signals assignment were done in CARA 1.5.²⁶ Automatic NOE assignments and structure calculations were performed in CYANA 2.1²⁷ using the 300 ms mixing time NOESY signals. To confirm the results, the chemical shift indexes were calculated for H_α protons (see Fig. S4). The protein model was constructed with a set of 20 structural representations selected from calculations using the lowest energy value and no violations constraints. Individually, the assemblies were submitted to a molecular dynamics refinement in AMBER 9.²⁸ Refined structural model of rMagi3 was validated and uploaded to the Protein Data Bank (PDB ID: 6AX2) and Biological Magnetic Resonance Bank (BMRB ID 30339).

Bioassays

Insects. House crickets (*Gryllus bimaculatus*) were injected intrathoracically between the second and

third pair of legs, with different doses of rMagi3 previously dissolved in dH_2O . Controls were done with distilled water only and the crickets were closely observed up to 2 h. The classification “paralysis” was considered when the cricket lost all mobility for 2 h, but finally recovered. The lethal effect was said, when the crickets died after injection (see Table II).

Mice. The protocol used for assaying the activity of peptides *in vivo* was followed according to the guidelines of our Institute Committee of Animal Welfare, keeping the number of animals to a minimum. Male mice (CD-1, 20 g body weight) were tested by intracranial rMagi3 injection, where the symptoms were dose-dependent. rMagi3 was diluted to achieve a 1 $\mu\text{g}/\mu\text{L}$ concentration, with 0.15 M saline containing 0.25 mg/mL of BSA. The injections (1 and 5 μL), with a 10 μL micro-syringe fitted with a glass capillary, was performed mid-way between the left eye and the left ear (intracranial). Negative controls were done with 0.15 M saline containing 0.25 mg/mL of BSA only and positive controls with the neurotoxic scorpion peptide Cssi isolated in our laboratory.²⁹ Mice were observed for toxicity symptoms up to 24 h. Toxicity assessments in mice were classified as follows; mild toxicity was considered when the mice showed alterations such as restless and piloerection in its behavior after venom fraction injection; moderate toxicity was when the mice showed additional symptoms of intoxication such as salorrhea and/or hyperactivity, but recovered before 24 h; severe toxicity was considered when the post-24 h injection the mice still showed signs of intoxication (hypoactivity, forced abdominal breathing, partial paralysis of limbs, and exophthalmus), but it finally was recovered; and, lethal when the mice died.

Voltage-gated ion channels assay in oocytes

cRNA of the clone for rat Nav1.4 in vector pBS-SK (+/-) resistant to ampicillin was linearized with NotI and transcribed *in vitro* with T7 RNA polymerase (mMessage mMachine kit; Ambion, TX, USA). Electrophysiology was performed as described in detail previously.³⁰ Each defolliculated *Xenopus laevis* oocyte was injected with 10 ng of cRNA.

Voltage-clamp recordings were done (1–3 days after injection) with the membrane potential kept at -100 mV using a two-electrodes voltage-clamp amplifier (model OC-725 C, Warner Instruments Corp., Hamden, CT, USA) at room temperature. Sodium currents were generated by de-polarizations between -60 and 60 mV, in 10 mV increments by 20 ms. The composition of the ND96 electrophysiological solution was: NaCl 96 mM, KCl₂ mM, CaCl₂ 1.8 mM, MgCl₂ 1 mM, HEPES 5 mM, and pH 7.5.



The glass microelectrodes were filled with 3M KCl. The effect of rMagi3 was tested at least in three different oocytes.

Acknowledgments

GTD, MMF y FCEG acknowledges the National Council of Science and Technology (CONACyT) for scholarship numbers 273490, 289035 y 598794. EC was supported by a fellowship from Programa de Becas Posdoctorales, Dirección General de Asuntos del Personal Académico (DGAPA). This work was partially supported by Dirección General de Asuntos del Personal Académico with grants DGAPA-PAPIIT number 207516 (FRP) and IN204116 (ELV). This study was made possible in part thanks to the UNAMs NMR lab: LURMN at IQ-UNAM which is funded by CONACyT (Project 0224747) and UNAM. We also acknowledge to Alan Goldin (University of California, Irvine) for Nav1.4 clone.

References

- Escoubas P, Bosmans F (2007) Spider peptide toxins as leads for drug development. *Expert Opin Drug Discov* 2:823–835.
- Nicholson GM (2007) Insect-selective spider toxins targeting voltage-gated sodium channels. *Toxicon* 49: 490–512.
- Nicholson GM (2013) *Spider Peptides*, 2nd Ed., Elsevier Inc. doi: 10.1016/B978-0-000-0000-0000-00063-0000.
- Norton RS, Pallaghy PK (1998) The cystine knot structure of ion channel toxins and related polypeptides. *Toxicon* 36:1573–1583.
- Wang X, Connor M, Smith R, Maciejewski MW, Howden ME, Nicholson GM, Christie MJ, King GF (2000) Discovery and characterization of a family of insecticidal neurotoxins with a rare vicinal disulfide bridge. *Nat Struct Biol* 7:505–513.
- Tedford HW, Sollod BL, Maggio F, King GF (2004) Australian funnel-web spiders: Master insecticide chemists. *Toxicon* 43:601–618.
- Corzo G, Gilles N, Satake H, Villegas E, Dai L, Nakajima T, Haupt J (2003) Distinct primary structures of the major peptide toxins from the venom of the spider *Macrothele gigas* that bind to sites 3 and 4 in the sodium channel. *FEBS Lett* 547:43–50.
- Estrada G, Silva AO, Villegas E, Ortiz E, Beirao PSL, Corzo G (2016) Heterologous expression of five disulfide-bonded insecticidal spider peptides. *Toxicon* 119:152–158.
- Clement H, Flores V, Diego-Garcia E, Corrales-Garcia L, Villegas E, Corzo G (2015) A comparison between the recombinant expression and chemical synthesis of a short cysteine-rich insecticidal spider peptide. *J Venom Anim Toxins Incl Trop Dis* 21:19.
- Clement H, Flores V, De la Rosa G, Zamudio F, Alagon A, Corzo G (2016) Heterologous expression, protein folding and antibody recognition of a neurotoxin from the Mexican coral snake *Micrurus laticorallus*. *J Venom Anim Toxins Incl Trop Dis* 22:25.
- Craik DJ, Daly NL, Waite C (2001) The cystine knot motif in toxins and implications for drug design. *Toxicon* 39:43–60.
- Schmoltdt HU, Wentzel A, Becker S, Kolmar H (2005) A fusion protein system for the recombinant production of short disulfide bond rich cystine knot peptides using barnase as a purification handle. *Protein Expr Purif* 39:82–89.
- Wishart DS, Sykes BD, Richards FM (1992) The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry* 31:1647–1651.
- Daly NL, Craik DJ (2011) Bioactive cystine knot proteins. *Curr Opin Chem Biol* 15:362–368.
- Nicholson GM, Little MJ, Tyler M, Narahashi T (1996) Selective alteration of sodium channel gating by Australian funnel-web spider toxins. *Toxicon* 34:1443–1453.
- Lampe RA, Defeo PA, Davison MD, Young J, Herman JL, Spreen RC, Horn MB, Mangano TJ, Keith RA (1993) Isolation and pharmacological characterization of omega-grammotoxin SIA, a novel peptide inhibitor of neuronal voltage-sensitive calcium channel responses. *Mol Pharmacol* 44:451–460.
- Villegas E, Adachi-Akahane S, Bosmans F, Tytgat J, Nakajima T, Corzo G (2008) Biochemical characterization of cysteine-rich peptides from *Oxyopes* sp. venom that block calcium ion channels. *Toxicon* 52: 228–236.
- Branton WD, Kolton L, Jan YN, Jan LY (1987) Neurotoxins from Plectreurus spider venom are potent presynaptic blockers in *Drosophila*. *J Neurosci* 7:4195–4200.
- Klint JK, Senff S, Rupasinghe DB, Er SY, Herzig V, Nicholson GM, King GF (2012) Spider-venom peptides that target voltage-gated sodium channels: Pharmacological tools and potential therapeutic leads. *Toxicon* 60:478–491.
- Pallaghy PK, Alewood D, Alewood PF, Norton RS (1997) Solution structure of robustoxin, the lethal neurotoxin from the funnel-web spider *Atrax robustus*. *FEBS Lett* 419:191–196.
- Fletcher JI, Chapman BE, Mackay JP, Howden ME, King GF (1997) The structure of versutoxin (δ -atractoxin-Hv1) provides insights into the binding of site 3 neurotoxins to the voltage-gated sodium channel. *Structure* 5:1525–1535.
- Zhang L, Chou CP, Moo-Young M (2011) Disulfide bond formation and its impact on the biological activity and stability of recombinant therapeutic proteins produced by *Escherichia coli* expression system. *Biotechnol Adv* 29:923–929.
- Stapleton A, Blankenship DT, Ackermann L, Gorder GW, Manley GD, Palfreyman G, Coutant JE, Cardint AD (1990) Neurotoxic insecticidal polypeptides isolated from the funnel-web spider *Hololena curta*. *J Biol* 265: 2054–2059.
- Nguyen BD, Meng X, Donovan KJ, Shaka AJ (2007) SOGGY: solvent-optimized double gradient spectroscopy for water suppression. A comparison with some existing techniques. *J Magn Reson* 184:263–274.
- Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6:277–293.
- Keller R (2004) The computer aided resonance assignment tutorial. Goldau, Switzerland: Cantina Verlag, pp. 1–81.
- López-Méndez B, Güntert P (2006) Automated protein structure determination from NMR spectra. *J Am Chem Soc* 128:13112–13122.



28. Salomon-Ferrer R, Case DA, Walker RC (2013) An overview of the Amber biomolecular simulation package. *Wiley Interdiscip Rev Comput Mol Sci* 3:198–210.
29. Hernández-Salgado K, Estrada G, Olvera A, Coronas FI, Possani LD, Corzo G (2009) Heterologous expressed toxic and non-toxic peptide variants of toxin C_{ss}II are capable to produce neutralizing antibodies against the venom of the scorpion *Centruroides suffusus suffusus*. *Immunol Lett* 125:93–99.
30. Cartier GE, Yoshikami D, Gray WR, Luo S, Olivera BM, McIntosh JM (1996) A new α -conotoxin which targets α 3 β 2 nicotinic acetylcholine receptors. *J Biol Chem* 271:7522–7528.

