



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Maestría y Doctorado en Ciencias Bioquímicas

Instituto de Química

**BIOSÍNTESIS, PURIFICACIÓN Y CARACTERIZACIÓN ESTRUCTURAL DE
LAS TOXINAS ROXYTX1 Y ROXYTX2 DE LA ARAÑA OXYOPES LYENATUS**

TESIS

QUE PARA OPTAR POR EL GRADO DE:
Maestro en Ciencias

PRESENTA:

Elisa Norzagaray Villarreal

Dr. José Federico del Río Portilla
[Instituto de Química](#)

Dr. Mauricio Alberto Trujillo Roldán
[Instituto de Investigaciones Biomédicas](#)

Dr. Enrique García Hernández
[Instituto de Química](#)

Ciudad Universitaria, Ciudad de México. Junio, 2022



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

Agradecimientos

A los miembros del jurado por el tiempo y esfuerzo invertido durante la revisión del trabajo, y por todos sus comentarios y observaciones que sirvieron para enriquecer este proyecto de tesis: Dr. Lorenzo Patrick Segovia Forcella, Dr. Luis Fernando Olguín Contreras, Dr. Daniel Alejandro Fernández Velasco, Dra. Nuria Victoria Sánchez Puig, Dr. Alejandro Alagón Cano.

A mi comité tutor, quienes semestre tras semestre me guiaron y acompañaron en el desarrollo de esta investigación.

Al Dr. Mauricio Alberto Trujillo Roldán por facilitarme los recursos y el espacio necesarios para llevar a cabo el escalamiento del proceso de expresión de proteínas.

Al Dr. Enrique García Hernandez y sus alumnos Jesus Rauda y Fernando Cofias por su apoyo para la realización de los experimentos de dinámica molecular.

Al Dr. Federico del Rio Portilla por su apoyo incondicional, su guía y paciencia a lo largo de la realización de este trabajo.

Este proyecto fue realizado gracias al apoyo otorgado por el Consejo nacional de Ciencia y Tecnología (CONACyT) a través del programa de becas para estudios de posgrado, número de becario 925037.

Contenido

Índice de figuras	iv
Índice de tablas	vi
Acrónimos y abreviaturas	vii
1. Resumen	1
2. Marco Teórico	
2.1 Arañas, veneno y toxinas	2
2.2 Canales iónicos	3
2.3 Producción recombinante de proteínas	6
2.4 Resonancia magnética nuclear en el estudio de proteínas	8
2.5 Predicción de estructura de proteínas por métodos <i>in silico</i>	15
3. Antecedentes y Justificación del Proyecto	
2.1 Oxitoxinas	20
2.2 Justificación del Proyecto	21
4. Hipótesis	22
5. Objetivos	
5.1 Objetivo general	22
5.2 Objetivos particulares	22
6. Diseño Experimental y Metodología	
6.1 Expresión y purificación de las toxinas HisOxyTx1 y HisOxyTx2	22
6.2 Caracterización de las toxinas purificadas	25
6.3 Predicción de estructuras por experimentos <i>in silico</i>	26
7. Resultados y Discusión	
7.1 Optimización de expresión	27
7.2 Purificación de las toxinas a partir de fase soluble e insoluble	33
7.3 Escalamiento de expresión	40
7.4 Cuantificación de las toxinas	45
7.5 Espectrometría de masas MALDI-TOF	45
7.6 Dicroísmo Circular	46
7.7 Resonancia magnética nuclear	51
7.8 Predicción de estructura tridimensional por modelaje por homología	60
7.9 Refinación de modelos por dinámica molecular.	75
8. Conclusiones	86
9. Perspectivas	87
10. Referencias	88

Índice de Figuras

Figura 2.1.1 Representación esquemática del motivo estructural ICK	3
Figura 2.2.1 Estructura de la subunidad de canales iónicos de calcio	5
Figura 2.2.2 Arquitectura tridimensional propuesta de canales de Ca^{2+}	6
Figura 2.4.1 Diagrama esquemático de la alineación de los espines nucleares con el campo magnético	9
Figura 2.4.2 Representación gráfica de los pasos a seguir para la obtención de un espectro de ^1H	11
Figura 2.4.3 Espectros unidimensionales de protón	14
Figura 3.1.1 Secuencias de aminoácidos de OxyTx1 y OxyTx2	21
Figura 7.1.1 SDS-PAGE de la prueba de expresión de HisOxyTx2 y HixOxyTx2 en BL21	28
Figura 7.1.2 SDS-PAGE de la prueba de expresión de HisOxyTx1 y HisOxyTx2 en RosettaGami	30
Figura 7.1.3 Cinética de expresión de la toxina HisOxyTx1	31
Figura 7.1.4 Cinética de expresión de la toxina HisOxyTx2	32
Figura 7.2.1 Pruebas de purificación de HisOxyTx1 y HisOxyTx2	34
Figura 7.2.2 Cromatograma fase reversa de HisOxyTx1 purificado de fase soluble	35
Figura 7.2.3 Cromatograma fase reversa de HisOxyTx1 purificado de cuerpos de inclusión	36
Figura 7.2.4 Comparación de los cromatogramas de HisOxyTx1	37
Figura 7.2.5 Cromatograma fase reversa de HisOxyTx2 purificado de fase soluble	38
Figura 7.2.6 Cromatograma fase reversa de HisOxyTx2 purificado de cuerpos de inclusión	39
Figura 7.2.7 Comparación de los cromatogramas de HisOxyTx2	40
Figura 7.3.1 Cinética de crecimiento a 20°C de RG transformada con pQE30XaOxyTx1	42
Figura 7.3.2 Cinética de crecimiento a 20°C de RG transformada con pQE30XaOxyTx2	43
Figura 7.5.1 Espectro de masas MALDI-TOF de HisOxyTx1	45
Figura 7.5.2 Espectro de masas MALDI-TOF de HisOxyTx2	46
Figura 7.6.1 Espectro de dicroísmo circular reconstruido para HisOxyTx1	47
Figura 7.6.2 Espectro de dicroísmo circular reconstruido para HisOxyTx1	48
Figura 7.6.3 Espectros de dicroísmo circular esperados para proteínas con distintas estructuras secundarias predominantes	49
Figura 7.6.4 Comparación de las curvas de dicroísmo circular obtenidas para HisOxyTx1 y HisOxyTx2	51
Figura 7.7.1 Espectro unidimensional de ^1H en $\text{H}_2\text{O}/\text{D}_2\text{O}$ correspondiente a HisOxyTx1 purificada de la fase soluble.	52
Figura 7.7.2 Espectro unidimensional de ^1H en $\text{H}_2\text{O}/\text{D}_2\text{O}$ correspondiente a HisOxyTx2 purificada de la fase soluble.	53
Figura 7.7.3 Espectro unidimensional de ^1H en $\text{H}_2\text{O}/\text{D}_2\text{O}$ correspondiente a HisOxyTx2 purificada en presencia de agente caotrópico y posterior replegamiento	54
Figura 7.7.4 Comparación de los espectros de ^1H -RMN correspondientes a HisOxyTx2	55
Figura 7.7.5 Espectro unidimensional de ^1H correspondiente a HisOxyTx1 purificado de fase soluble obtenida a partir del proceso de escalamiento de expresión en $\text{H}_2\text{O}/\text{D}_2\text{O}$.	56
Figura 7.7.6 Espectro unidimensional de ^1H correspondiente a HisOxyTx1 en $\text{PBS}/\text{D}_2\text{O}$.	57
Figura. 7.7.7 Comparación de los espectros de HisOxyTx1 purificada a partir de fase soluble en $\text{H}_2\text{O}/\text{D}_2\text{O}$ y PBS	58
Figura 7.7.8 Espectro unidimensional de ^1H correspondiente a HisOxyTx2 en $\text{H}_2\text{O}/\text{D}_2\text{O}$.	59
Figura 7.7.9 Comparación de los espectros de ^1H -RMN correspondientes a HisOxyTx2	59
Figura 7.8.1 Modelo de OxyTx1 obtenido por I-Tasser.	62
Figura 7.8.2 Modelo de OxyTx1 obtenido por SwissModel.	65
Figura 7.8.3 Modelo de OxyTx1 obtenido por RaptorX.	67
Figura 7.8.4 Comparación visual de los tres modelos obtenidos para la estructura de OxyTx1.	69
Figura 7.8.5 Modelo de OxyTx2 obtenido por I-Tasser.	70
Figura 7.8.6 Modelo de OxyTx2 obtenido por SwissModel.	72
Figura 7.8.7 Modelo de OxyTx2 obtenido por RaptorX.	73

Figura 7.8.8 Comparación visual de los tres modelos obtenidos para la estructura de OxyTx2.	74
Figura 7.9.1 Modelo tridimensional de los centroides de los tres clústeres más poblados de la DM de OxyTx1SM	77
Figura 7.9.2 Valores de RMSD de la toxina a lo largo de la DM de OxyTx1SM.	78
Figura 7.9.3 Comparación de los valores de RMSD por regiones a lo largo de la DM	78
Figura 7.9.4 Valores de RMSF por residuo durante la DM de OxyTx1SM.	79
Figura 7.9.5 Fluctuación de los residuos a lo largo de la DM indicada por el valor de RMSF para cada residuo a lo largo de los 500 ns de la trayectoria de OxyTx1SM.	80
Figura. 7.9.6 Modelo tridimensional de los centroides de los tres clústeres más poblados de la DM de OxyTx2SM.	82
Figura 7.9.7 Valores de RMSD de la toxina a lo largo de la DM de OxyTx2SM.	83
Figura 7.9.8 Comparación de los valores de RMSD por regiones a lo largo de la DM	83
Figura 7.9.9 Valores de RMSF por residuo durante la DM de OxyTx2SM.	85
Figura 7.9.10 Fluctuación de los residuos a lo largo de la DM indicada por el valor de RMSF para cada residuo a lo largo de los 500 ns de la trayectoria de OxyTx2SM.	85

Índice de tablas

Tabla 7.3.1 Monitoreo de la expresión de HisOxyTx1 a 20°C en bioreactor de 10 L	41
Tabla 7.3.2 Monitoreo de la expresión de HisOxyTx2 a 20°C en bioreactor de 10 L	43
Tabla 7.4.1 Cuantificación gravimétrica de las toxinas obtenidas por diferentes metodologías.	45
Tabla 7.6.1 Distribución de estructura secundaria predicha por los resultados de DC.	50
Tabla 7.8.1 Alineación de secuencias de OxyTx1 y OtTx1a	60
Tabla 7.8.2 Alineación de secuencias de OxyTx2 y OtTx1a	61
Tabla 7.8.3 Evaluación de PDB 2N86 por MolProbity	61
Tabla 7.8.4 Evaluación de ITasser-OxyTx1 por MolProbity	63
Tabla 7.8.5 Conectividad de puentes disulfuro en ITasser-OxyTx1	64
Tabla 7.8.6 Evaluación de interacciones hidrofóbicas de ITasser-OxyTx1 por PIC	64
Tabla 7.8.7 Evaluación de SwissModel-OxyTx1 por MolProbity	65
Tabla 7.8.8 Enlaces disulfuro predichos por el servidor SwissModel para OxyTx1	66
Tabla 7.8.9 Evaluación de interacciones hidrofóbicas de SwissModel-OxyTx1 por PIC	66
Tabla 7.8.10 Evaluación de RaptorX-OxyTx1 por MolProbity	68
Tabla 7.8.11 Evaluación de interacciones hidrofóbicas de RaptorX-OxyTx1 por PIC	68
Tabla 7.8.12 Comparación de las secuencias de aminoácidos de OxyTx1 y OtTx1a	69
Tabla 7.8.13 Comparación de los valores de RMS entre los modelos de OxyTx1 obtenidos divididos por regiones de la toxina.	69
Tabla 7.8.14 Evaluación de ITasser-OxyTx2 por MolProbity	71
Tabla 7.8.15 Conectividad de puentes disulfuro en ITasser-OxyTx2	71
Tabla 7.8.16 Evaluación de interacciones hidrofóbicas de ITasser-OxyTx2 por PIC	71
Tabla 7.8.17 Evaluación de SwissModel-OxyTx2 por MolProbity	72
Tabla 7.8.18 Evaluación de interacciones hidrofóbicas de SwissModel-OxyTx2 por PIC	72
Tabla 7.8.19 Evaluación de RaptorX-OxyTx2 por MolProbity	73
Tabla 7.8.20 Evaluación de interacciones hidrofóbicas de RaptorX-OxyTx2 por PIC	74
Tabla 7.8.21 Comparación de los valores de RMS entre los modelos obtenidos de OxyTx2 divididos entre regiones de la toxina.	75
Tabla 7.9.1 Análisis de clústeres de la dinámica de OxyTx1SM.	76
Tabla 7.9.2 Comparación de las evaluaciones de MolProbity de OxyTx1SM antes y después del proceso de refinamiento por DM.	81
Tabla 7.9.3 Análisis de clústeres de la dinámica de OxyTx2SM.	81
Tabla 7.9.4 Comparación de las evaluaciones de MolProbity de OxyTx2SM antes y después del proceso de refinamiento por DM.	86

Acrónimos y Abreviaturas

ACN	Acetonitrilo
CAMQ	Cromatografía de afinidad metal-quelato
CD	Dicroísmo circular
COSY	Correlation Spectroscopy
DM	Dinámica molecular
DTT	Ditriotreitol
<i>E. coli</i>	<i>Escherichia coli</i>
FID	Free Induction Decay
FPLC	Fast protein liquid chromatography
GSH/GSSG	Glutación reducido/glutación oxidado
HPLC	High performance liquid chromatography
HT	High Tension
ICK	Inhibitory cysteine knot
MALDI-TOF	Matrix-assisted laser desorption-ionization-time of flight
NOESY	Nuclear Overhauser Effect Spectroscopy
ppm	Partes por millón
RMN	Resonancia magnética nuclear
RMSD	Root Mean Square Deviation
RMSF	Root Mean Square Fluctuation
SDS-PAGE	Sodium Dodecyl Sulfate-Polyacrylamide Gel
TA	Temperatura ambiente
TFA	Ácido trifluoroacético
TOCSY	Total Correlation Spectroscopy

1. Resumen

Las neurotoxinas obtenidas de venenos de animales como los alacranes, arañas, conos, serpientes, entre otros son péptidos de gran interés para el estudio de canales iónicos. Entre sus características de interés se incluyen una alta selectividad hacia sus blancos moleculares y una gran estabilidad, en especial cuando se trata de toxinas que presentan enlaces disulfuro.

Para estudiar la interacción de las neurotoxinas con sus blancos moleculares, en particular canales iónicos, es necesario elucidar la estructura tridimensional de estos péptidos. Con este fin, se busca obtener cantidades grandes de la toxina de interés en su forma pura. Esto representa un problema, ya que los venenos de animales son mezclas complejas de cientos o a veces miles de moléculas, entre las cuales las neurotoxinas representan un porcentaje sumamente bajo. Por esta razón se utilizan técnicas de biología molecular para obtener los péptidos de manera recombinante en sistemas heterólogos. Una vez que se ha obtenido una cantidad suficiente de la muestra de interés, se pueden realizar diversos experimentos para determinar la estructura de las toxinas.

Dos técnicas de gran importancia para la elucidación de estructuras protéicas son la difracción de rayos X y la resonancia magnética nuclear (RMN). La difracción de rayos X requiere la obtención previa de cristales a partir de la proteína de interés. Las neurotoxinas son péptidos pequeños (por lo general contienen menos de 100 aminoácidos), y por lo tanto a menudo es imposible obtenerlos en forma cristalizada. Por esta razón, se utiliza la RMN para estudiarlos en disolución.

En este trabajo se describe la obtención mediante un sistema heterólogo de las toxinas rOxyTx1 y rOxyTx2, previamente aisladas del veneno de la araña *Oxyopes lyneatus*. Estas toxinas presentan actividad sobre canales de calcio de insectos. Con el fin de determinar su estructura, se utilizaron RMN y técnicas *in silico* (modelado por homología con posterior refinamiento por dinámica molecular). Los resultados indican que estas toxinas presentan un motivo estructural denominado nudo inhibitorio de cisteínas (*inhibitory cysteine knot*, o ICK), el cual les confiere una alta estabilidad. También se pudo observar que la presencia de cinco

enlaces disulfuro en cada toxina hace que el motivo sea conservado y su movilidad es baja al simularlas en un sistema en disolución.

Los resultados aquí reportados sirven como una propuesta para estas toxinas, de las cuales no se cuenta con estudios estructurales previos. A partir de estos se pueden realizar más experimentos de RMN para obtener una propuesta experimental de las estructuras de interés.

2. Marco teórico

2.1. Arañas, veneno y toxinas

Las arañas son los animales venenosos con más especies conocidas, con >46,000 especies descritas hasta ahora. Como su supervivencia depende completamente en la depredación, han desarrollado una combinación compleja de enzimas, neurotoxinas y péptidos antimicrobianos y citolíticos en sus venenos¹. Existe un gran interés por estudiar los componentes de sus venenos, ya que éstos presentan actividad sobre un gran número de especies animales, entre ellos insectos y mamíferos, de manera potente y selectiva. Se ha propuesto su uso en tratamientos farmacológicos, bioinsecticidas, entre otros.²

El componente principal de la mayoría de los venenos de arañas son toxinas peptídicas ricas en enlaces disulfuro, aunque existen excepciones. Estas toxinas presentan propiedades farmacológicas variadas; se ha observado que muestran actividad en un gran número de canales iónicos neuronales de manera selectiva y potente.³ Se cree que muchas de estas toxinas provienen de un gen ancestral que codifica una proteína tipo “nudo” (*knottin* en inglés) rica en enlaces disulfuro². Los nudos son péptidos pequeños (por lo general 30-50 residuos) que conforman un motivo estructural conocido como nudo inhibitorio de cisteínas (ICK por sus siglas en inglés). El motivo ICK se conforma por una hoja- β de dos o tres láminas- β estabilizadas por tres enlaces disulfuro que forman un pseudo-nudo en el que dos de los enlaces disulfuro y las secciones que intervienen de la cadena peptídica forman un lazo cerrado que es intersectado por el tercer enlace disulfuro.⁴ Una representación esquemática del motivo ICK se puede observar en la figura 2.1.1. Se debe notar que el motivo no cumple con la definición matemática de nudo, ya que se puede deshacer por medio de una transformación geométrica, sin la necesidad de romper

alguno de los enlaces. El motivo ICK permite a los péptidos tipo nudo presentar una alta estabilidad química y térmica, así como una buena resistencia a proteasas. Se han descrito toxinas de araña con este motivo estructural que se mantienen estables por largos períodos (>1 semana) en disolución y cuya temperatura de fusión es mayor a 80°C². Otra característica del motivo ICK es que constriñe fuertemente la estructura del péptido, de manera que la mayor parte de los residuos que no son cisteína pueden ser mutados sin comprometer la integridad estructural de la toxina. Esto facilita la adaptación evolutiva de las toxinas, así como la ingeniería biotecnológica en búsqueda de actividades novedosas.

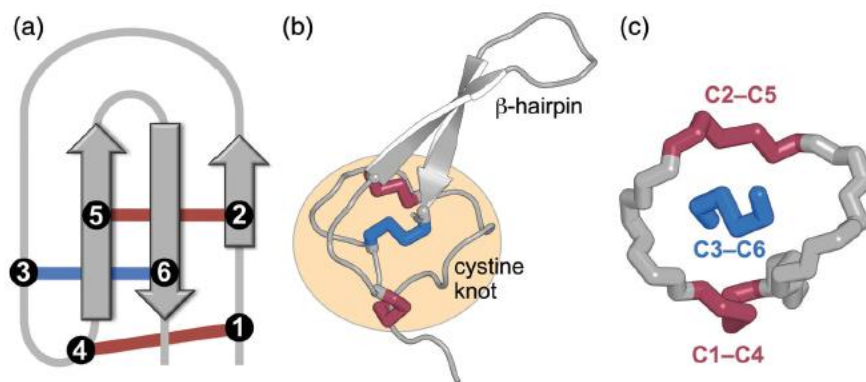


Fig. 2.1.1 (A) Representación esquemática del motivo estructural ICK. (B) Representación tridimensional de la toxina Hv1a, en la que el motivo estructural ICK se muestra en la región marcada en naranja. (C) Figura simplificada de esta misma región ampliada, donde los enlaces disulfuro que forman el lazo se muestran en rojo y el enlace que atraviesa para formar el pseudo nudo se muestra en azul. Tomada de la referencia 2.

2.2. Canales iónicos

Los canales iónicos son proteínas transmembranales que permiten el paso de moléculas inorgánicas (por lo general iones Na⁺, K⁺, Ca²⁺ y Cl⁻) que no son capaces de penetrar la bicapa lipídica por su carga y su estado de hidratación. Estos canales forman poros pequeños y altamente selectivos que se abren y cierran rápidamente. La apertura del canal depende de un estímulo, el cual puede ser un cambio en el potencial de la membrana (*voltage gated*), la presencia de un ligando o un estrés mecánico.⁵ Las toxinas de araña que se estudiarán en este proyecto presentan actividad bloqueadora sobre canales de Ca²⁺ mediados por potencial (Ca_v).

Estos canales son un blanco terapéutico de interés, ya que tienen un rol importante en procesos de proliferación celular, secreción hormonal, expresión genética, regulación de actividad enzimática, entre otros.⁶

Los canales Ca_v se dividen en subtipos según sus características electrofisiológicas. Se han descrito seis categorías principales: L (*long-lasting*, de larga duración), T (*transient*, transitorios), N (no corresponde a L o T, o neuronal), P (células Purkinje), Q (después de P) y R (remanentes o resilientes).⁷

Los canales L se encuentran en músculo liso, esquelético y cardíaco; sus corrientes se distinguen por un alto potencial de activación, alta conductancia en un solo canal, inactivación lenta mediada por voltaje y regulación por vías de fosforilación de proteínas dependientes de cAMP. Estas corrientes son inhibidas por fármacos antagonistas a Ca^{2+} , incluyendo dihidropiridinas, fenilalquilaminas y benzodiazepinas. Este tipo de canales realizan funciones como regulación de la transcripción, secreción endócrina, regulación de actividad enzimática, regulación del ritmo cardíaco, entre otras.⁸

Los canales tipo T se observan en muchas de las mismas células que los tipo L; sin embargo, las corrientes observadas en éstos se activan a potenciales mucho más negativos; se inactivan rápidamente; se desactivan lentamente; tienen baja conductancia de un sólo canal y no son inhibidas por antagonistas de Ca^{2+} convencionales. Este tipo de canales está involucrado principalmente en la regulación del ritmo cardíaco y la tasa de disparo neuronal.⁹

Los canales tipo N se definieron originalmente por su dependencia de voltaje y velocidad de inactivación intermedia; más negativo y rápido que los de tipo L pero más positivo y lento que los de tipo T. No son bloqueados por moléculas orgánicas como las que bloquean al tipo L, y son sensibles al péptido ω -conotoxina-GVIA (del veneno del cono *Conus geographus*) y péptidos similares. Debido a que la dependencia de voltaje y propiedades cinéticas de este tipo de canales varían ampliamente, la sensibilidad a este tipo de péptidos es su característica determinante.¹⁰

Los canales tipo P, Q y R se caracterizan principalmente por el tipo de moléculas que bloquean sus corrientes. Las de tipo P y Q son sensibles a la toxina ω -agatoxina-IVA con diferentes niveles de afinidad (la mayor afinidad la presentan los tipo P). Los canales tipo R son sensibles a la mayor parte de los bloqueadores de canales de calcio, incluyendo moléculas orgánicas y péptidos; este tipo de canales se han encontrado en neuronas granulares del cerebelo. Estos últimos tres tipos de canales están involucrados principalmente en la liberación de neurotransmisores.¹⁰

Los canales de calcio son complejos de unidades funcionales denominadas $\alpha 1$, $\alpha 2$, β , γ y δ . El modelo propuesto por Wolf et al¹¹ consiste en una unidad funcional transmembranal principal $\alpha 1$ de 190 kDa asociada con un dímero $\alpha 2\delta$ ligado por puentes disulfuro de 170 kDa, una unidad funcional transmembranal β de 55 kDa y una unidad funcional transmembranal γ de 33 kDa.¹⁰ Una representación gráfica de este modelo se observa en las figuras 2.2.1 y 2.2.2.

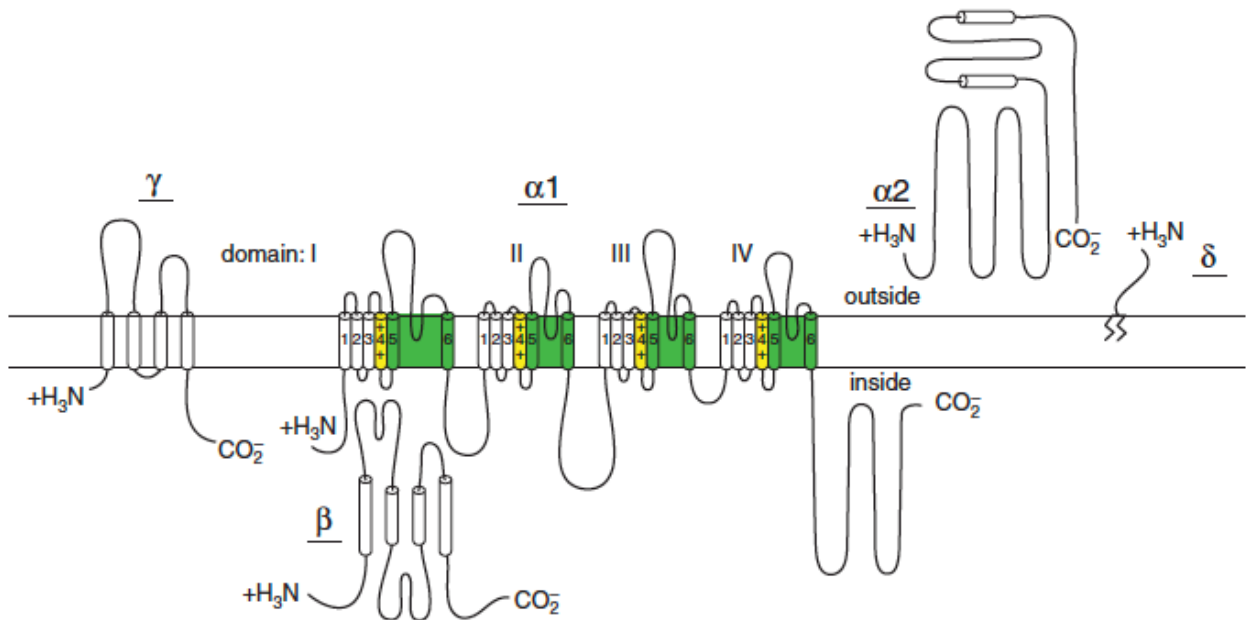


FIG 2.2.1. Estructura de la subunidad de canales iónicos de calcio. Las subunidades se representan como modelos transmembranales de plegamientos donde los cilindros corresponden a hélices α y el largo de las líneas corresponde aproximadamente al largo de los segmentos. Figura tomada de la referencia 10

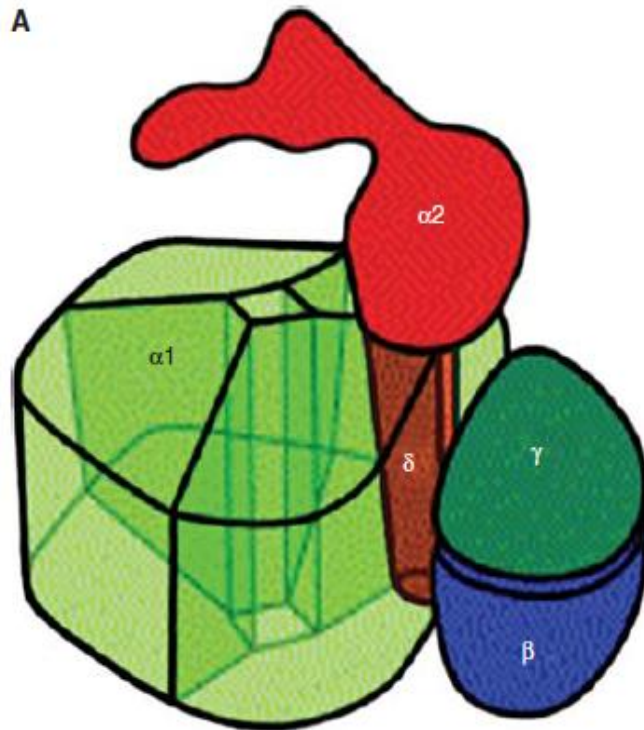


FIG 2.2.2. Arquitectura tridimensional propuesta de canales de Ca^{2+} , basada en criomicroscopía electrónica de $\text{Ca}_v1.1$. Figura tomada de la referencia 10

Es importante notar que la mayor parte de los estudios sobre canales de calcio que se encuentran en la literatura se han realizado sobre canales de mamíferos. Sin embargo, existen también canales de calcio en otras especies; en particular, los canales de insectos son relevantes para este trabajo. El modelo estructural en general se conserva entre canales de mamífero y de insecto, aunque existen algunas diferencias, en particular en los dominios citoplásmicos de las regiones amino- y carboxilo-terminal, las cuales son más largas en el caso de canales de insectos. Además, se reportan diferencias en la sensibilidad a bloqueadores orgánicos (por ejemplo, los canales tipo L de insecto no son sensibles a dihidropiridinas). La clasificación de los tipos de canales se conserva en el caso de canales de insecto.¹²

2.3. Producción recombinante de proteínas

Los estudios estructurales de proteínas requieren de cantidades considerables de la proteína de interés en disolución; muchas veces estas cantidades son mayores a las que se pueden obtener fácilmente a partir de la fuente natural. Este problema está presente en el estudio de toxinas animales, ya que los venenos están constituidos por cientos de compuestos y por lo general las proteínas de interés representa un porcentaje muy bajo de éstos. Por esta razón, sería necesario obtener muestras de veneno de decenas o en ocasiones cientos de especímenes para acumular cantidades suficientes de muestras para realizar los experimentos.¹³

Una manera de solucionar este problema es utilizar sistemas de expresión bacteriana para producir proteínas recombinantes. Esta es una técnica que consiste en utilizar ingeniería genética para introducir un gen de interés en un organismo heterólogo y así producir cantidades elevadas de la proteína de interés¹⁴. El organismo más utilizado para estos sistemas es *Escherichia coli*, ya que presenta numerosas ventajas tales como un ciclo de vida corto, fácil manipulación genética y un genoma y biología celular bien estudiados.¹⁵

Los vectores utilizados para la transformación de la célula huésped contienen genes que facilitan la producción de proteínas recombinantes, tales como un promotor a partir del cual se expresa el gen de interés y un origen de replicación que regular el número de copias del plásmido en cada célula huésped.¹⁵

Una vez producida la proteína, se debe llevar a cabo el proceso de purificación. Otra ventaja de los sistemas de expresión recombinante de proteínas es permitir la introducción de secuencias al péptido que facilitan su purificación. Un método común es la adición de una secuencia de histidinas a la construcción de la proteína. Este residuo es capaz de formar complejos (quelatos) con iones metálicos como Ni (II) o Cu (II), por lo que se puede utilizar la técnica de cromatografía de afinidad metal-quelato (CAMQ).¹⁶

Un problema común al realizar la producción recombinante de proteínas es la obtención de agregados insolubles. Cuando este es el caso, existen métodos de purificación de proteínas a partir de cuerpos de inclusión, que consisten en utilizar agentes caotrópicos que permitan la solubilización de la proteína en distintas conformaciones estructurales. Una vez disuelta la proteína, se utilizan agentes como pares redox GSH/GSSH, DTT/GSSH, cisteína/cistina y cisteamina/cistamina¹⁷ que favorecen el repliegamiento controlado. Otra opción cuando se trabaja una proteína que tiende a formar agregados insolubles es modificar la metodología de expresión, ya sea disminuyendo la temperatura, concentración del cultivo, utilizando cepas bacterianas que produzcan ambientes oxidantes, entre otros.¹⁸

Una vez que se ha logrado obtener la proteína de manera soluble, se debe purificar utilizando HPLC o FPLC para asegurar un nivel de pureza alto. La proteína pura se debe caracterizar por métodos espectroscópicos y funcionales (e.g. actividad) para corroborar su identidad y correcto plegamiento antes de realizar experimentos biológicos o estructurales¹⁷

2.4. Resonancia magnética nuclear en el estudio de proteínas

2.4.1 ¿Qué es la RMN?

La RMN es una técnica espectroscópica utilizada para la elucidación de estructuras moleculares. El principio de la RMN se basa en la detección del espín nuclear en campos magnéticos localizados alrededor de núcleos atómicos. El espín nuclear (I) representa el momento angular total de espín. Si se coloca la muestra en presencia de un campo magnético, los núcleos se acomodan en orientaciones específicas. El número de orientaciones que un núcleo puede adoptar es $2I + 1$.^{19,20}

Todos los núcleos que tienen un número atómico o número de masa impar presentan espín nuclear, y por lo tanto se comportan como objetos magnéticos. Esto se debe a que los núcleos no son esferas perfectas, sino esferoides oblatas. Cuando se encuentran en un campo magnético, estos núcleos absorben radiación en la región del espectro de radiofrecuencias. La frecuencia exacta (conocida como frecuencia de resonancia) varía y depende de diversos

factores. El más importante de estos es el entorno químico de cada núcleo; es por esta razón que en una molécula, cada núcleo con espín genera una señal característica. Es así como la técnica de RMN permite "evidenciar" cada uno de los átomos que conforman a la molécula, y de esta manera determinar su estructura tridimensional. Los átomos que tienen número de masa impar pueden tener valores de I de $1/2$, $3/2$, $5/2$, ..., etc. En RMN, los núcleos más comúnmente utilizados son ^1H y ^{13}C , seguidos por ^{19}F , ^{31}P , ^{15}N y ^{29}Si , todos ellos con un valor de $I=1/2$. Por lo tanto, se ejemplificará sólo este caso por simplicidad^{19,21,22}

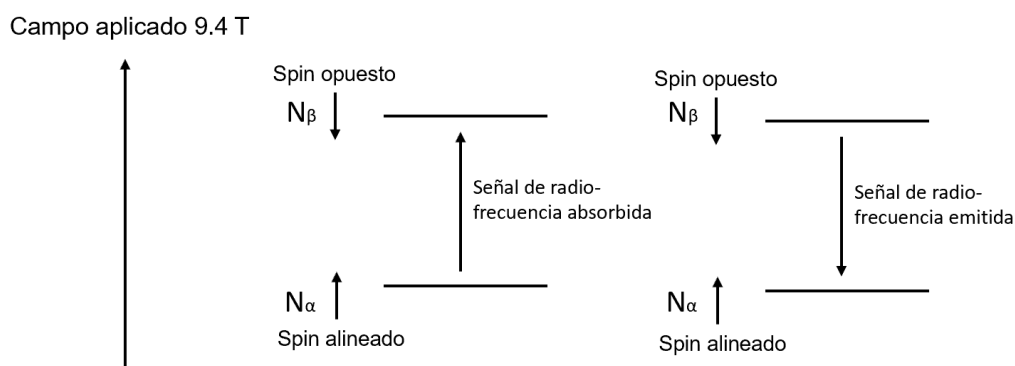


FIG. 2.4.1. Diagrama esquemático de la alineación de los espines nucleares con el campo magnético.

Un núcleo con un espín nuclear de $1/2$ da lugar a dos orientaciones al interactuar con un campo magnético externo: una orientación de baja energía (núcleos alineados con el campo) y una de alta energía (núcleos opuestos al campo), como se ilustra en la figura 2.4.1. Las transiciones energéticas entre estos dos niveles se describen con la ecuación:

$$\Delta E = \frac{h\gamma B_0}{2\pi}$$

Donde γ es la constante giromagnética del núcleo a observar y B_0 es la fuerza del campo magnético externo.

Para describir la frecuencia ν correspondiente a la diferencia energética se utiliza la ecuación:

$$\nu = \frac{\gamma B_0}{2\pi}$$

La distribución de Boltzmann del sistema se describe por la ecuación:

$$\frac{N_\beta}{N_\alpha} = e^{\frac{\Delta E}{kT}}$$

Donde N_α es el número de núcleos en el estado de baja energía, y N_β es el número de núcleos en el estado de alta energía.

Si al sistema se le aplica una radiofrecuencia del mismo valor que ν , ocurre una transición de núcleos del estado de baja energía al de alta energía, y la distribución de Boltzmann cambia. Este valor se denomina frecuencia de resonancia o frecuencia de Larmor, y es específico para cada núcleo. La frecuencia de Larmor depende de la naturaleza del núcleo y de la fuerza del campo magnético aplicado.²²

Cuando la muestra se coloca en el espectrómetro, los núcleos magnéticos alcanzan una distribución de equilibrio, con un mayor número de átomos en el nivel correspondiente a N_α que en N_β . Esto genera un vector de magnetización M_0 en dirección del campo aplicado. Cuando se aplica un pulso de frecuencia ν , el vector de magnetización gira en un ángulo θ . El vector de magnetización subsecuentemente regresa a la posición de equilibrio girando sobre el eje z . El equipo de resonancia detecta los vectores magnéticos a lo largo de los ejes x y y ; estas señales se denominan *Free Induction Decay* (FID). Cada pulso aplicado a la muestra genera un FID, y se utilizan numerosos pulsos en cada experimento de RMN. Este proceso se ilustra en la figura 2.4.2. El resultado del espectro proviene de la suma de todos los FID generados, a través de una Transformada de Fourier.²³

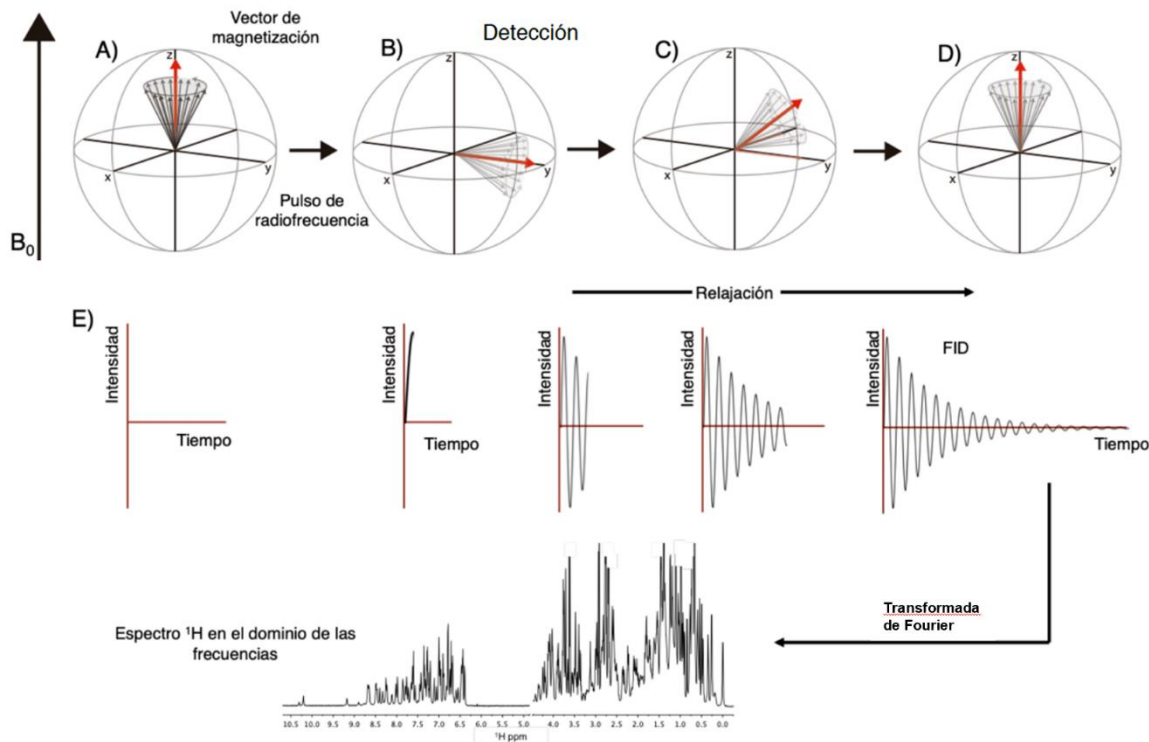


FIG. 2.4.2. Representación gráfica de los pasos a seguir para la obtención de un espectro de ^1H .

- A) Alineación del vector de magnetización con el campo electromagnético.
 B) Reacomodo del vector de magnetización luego de girar sobre un ángulo θ gracias al pulso aplicado
 C) Relajación del vector de magnetización
 D) Reequilibrio del vector de magnetización a la posición inicial.
 E) Detección de la señal a lo largo del proceso de relajación del vector de magnetización para generar la señal FID

En un experimento unidimensional, la señal obtenida directamente del equipo es función del tiempo t . En el caso de un experimento multidimensional, el espectro es una función de dos tiempos variables, la señal se obtiene en función de tiempos t_n , y se debe aplicar la transformada de Fourier n veces para obtener un espectro en n dimensiones. En el caso de un experimento en dos dimensiones, la muestra se excita por uno o varios pulsos de radiofrecuencia, permitiendo que la magnetización evolucione durante el tiempo t_1 . Posteriormente, se vuelve a aplicar uno o varios pulsos para obtener la señal como función del tiempo t_2 . El conjunto de estos eventos conforma una función de dos tiempos variables; la variación en la secuencia de pulsos determina la información obtenida del sistema.²²

2.4.2 ¿Por qué utilizar RMN?

A lo largo de las primeras décadas del siglo XX, la elucidación de estructuras de productos naturales era un proceso extremadamente laborioso que involucraba numerosas reacciones sintéticas y degradativas. Los avances que se han logrado en técnicas espectroscópicas como espectrometría de masas, cristalografía de rayos X, y resonancia magnética nuclear (RMN) han permitido simplificar enormemente esta tarea.

En particular en el estudio de proteínas, la RMN presenta ventajas sobre la cristalografía para el estudio de proteínas pequeñas (<30 kDa) en disolución, ya que estas proteínas por lo general no permiten la generación de cristales. Además, utilizando RMN es posible determinar relaciones proteína-función, distinguir entre diferentes conformaciones, encontrar sitios específicos de interacción, entre otros.²² La preparación de muestras para su estudio por RMN es sencilla, y la técnica no es destructiva, por lo que la muestra se puede recuperar y utilizar para realizar distintos experimentos si es necesario.²⁴

2.4.3 ¿Cómo se determina la estructura de una proteína por RMN?

Existen diferentes tipos de experimentos de RMN, cada uno con características particulares que permiten desarrollar diversas estrategias para la determinación estructural de las moléculas de interés. Por lo general se utilizan tres tipos de experimentos de RMN, cada uno para un propósito diferente:

1. Experimentos que detectan todos los núcleos presentes en un sistema de espín (y por lo tanto todo el sistema de espín).
2. Experimentos que detectan núcleos vecinos secuenciales dentro de un sistema de espín (es decir, átomos unidos por enlaces covalentes).
3. Experimentos que detectan núcleos cercanos en el espacio (no necesariamente enlazados).

En el caso del estudio de macromoléculas es esencial la distinción entre el segundo y tercer caso; esto sólo es posible mediante la utilización de experimentos multidimensionales.²⁵

El primer paso para el análisis de la muestra de interés consiste en un experimento unidimensional de ^1H . Este experimento permite conocer el estado de plegamiento de la proteína. En los espectros de ^1H -RMN se pueden observar picos en regiones características que corresponden a distintos tipos de protones que conforman a los aminoácidos. Las señales que se encuentran a valores de desplazamiento entre 8.5 y 11 ppm corresponden a los protones de amidas en la cadena; si estas señales se encuentran dispersas, se concluye que la proteína se encuentra plegada. En cambio, la existencia de señales intensas y con poca dispersión alrededor de 8.3 ppm (la región característica de amidas en motivos “random coil”) es indicador de una proteína desplegada. Un ejemplo de un espectro unidimensional que corresponde a una proteína plegada y uno que corresponde a una proteína desplegada se observan en la figura 2.4.3. Otro indicador del estado de plegamiento de la proteína es el experimento bidimensional TOCSY, en el cual se esperaría la presencia de tantas señales en el espectro como sistemas de espín (aminoácidos) presenta una proteína plegada.²⁶

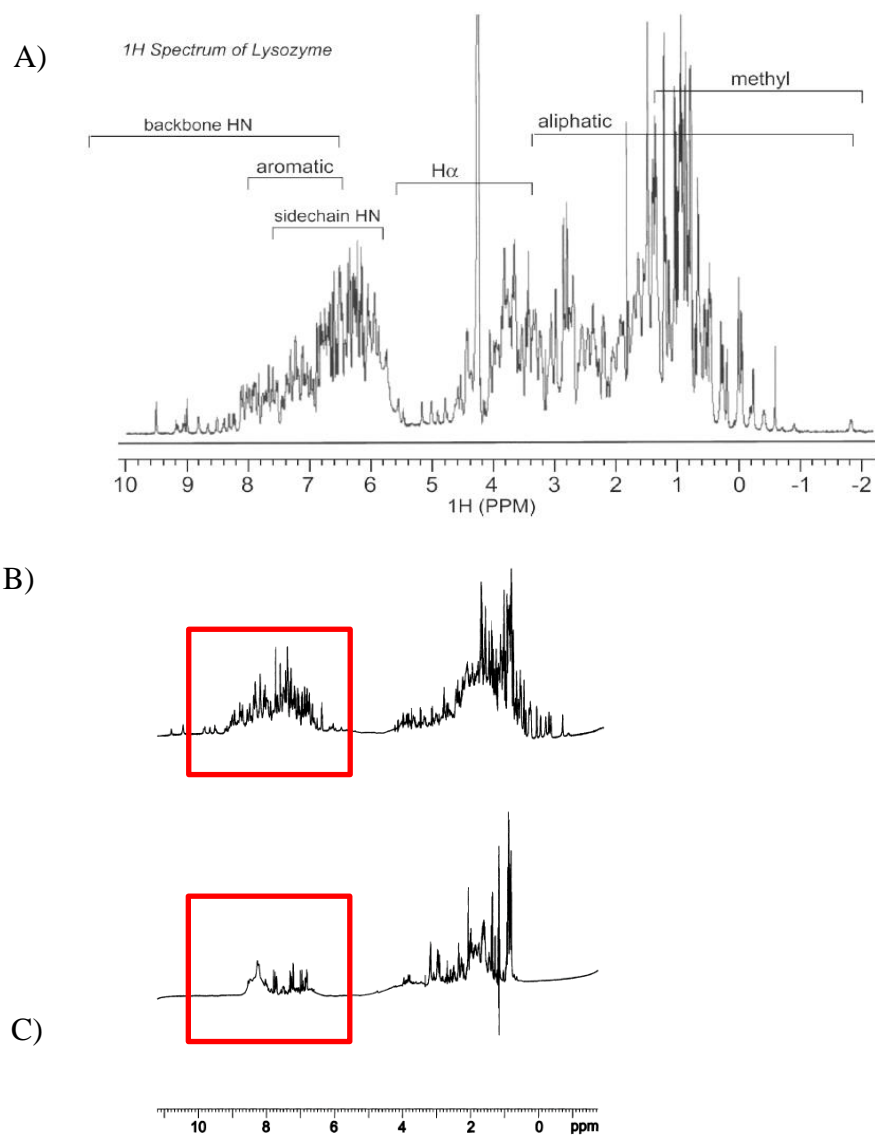


FIG 2.4.3. Espectros unidimensionales de protón. A) Espectro con la región correspondiente a cada fragmento de la proteína marcado. B) Espectro de una proteína plegada con la zona correspondiente a protones aromáticos marcada en rojo. C) Espectro de una proteína desplegada con la región correspondiente a protones aromáticos marcada en rojo. Figura tomada de la referencia 23

Una vez que se determinó que la proteína se encuentra plegada, se procede a la asignación de señales en el experimento $^1\text{H}, ^1\text{H}$ -COSY, donde cada señal representa la correlación entre núcleos que se encuentran a no más de tres enlaces de distancia. Cada sistema de espín tiene una topología característica en un espectro $^1\text{H}, ^1\text{H}$ -COSY y puede ser asociado con un tipo de residuo específico. Siete topologías pueden ser asociadas únicamente con un aminoácido; los trece aminoácidos restantes comparten tres topologías diferentes y por lo tanto se requiere la secuencia de aminoácidos de la proteína que se estudia para realizar su asignación.

Posteriormente se utiliza el experimento bidimensional $^1\text{H}, ^1\text{H}$ -NOESY, el cual muestra una señal por cada par de núcleos con una distancia internuclear menor a 5 Å. Si se conoce la secuencia de la proteína, se puede determinar cuáles son los átomos que se relacionan con núcleos “arbitrarios” (i.e. que no corresponden al sistema de espín). Esta información permite entonces establecer restricciones a la conformación de la molécula, permitiendo calcular su estructura. Cada una de estas restricciones representa un par de átomos con una distancia internuclear menor a 5 Å. El volumen de las señales del NOESY se utilizan para calcular los límites de distancia de las restricciones (entre 2 a 5 Å). Con esta información un programa de cálculo estructural (e.g. CYANA) es capaz de simular plegamientos de la estructura primaria de la proteína y encontrar aquellos que cumplan con las restricciones de forma óptima.^{25,27}

2.5. Predicción de estructura de proteínas por métodos *in silico*.

2.5.1 Métodos de modelaje comparativo

Los métodos computacionales de predicción de estructura de proteínas se dividen en dos categorías: aquellas basadas en plantillas, y los métodos “*template-free*” o sin plantillas. Algunos protocolos combinan aspectos de ambas categorías. Este trabajo utilizará sólo aquellos métodos basados en plantillas. Éstos se basan en métodos de modelaje comparativo, los cuales parten de una secuencia de aminoácidos e identifican plantillas con relación evolutiva por comparación de secuencias o perfiles (e.g. BLAST y HHpred). Una vez identificada la plantilla o plantillas adecuada, se genera una propuesta de estructura para la proteína problema.²⁸ Estos métodos se

basan en la observación de que las estructuras proteicas conocidas comprenden un número limitado de plegamientos estables; por esta razón, se pueden compartir elementos estructurales en proteínas evolutivamente distantes.²⁹

2.5.2 Homología, similitud e identidad.

La similitud es una propiedad cuantitativa entre dos proteínas que determina el nivel de correspondencia entre las dos secuencias al nivel de aminoácidos. Una alineación de secuencias muestra si las secuencias son idénticas, similares o diferentes y cuantifica el número de aminoácidos correspondientes.

El porcentaje de identidad de una secuencia con respecto a otra determina el número de aminoácidos idénticos que coinciden exactamente entre ambas, sin tomar en cuenta espacios. El resultado es reportado con respecto a la secuencia de menor longitud. Esta propiedad no es transitiva, es decir, $A=B$ y $B=C$ no implica que $A=C$.

La similitud es una medida de la relación entre dos secuencias. Toma en cuenta el porcentaje de identidad, así como la sustitución de aminoácidos con propiedades similares entre secuencias.

En estudios comparativos, se utiliza el término “homología”. Homología indica la presencia de un ancestro común entre las proteínas a estudiar y se refiere a sus características estructurales. En anatomía, se define homología como “el mismo órgano bajo toda variedad de forma y función”³⁰; en el estudio de proteínas se mantiene el significado de tener un origen evolutivo en común y se usa para clasificarlas taxonómicamente por su distribución espacial. El modelado de proteínas por homología genera estructuras tridimensionales a partir de secuencias bidimensionales bajo el entendido de que la similitud en 3D refleja la similitud en 2D. Sin embargo, es importante notar que la homología entre proteínas no implica similitud en la secuencia como condición necesaria.³¹

2.5.3 Principios de dinámica molecular

Los experimentos de dinámica molecular (DM) consisten en un conjunto de técnicas de simulación molecular que se utilizan para entender el movimiento físico y los ensamblados de átomos para muchos sistemas biológicos de interés en términos de su estructura a nivel microscópico a lo largo de un periodo de tiempo determinado. Asimismo, este tipo de experimentos son capaces de realizar seguimientos de procesos que ocurren en menos de un milisegundo.³²

Las simulaciones de DM son utilizadas para diversos propósitos, tales como la exploración dinámica de la superficie de las proteínas, estudios de complejos macromoleculares (proteína-proteína, proteína-DNA, proteína-ligando), y sistemas lipídicos. Las dinámicas moleculares sirven para resolver cuestiones como la termodinámica de complejos proteína-ligando, plegamiento de proteínas, cálculos de energía libre de moléculas, simulaciones de proteínas en la membrana, transporte de iones, etc.

Al realizar una DM, el objetivo es imitar el comportamiento natural de los átomos asumiendo una función de energía potencial específica. Esta función nos permite calcular la fuerza experimentada por cualquier átomo según las posiciones de los otros átomos. Para realizar estos experimentos, el tiempo se divide entre pasos discretos del orden de femtosegundos (10^{-15}). En cada paso se calculan las fuerzas que actúan sobre cada átomo utilizando un campo de fuerza. En una DM, así como en la vida real, los átomos se encuentran en constante movimiento y no llegan a un estado permanente de mínima energía. Sin embargo, si se observa durante un tiempo suficientemente largo, la simulación se aproxima a la distribución de Boltzmann.^{33,34}

Durante estos experimentos se debe tomar en cuenta la energía del sistema y la influencia del disolvente. Aunque se esperaría que la energía se conservara al ser un sistema cerrado, lo que sucede en la práctica es que la energía total del sistema suele aumentar debido a errores de redondeo. Por lo tanto, se utilizan mecanismos (“termostatos”)³⁵ que permiten mantener la temperatura constante. Para el caso del disolvente, se pueden utilizar solventes explícitos (representar moléculas de agua, iones, lípidos, etc.) o implícitos (utilizar modelos matemáticos para simular los efectos promedio del disolvente).³⁶ Los disolventes explícitos hacen que el experimento sea computacionalmente demandante, pero generan resultados más cercanos a la

realidad, mientras que los disolventes implícitos generan resultados menos confiables en un menor tiempo.³⁷

Existen ciertas desventajas en los experimentos de DM. Una de ellas es la escala de tiempo que se puede utilizar. Muchos procesos de interés que se busca estudiar por estos métodos ocurren en el orden de microsegundos o milisegundos, o incluso más tiempo. Sin embargo, un experimento del orden de microsegundos puede ser extremadamente largo y demandante computacionalmente, y por lo tanto no es factible si no se tiene acceso a un equipo con las características adecuadas. Otra desventaja es la precisión de los campos de fuerza utilizados. Aunque se han logrado avances significativos en la calidad de los campos de fuerza disponibles, incluso aquellos con excelentes evaluaciones son imperfectos. Por lo tanto, los resultados obtenidos con estas herramientas son inherentemente aproximaciones.³⁸

Otro aspecto que se debe considerar es la imposibilidad de formar o romper enlaces covalentes durante una DM estándar. Si se tiene una propuesta inicial de estructura para la proteína a estudiar, esto no suele ser un problema ya que los enlaces covalentes no se suelen romper o crear en una proteína que se ha terminado de formar. Sin embargo, si se estudian proteínas con enlaces disulfuro o residuos ácidos o básicos que jueguen un papel importante en la interacción de interés, esto puede representar una limitante.³²

2.5.4 Refinamiento de estructuras por dinámica molecular

Los modelos estructurales de proteínas obtenidos por homología son a menudo propuestas de baja resolución; es común encontrar errores de 3 Å o más con respecto a la estructura nativa.³⁹ Por esta razón, existen distintos protocolos de refinamiento de estructura basados en principios físicos o familiaridad con el tipo de proteína. Uno de los métodos que da mejores resultados es el uso de experimentos de dinámica molecular que permiten observar la proteína en un ambiente que se aproxima al nativo y por lo tanto favorece el acomodo correcto de la cadena de aminoácidos.

Se debe comenzar con una estructura modelada por homología que sea evaluada positivamente como cercana a lo esperado en la estructura nativa. Existen diversos protocolos de refinamiento, y se han reportado experimentos con una amplia variedad de tiempo de experimento (desde 100 ns hasta 100 μ s).

Los pasos por seguir pueden variar al refinar una estructura, pero en general siguen esta secuencia:

- Minimización de la estructura
- Equilibrado del sistema
- Generación de trayectorias por simulaciones de DM
- Calcular la estructura “promedio” a partir de las trayectorias generadas.
- Análisis de las trayectorias, en especial valores de RMSD y valores de energía en función del tiempo.^{39,40}

3. Antecedentes y justificación del Proyecto

3.1. Oxitoxinas

Las oxitoxinas (OxyTxS) son péptidos ricos en enlaces disulfuro aislados del veneno de la araña *Oxyopes lineatus* que actúan como bloqueadores de canales de calcio dependientes de voltaje. OxyTx1 se encuentra en mayor proporción en el veneno de la araña y consiste de 69 aminoácidos, con un peso molecular experimental de 8059.1 Da. OxyTx2 es la segunda principal neurotoxina encontrada en el veneno de *O. lineatus*. Consiste de 55 aminoácidos y su peso molecular experimental es de 6175.2 Da (Fig 3.1.1). Se puede asumir que ambas toxinas presentan amidación en el C-terminal y forman 5 enlaces disulfuro en su forma nativa por la discrepancia entre sus pesos moleculares experimentales y teóricos (8069.1 Da y 6186.1 Da en sus formas reducidas, respectivamente).⁴¹

Al probarse la actividad de las oxitoxinas en canales Ca_v expresados *in vitro*, se observó una inhibición de la corriente de Ca^{2+} en canales tipo P/Q hasta 53%. Además, OxyTx1 produjo un bloqueo reversible de corrientes de Ca^{2+} en canales tipo N y L (65% y 63%, respectivamente). Se observó además que la actividad bloqueadora de OxyTx1 no puede ser reversible por pulsos despolarizantes, como es el caso de ω -AgaIVA⁴². Este resultado es relevante ya que, a pesar de tener menor afinidad por los canales iónicos en comparación con ω -AgaIVA, las oxitoxinas poseen mejores propiedades paráliticas. Además, la actividad de estas toxinas es específica en canales Ca_v de insectos, ya que no bloquean canales Na_v y no presentan toxicidad en ratones.⁴¹

La estructura tridimensional de estas toxinas no se ha estudiado. Se han realizado estudios de su estructura secundaria por dicroísmo circular en los que además se compararon con las toxinas ω -AgaIVA y PltxII.⁴³ Se encontró que las oxitoxinas presentan una estructura similar entre ellas y que su contenido de α -hélice, hoja β paralela y giros β es comparable con la de ω -AgaIVA y PltxII; por otro lado, las oxitoxinas presentan mayor proporción de regiones no

estructuradas y menor proporción de hoja β antiparalela.^{41,43} En el año 2017, Nadezhdin *et al.* reportaron la estructura de la toxina modular OtTx1a (Spiderine 1a), la cual presenta dos dominios, OtTx1a-AMP y OtTx1a-ICK. Este segundo dominio tiene alta similitud con OxyTx1 y OxyTx2; su estructura es tipo ICK y es estabilizada por cinco puentes disulfuro, lo que le confiere estabilidad en disolución acuosa.⁴⁴ Este dominio es el primer caso de un péptido tipo ICK con cinco puentes disulfuro; la elucidación de las estructuras de las oxitoxinas podría llevar a establecer una nueva familia de toxinas estabilizadas por cinco puentes disulfuro con actividad sobre canales iónicos.

Oxytoxin1	DWECLPLHSSCDNDVCCKNHHCHCPYSNVSKLEKWLPEWAKIPDALKRCSQQRNDKDGKINTCDKYKN
Oxytoxin2	AWKCLPKDSTCGDDCDCEGLHCHCPLRNMLPAILRCSQSKDDHINTCPKYKKS

FIG 3.1.1. Secuencias de aminoácidos de OxyTx1 y OxyTx2.

3.2. Justificación del proyecto

La elucidación de la estructura tridimensional de las toxinas OxyTx1 y OxyTx2 permitirá realizar análisis más detallados de la interacción de la toxina con los canales de calcio sobre los que tienen actividad. Estos tipos de análisis pueden ser utilizados para realizar caracterizaciones de los canales en cuestión, así como para diseñar experimentos para estudiar la contribución de cada uno de los dominios a los enlaces toxina-canal.

4. Hipótesis

La elucidación de la estructura tridimensional de las toxinas recombinantes rOxyTx1 y rOxyTx2 permitirá confirmar la existencia de una nueva familia estructural de toxinas de araña, y servirá como punto de partida para proyectos futuros en los que se busque estudiar la interacción de las toxinas con sus blancos moleculares.

5. Objetivos

5.1. Objetivo General

Obtener las toxinas HisOxyTx1 y HisOxyTx2 por métodos recombinantes y determinar su estructura tridimensional utilizando resonancia magnética nuclear y experimentos *in silico*.

5.2. Objetivos Particulares

- Expresar las toxinas HisOxyTx1 y HisOxyTx2 en *E. coli*.
- Extraer las proteínas de fusión de cuerpos de inclusión y replegarlas.
- Confirmar la identidad de las toxinas mediante espectrometría de masas.
- Adquirir experimentos de RMN.
- Obtener propuestas de estructuras utilizando modelado por homología.
- Refinar propuestas utilizando dinámica molecular.
- Determinar la mejor propuesta de estructura de OxyTx1 y OxyTx2.

6. Diseño Experimental y Metodología

6.1. Expresión y purificación de las toxinas HisOxyTx1 y HisOxyTx2.

Para la expresión de las toxinas en *E. coli* se utilizaron plásmidos cuya construcción fue previamente reportada por Estrada *et al.*⁴³, denominados pQE30XaOxyTx1 y pQE30XaOxyTx2. Ambas construcciones contienen una secuencia de 6 residuos de histidina denominada HisTag en el extremo N-terminal, la cual facilita la purificación de la proteína de fusión por cromatografía de afinidad metal-quelato (CAMQ), y una secuencia de corte para la proteasa FactorXa para permitir la separación de la toxina madura y la etiqueta de histidinas.

Se utilizaron células competentes de la cepa BL21 o RosettaGami para transformar con los plásmidos pQE30XaOxyTx1 y pQE30XaOxyTx2. Las células transformadas se incubaron en

medio Luria Bertani (LB) con los antibióticos correspondientes (Amp en el caso de BL21, Amp, Kan, Cam y Tet en el caso de RosettaGami) a 37°C hasta alcanzar una densidad óptica de 0.5 unidades de absorción a 600 nm. En este punto el cultivo se indujo con isopropil tiogalactpiranósido (IPTG) a una concentración de 0.5 mM por 12 h a 25°C.

6.1.1 Purificación de las toxinas a partir de fase soluble

Las células se precipitaron por centrifugación y posteriormente se resuspendieron en amortiguador de lisis (**Tris 50 mM, NaCl 300 mM pH 8.0**) y se lisaron por sonicación. El material celular lisado se centrifugó por 40 min a 16500 RPM. La fracción soluble se clarificó por filtración a 0.45 µm.

La fracción soluble se purificó por CAMQ utilizando una columna Hi-Trap cargada con Ni²⁺ y equilibrada con el mismo amortiguador utilizado para lisar el material celular. La proteína de fusión se eluyó utilizando un amortiguador con imidazol (**Tris 50 mM, NaCl 300 mM, imidazol 500 mM, pH 8.0**) y se desaló y concentró por columna de extracción en fase sólida (Strata C-18) con un gradiente de 0.05% TFA en H₂O y 0.05% TFA en acetonitrilo como disolvente polar y no polar respectivamente (los pasos del gradiente fueron 0, 20, 40, 60, 80 y 100% acetonitrilo). La toxina eluye a una concentración aproximada de 60% de acetonitrilo.

Las toxinas se purificaron a partir del material resultante por HPLC fase reversa con una columna Phenomenex C12, utilizando como eluyente un gradiente de acetonitrilo/0.05%TFA y H₂O/0.05%TFA con flujo de 1 mL/min y un gradiente de 20 – 80% de acetonitrilo. Las fracciones colectadas se liofilizaron y la identidad de las toxinas se confirmó por espectrometría de masas MALDI-TOF y resonancia magnética nuclear.

6.1.2. Purificación de las toxinas a partir de cuerpos de inclusión.

Las células se precipitaron por centrifugación y posteriormente se resuspendieron en amortiguador de resuspensión (**Tris 50 mM, NaCl 300 mM, urea 8M, pH 8.0**) y se lisaron por sonicación. El material celular lisado se centrifugó por 40 min a 16,500 RPM. El sobrenadante se clarificó por filtración a 0.45 µm.

Esta mezcla se purificó por CAMQ utilizando una columna Hi-Trap cargada con Ni²⁺ y equilibrada con el mismo amortiguador utilizado para lisar el material celular. La proteína de fusión se eluyó utilizando un amortiguador con imidazol (**Tris 50 mM, NaCl 300 mM, imidazol 500 mM, urea 8M, pH 8.0**). A la fracción eluída se le retiró el imidazol por diálisis y se realizó la reducción de los puentes disulfuro de la toxina agregando DTT (50 mM) a la disolución, incubando a temperatura ambiente (TA) por 24 h.

Posteriormente se realizó un proceso de replegamiento controlado. Para este fin, se retiró la urea y DTT de la disolución por diálisis, y ésta se añadió lentamente a un amortiguador de replegamiento (**50 mM Tris, 300 mM NaCl, 0.5 mM glutatión reducido, 5 mM glutatión oxidado, 5% glicerol**) en un volumen 10X de la disolución inicial para una mayor reducción de la concentración de urea y DTT. Esta disolución se incubó con agitación constante por 48 h.

Una vez realizada esta acción se purifica la toxina una vez más por CAMQ; se desaló y concentró por columna de extracción en fase sólida (Strata C-18), utilizando un gradiente de 0.05% TFA en H₂O y 0.05% TFA en acetonitrilo como disolvente polar y no polar respectivamente, como previamente descrito. La toxina eluye a una concentración aproximada de 60% de acetonitrilo.

Por último, las toxinas HisOxyTx se purificaron por HPLC fase reversa, utilizando la metodología descrita en la sección anterior. Las fracciones colectadas se liofilizan y la identidad de las toxinas se confirma por espectrometría de masas MALDI-TOF y resonancia magnética nuclear.

6.1.3. Escalamiento de la expresión de las toxinas para su purificación en fase soluble.

Para maximizar la cantidad de biomasa generada para la purificación de la toxina, se realizó la expresión de las toxinas en un volumen de 10 L en biorreactores. Para este fin se preparó un preinóculo de 50 mL de cada una de las cepas transformadas, los cuales se incubaron toda la noche a 30°C. Con este cultivo se inocularon dos matraces Fernbach con un volumen de

500 mL a 20°C toda la noche. Estos cultivos se utilizaron para inocular los biorreactores que fueron equilibrados previamente a 20°C con agitación mecánica y presión positiva de O₂. Los biorreactores se incubaron a esta temperatura constante hasta alcanzar una densidad óptica (OD) de 0.5-0.6. En este punto se indujeron con una concentración de 0.5 mM de IPTG y se continuó incubando a 20°C hasta alcanzar una OD constante. Durante todo el tiempo de incubación se monitorearon los niveles de pH y concentración de oxígeno en disolución. Una vez que se alcanzó la fase estacionaria se centrifugaron las células y se almacenó la biomasa para su posterior purificación utilizando la metodología previamente descrita.

6.2 Caracterización de las toxinas purificadas.

Las toxinas purificadas y liofilizadas se cuantificaron por gravimetría. Después, se analizaron por espectrometría de masas MALDI-TOF utilizando como matriz una solución saturada de ácido alfa-ciano-4-hidroxicinámico en relaciones con la muestra de 1:5 y 1:10 en agua desionizada. Las mediciones se realizaron en un equipo Bruker Daltonics Microflex. El modo de adquisición fue lineal, en polaridad positiva.

Los experimentos de dicroísmo circular se obtuvieron en un espectropolarímetro Jasco J-1500 en concentración 1 mg/mL disueltos en agua desionizada, esto se debe a que se busca tener condiciones de disolución similares a las que se utilizarán para los experimentos de RMN. Se adquirió a 20°C y se midieron longitudes de onda de 240 – 190 nm en una celda de cuarzo de paso óptico de 0.1 cm de espesor.

Los espectros de Resonancia Magnética Nuclear de una dimensión se adquirieron en un equipo Varían 500 MHz. Las toxinas liofilizadas se resuspendieron en una mezcla H₂O/D₂O 5% o PBS/D₂O 5% según el experimento. Las concentraciones de muestra dependen del experimento y la cantidad disponible de cada toxina y varían entre 0.5 y 1.5 mg/mL

6.3. Predicción de estructuras por experimentos *in silico*

6.3.1. Predicción de estructuras por servidores de modelado comparativo.

Para determinar la estructura tridimensional de las toxinas se utilizaron servidores de predicción por homología. En todos los predictores utilizados, el primer paso debe ser encontrar una proteína con similitud alta a la proteína de interés. Ésta se denomina proteína plantilla. A partir de la proteína plantilla y la secuencia de la proteína de interés, cada servidor sigue un protocolo distinto.

Servidores utilizados

I-Tasser

I-Tasser (Iterative Threading ASSEmbly Refinement) es un algoritmo jerárquico de predicción de estructura de proteínas. Consiste en una comparación inicial de la secuencia introducida con una biblioteca de PDBs. Al identificar segmentos similares al PDB plantilla, el servidor ensambla los fragmentos utilizando simulaciones Monte Carlo y determina la estructura de las regiones sin correspondencia (por lo general giros) por medio de modelado *ab initio*.^{45,46}

SwissModel

Swiss-Model es un servidor para el modelado comparativo de estructuras tridimensionales de manera automática. Los modelos se procesan utilizando un pipeline que se basa en ProMod3, un sistema de comparación basado en OpenStructure. Después de encontrar una plantilla, este sistema resuelve inserciones y deleciones utilizando candidatos de una base de datos de estructuras. Si no se pueden encontrar estos, se realiza una búsqueda espacial usando técnicas Monte Carlo. Para cadenas laterales no conservadas se utiliza una librería de rotámeros dependientes en la cadena primaria, y su configuración óptima se estima usando el algoritmo TreePack. Como último paso, se resuelven choques o interacciones no favorables utilizando algoritmos de minimización de energía.⁴⁷

RaptorX

RaptorX es un servidor basado en plantillas que predice estructura secundaria y terciaria. Es útil en casos donde no existen varios homólogos cercanos a la secuencia de interés. Su primer paso es hacer un perfil de los homólogos disponibles para la secuencia. Después, utiliza campos condicionales aleatorios para integrar una variedad de señales biológicas en una función de

paralelismo no lineal. Por último, utiliza paralelismo de múltiples moldes para modelar una sola secuencia, usando similitudes interplantilla si estas existen para refinar la estructura.⁴⁸

6.3.2. Refinamiento de las estructuras por dinámica molecular.

Se realizaron 500 ns de tiempo de simulación para cada una de las toxinas con la suite AMBER 20⁴⁹, utilizando el campo de fuerza FF19SB⁵⁰. La protonación a pH 7.4 y fuerza iónica de 0.1 se realizó con el programa H++⁵¹. La caja de solvatación octaédrica se generó con una distancia mínima de 18 Å de la proteína al borde de la caja, con el modelo de disolvente TIP3P⁵². Se colocaron contraiones para neutralizar el sistema en posiciones aleatorias. Se utilizaron condiciones periódicas de frontera, y se realizó un ensamble NPT (37°C, 1 atm). Se realizó minimización por gradiente conjugado (5000 pasos) seguido por calentamiento NVT (0.8 ns), un equilibrado NPT (1 ns) usando dinámica de Langevin para equilibrar el disolvente y relajar la estructura. Los análisis de las trayectorias se realizaron utilizando el programa CPPTRAJ⁵³ y VMD⁵⁴. Las representaciones moleculares se produjeron con el programa Chimera⁵⁵.

7. Resultados y Discusión

7.1. Optimización de expresión

Se transformaron células de la cepa BL21 con los plásmidos correspondientes a HisOxyTx1 y HisOxyTx2 y se realizaron pruebas de expresión de ambas toxinas durante 15 h (Fig 7.1.1).

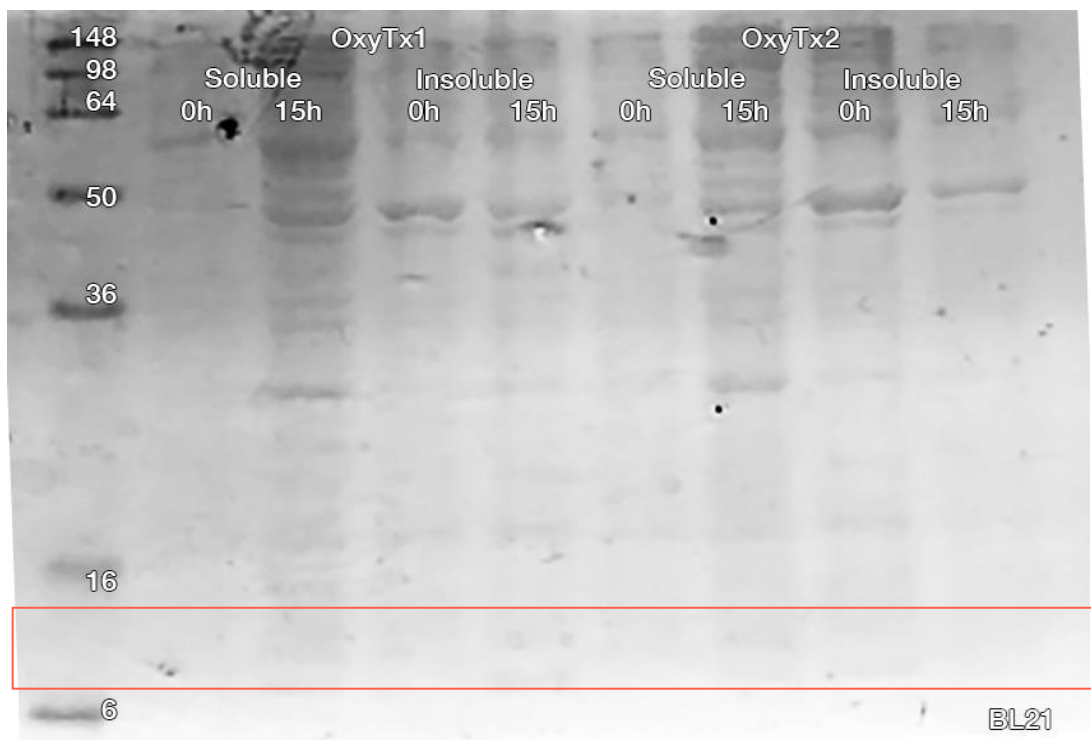


FIG 7.1.1. SDS-PAGE de la prueba de expresión de HisOxyTx1 y HisOxyTx2 en BL21. La fracción insoluble se resuspendió con amortiguador de lisis sin agente caotrópico. En rojo se muestra la región en la que se esperaría ver las bandas correspondientes a las toxinas de interés. La columna de la izquierda corresponde al marcador de masa molecular; todos los valores se presentan en kDa.

En la Fig. 7.1.1 no se observa una sobreexpresión de alguna proteína en la región correspondiente a la masa molecular, las cuales son 9.9 kDa para HisOxyTx1 y 8.0 para HisOxyTx2. Al haber adquirido las células transformadas la resistencia al antibiótico Amp, se espera que la transformación se haya llevado a cabo adecuadamente. Sin embargo, los resultados observados en el gel de poliacrilamida indican que la sobreexpresión de las proteínas no se llevó a cabo como se esperaba. BL21 no es una cepa que confiera propiedades oxidantes a *E. Coli*, las cuales permiten la correcta expresión de este tipo de neurotoxinas ricas en puentes disulfuro. Esta puede ser la razón por la cual no se logró obtener las neurotoxinas utilizando esta cepa bacteriana.

Con el fin de determinar una metodología adecuada para la expresión de estos péptidos, se repite la transformación en la cepa RosettaGami. La razón por la que se eligió esta cepa es porque presenta un ambiente oxidante que se espera que favorezca el repliegamiento de las

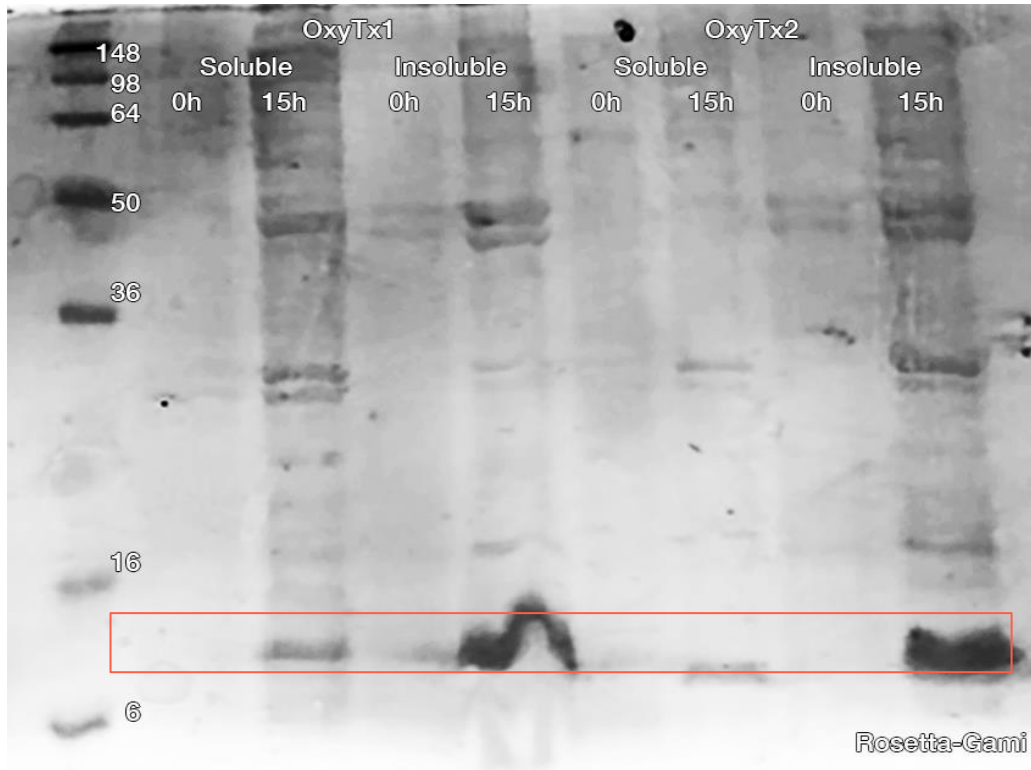


FIG 7.1.2. SDS-PAGE de la prueba de expresión de HisOxyTx1 y HisOxyTx2 en RosettaGami. La fracción insoluble se resuspendió con amortiguador de lisis sin agente caotrópico. El rectángulo en rojo muestra la región en la que se esperaría ver las bandas correspondientes a las toxinas de interés. La columna de la izquierda corresponde al marcador de masa molecular; todos los valores se presentan en kDa.

toxinas y, por lo tanto, el aumento de la concentración de estas proteínas en la fase soluble para que así se purifique sin necesidad de un proceso de repliegamiento controlado.

En la figura 7.1.2 se observa que en esta cepa hay una sobreexpresión de las toxinas por la presencia de una banda alrededor de 8-9 kDa en ambos casos. Este resultado confirma la capacidad de la cepa RosettaGami de sobreexpresar las proteínas de interés. También se observa que las toxinas se encuentran en mayor concentración en la fase insoluble del lisado.

Una vez que se confirmó la sobreexpresión de las toxinas en la cepa, se realizó una prueba de expresión a lo largo de 18 horas, monitoreando la concentración de las toxinas para establecer el tiempo adecuado de incubación. Con este fin se inoculó medio LB con 3% de stock de glicerol de células RG previamente transformadas con cada plásmido. Se indujo con IPTG 0.5

mM al alcanzarse una OD de 0.5-0.6 y se incubó a 25°C. Una vez que se indujo el cultivo se tomaron muestras de 1 mL por hora para monitorear la expresión. Estas muestras se centrifugaron para bajar el botón celular; se resuspendieron en 150 µL de amortiguador de lisis; se lisaron por sonicación; y, se centrifugaron nuevamente para precipitar la fase insoluble. Para preparar las muestras para su análisis en geles SDS-PAGE se utilizó la fase soluble sin diluir y la fase insoluble se resuspendió en 150 µL de amortiguador de lisis. En ambos casos se utilizaron 20 µL de muestra y un volumen igual de amortiguador de carga 2X.



FIG 7.1.3. Cinética de expresión de la toxina HisOxyTx1 a 25°C con IPTG 0.5 mM. La columna de la izquierda corresponde al marcador de masa molecular; todos los valores se presentan en kDa.

En el caso de HisOxyTx1 (Fig 7.1.3), se observó que existe una concentración apreciable de la toxina a partir de las 4 horas de incubación en la fase insoluble; mientras que en el caso de la fase soluble ésta se aprecia hasta después de 6 horas. Esta discrepancia parece indicar que la toxina se encuentra mayoritariamente en su forma insoluble y tiende a formar cuerpos de inclusión, donde se generan agregados moleculares de la proteína que no permiten que ésta se solubilice en agua. Por lo tanto, se observa una cantidad apreciable de la toxina de interés en la fase insoluble mientras que en la fase soluble la concentración sigue siendo baja. En el caso de HisOxyTx2 (Fig 7.1.4), no se observa una concentración apreciable de toxina en ninguna de las fases antes de 6 horas, indicando que puede existir una menor tendencia a formar cuerpos de inclusión. En ambos casos, la concentración parece alcanzar un máximo alrededor de las 8 horas, ya que no existe una diferencia apreciable en la concentración entre 8 y 16 horas. Este resultado indica que no parece haber una degradación de la toxina o una disminución en su solubilidad después de periodos largos de incubación, ya que la proporción de concentración en fase soluble e insoluble permanece constante.

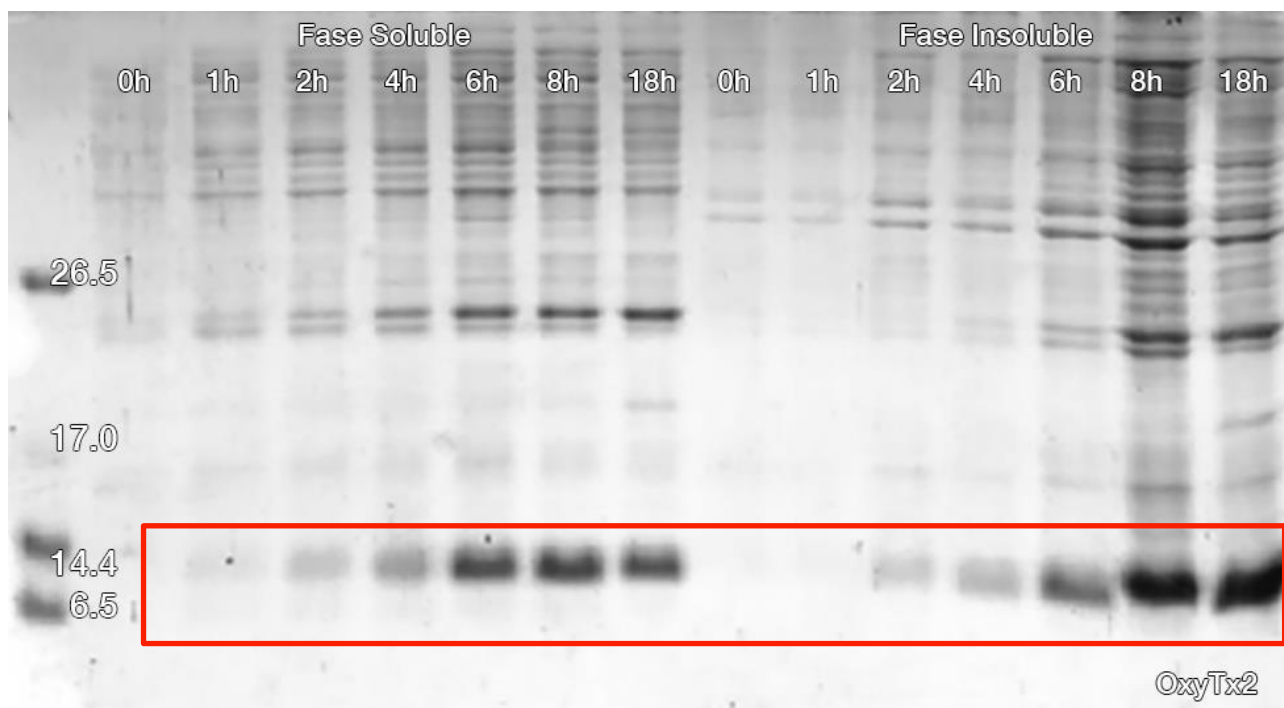


FIG 7.1.4. Cinética de expresión de la toxina HisOxyTx2 a 25°C con IPTG 0.5 mM. La columna de la izquierda corresponde al marcador de masa molecular; todos los valores se presentan en kDa.

7.2. Purificación de las toxinas a partir de fase soluble e insoluble.

Una vez que se determinaron las condiciones de expresión de las toxinas en RosettaGami, así como la tendencia de las toxinas a formar cuerpos de inclusión, se realizaron pruebas de purificación utilizando condiciones desnaturizantes con la intención de llevar a cabo un replegamiento controlado, siguiendo el protocolo establecido por Estrada et al⁴³, con algunas modificaciones en cuanto a los reactivos utilizados. Es importante mencionar que, a lo largo del proceso de obtención de las toxinas para su caracterización estructural, no se llevó a cabo el corte proteolítico de la etiqueta de purificación ya que se espera que esta no interfiera con el plegamiento de las proteínas.

Con el fin de simplificar el proceso de obtención de las toxinas, se realizaron pruebas de purificación de la fase soluble al lisar sin un agente desnaturizante. A pesar de observarse que la concentración en la fase soluble es menor a la de la fase insoluble, se espera obtener una mayor cantidad de toxina pura y con el plegamiento correcto, ya que el proceso de desnaturización de cuerpos de inclusión y posterior replegamiento controlado pueden ser fuentes de pérdida de proteína.

En todos los casos, el material celular que se sometió a pruebas de purificación fue obtenido a partir de cultivos de 1.5 L inducidos con 0.5 mM IPTG, incubados a 25°C durante 16 horas.

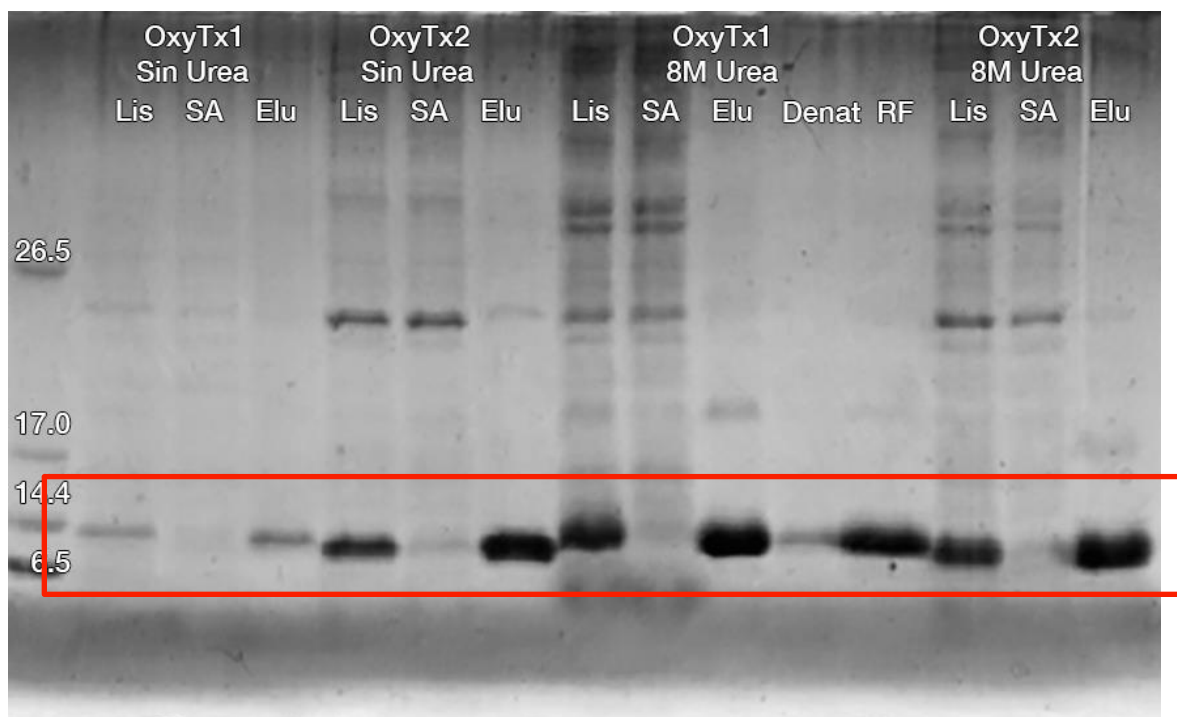


FIG. 7.2.1. Pruebas de purificación de HisOxyTx1 y HisOxyTx2 en presencia y ausencia de urea. En el caso de las pruebas sin urea, se trabajó exclusivamente la fracción soluble resultante de la centrifugación del material lisado. En todos los casos, *Lis* se refiere al material lisado previo a ser cargado en la columna HisTrap, *SA* se refiere a la mezcla que sale de la columna y que corresponde a las proteínas sin afinidad a ésta, y *Elu* a la fracción que sale de la columna al realizar el lavado con amortiguador de elución (500 mM imidazol). *Denat* hace referencia a la fracción solubilizada tratada con agentes desnaturizantes (DTT), y *RF* a esta misma fracción después del proceso de repliegamiento controlado. La columna de la izquierda corresponde al marcador de masa molecular; todos los valores se presentan en kDa.

En el caso de ambas toxinas se observa que hay una mayor concentración de proteína al lisar con urea (Fig 7.2.1); sin embargo, se puede asumir que ésta se encuentra desnaturizada al estar en contacto con una concentración alta de agente caotrópico. Aunque la concentración de las toxinas en fase soluble al lisar sin urea es menor, es factible esperar que se encuentren en una conformación lo suficientemente estable para ser solubilizadas sin ayuda de un agente caotrópico. Esto indicaría que se pueden encontrar en su estado de plegamiento nativo.

En el caso de las toxinas purificadas a partir de fracción soluble, se procede directamente al desalado y concentración de la muestra por columna de extracción en fase sólida (Strata) para su subsecuente purificación por RP-HPLC.

En el caso de las toxinas purificadas en presencia de agente caotrópico, se procede al replegamiento controlado. Para llevar a cabo este proceso es necesario reducir por completo los puentes disulfuro que pueden haberse formado en conformaciones intra e intermoleculares distintas a las esperadas. Esto se lleva a cabo incubando las fracciones eluidas de la columna con DTT 50 mM a TA por 24 horas. Una vez transcurrido este tiempo, se retira el agente reductor y el agente caotrópico de la disolución por diálisis y dilución; se coloca la proteína en amortiguador de replegamiento. Se incuba a TA con agitación constante durante 48 horas y se vuelve a aislar la proteína por medio de CAMQ. Una vez obtenida la toxina del replegamiento, se procede a desalar y concentrar por extracción en fase sólida para su purificación en RP-HPLC.

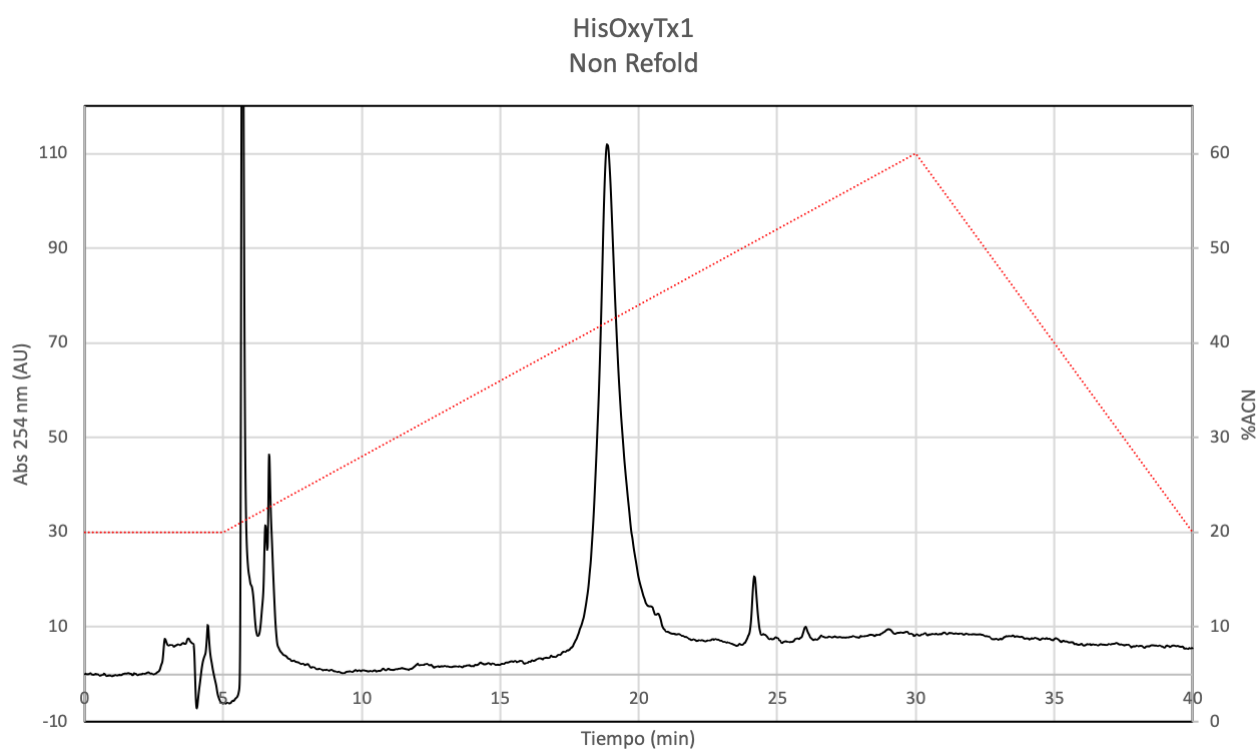


FIG 7.2.2. Cromatograma fase reversa de HisOxyTx1 purificado de fase soluble en ausencia de agente caotrópico obtenido a partir de una columna C12 con gradiente 20 – 60% acn.

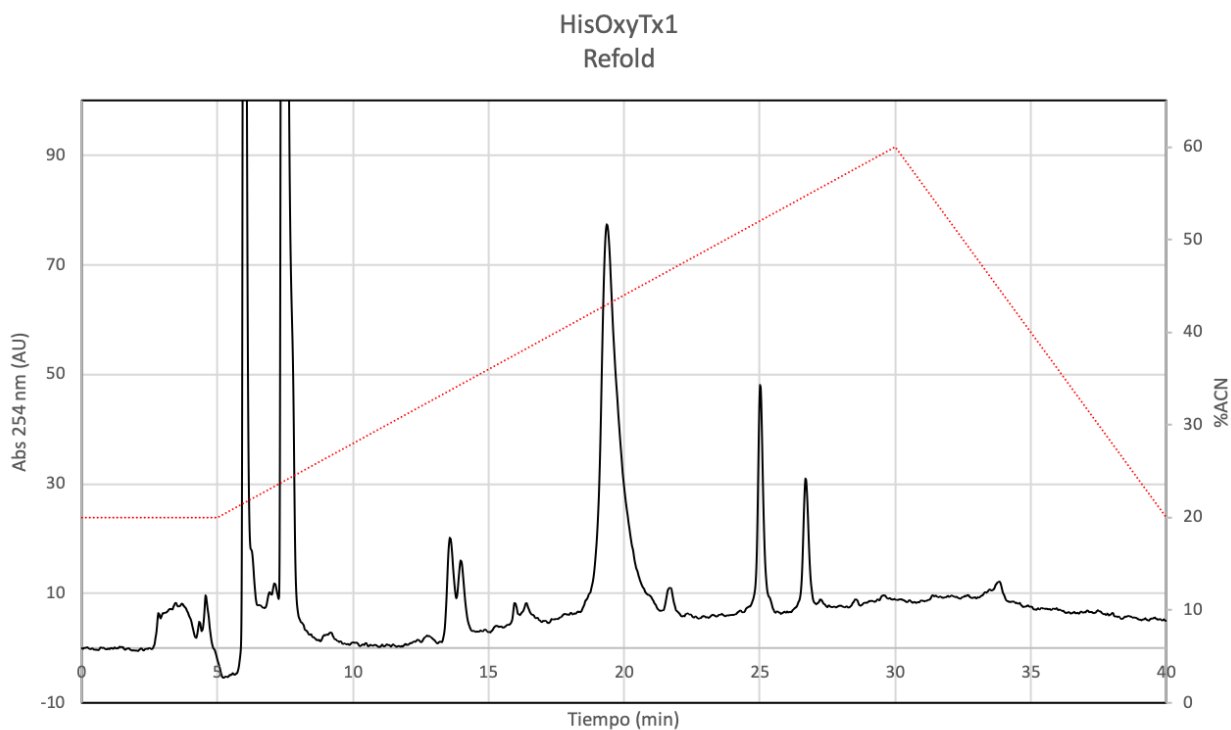


FIG 7.2.3. Cromatograma fase reversa de HisOxyTx1 purificado de cuerpos de inclusión y posterior repliegamiento controlado obtenido a partir de una columna C12 con gradiente 20 – 60% acn.

Las muestras de ambas toxinas, desaladas y concentradas tanto en su forma replegada como no replegada fueron inyectadas en el equipo de RP-HPLC utilizando una columna de fase estacionaria C12, con una fase móvil de H₂O/TFA 0.05% como disolvente polar y acetonitrilo/TFA 0.05% como disolvente no polar. Para la purificación de las toxinas de interés, se utilizó un gradiente de 20 a 60% de acetonitrilo por 30 minutos. En ambos casos se observaron múltiples picos que pueden indicar la presencia de isoformas de las proteínas (Fig 7.2.2 y 7.2.3); se colectaron sólo los picos predominantes (que en ambos casos eluyen alrededor de 18 minutos y 43% de acetonitrilo) para su posterior estudio bajo la expectativa de que éstos correspondan a la conformación más estable formada.

En ambos casos es evidente la presencia de picos minoritarios que eluyen alrededor del predominante; esto evidencia la tendencia de la toxina a formar isoformas. Sin embargo, se puede concluir que el proceso de replegamiento favorece la formación de varias isoformas, ya que a pesar de que la disolución eluida de la columna después de la purificación con agente caotrópico presenta una mayor concentración; el valor de absorbancia que se observa es menor comparado con el de la mezcla resultante de la purificación sin agente caotrópico. Esto indica que el replegamiento resultó en la formación de un mayor número de isoformas distintas a la más estable (la que se esperaría fuera la nativa), evidenciadas por la presencia de un mayor número de picos en el cromatograma de la figura 7.2.3 en comparación con el anterior.

Se observa en la fig 7.2.4 que en ambos casos los picos predominantes eluyen a tiempos y

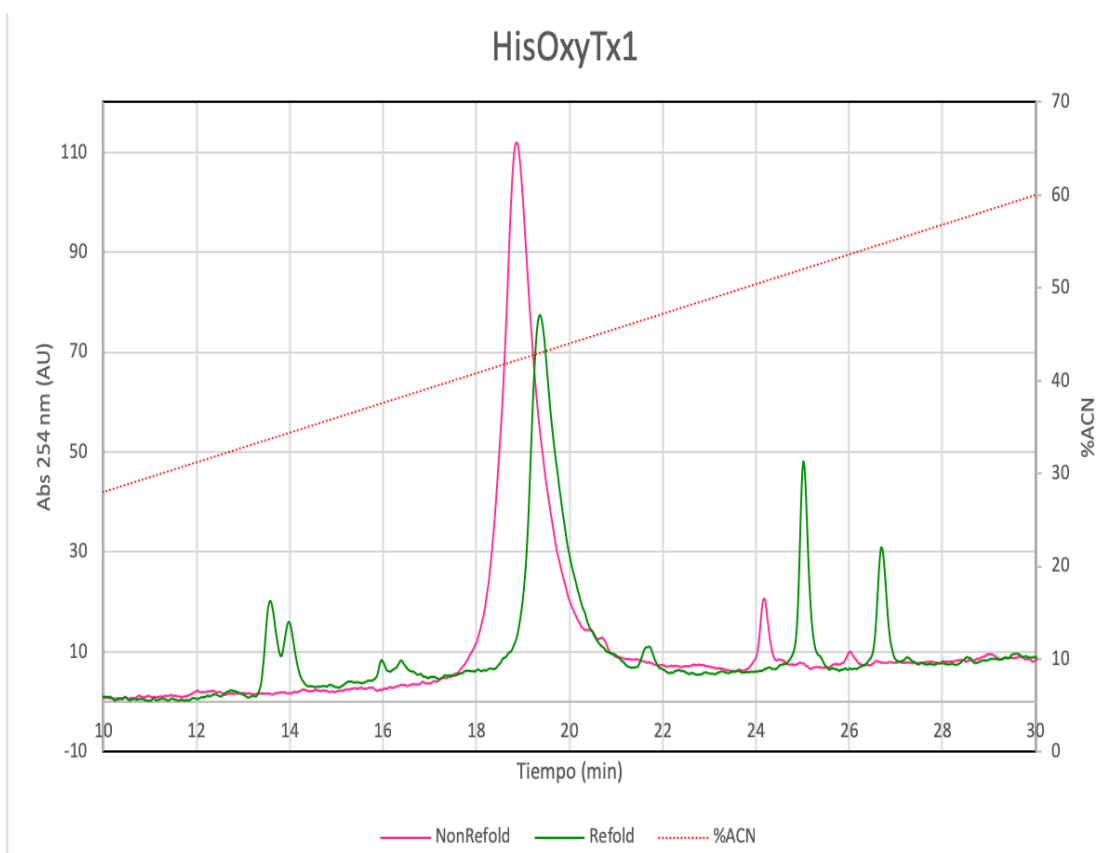


FIG 7.2.4. Comparación de los cromatogramas de HisOxyTx1. En rosa se muestran los picos correspondientes a la purificación de la toxina a partir de la fase soluble en ausencia de agente caotrópico y en verde los picos correspondientes a la purificación de la toxina sometida a un proceso de replegamiento controlado.

concentraciones de acetonitrilo similares, siendo la diferencia de tiempo de retención menor a 1 minuto y de concentración de acetonitrilo menor de 1%.

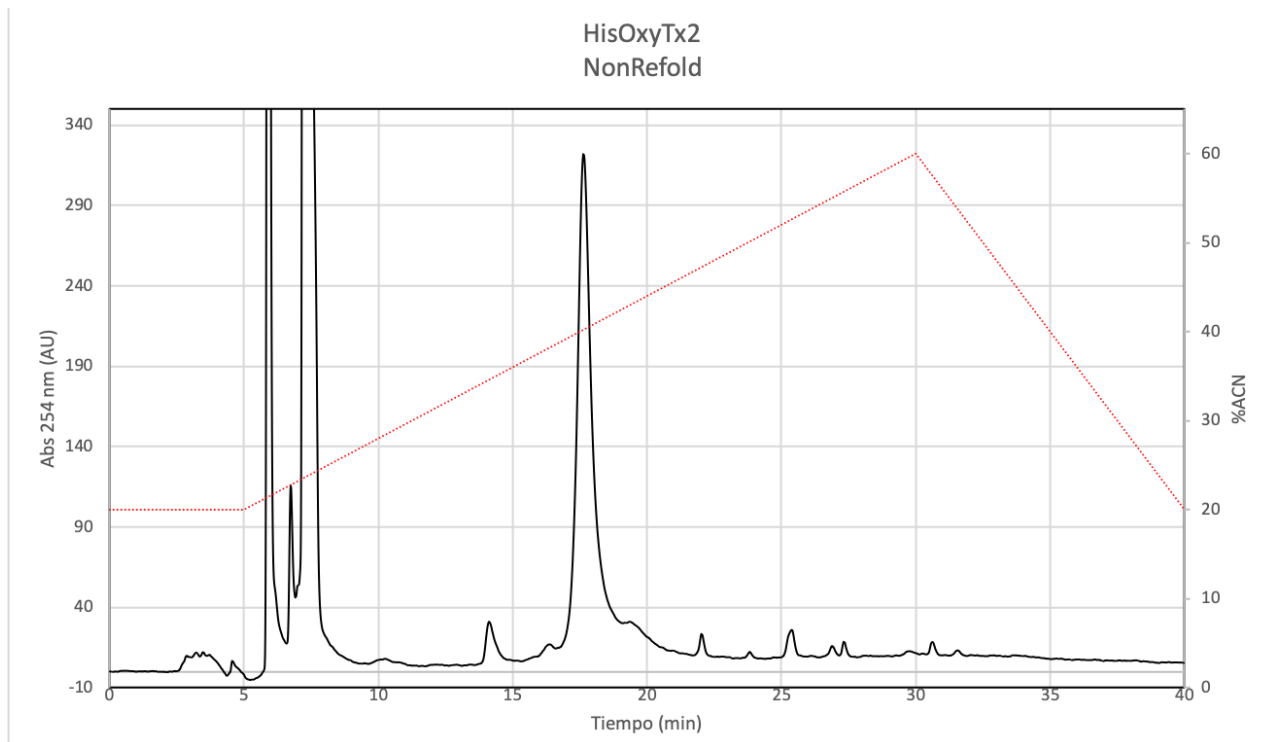


FIG 7.2.5. Cromatograma fase reversa de HisOxyTx2 purificado de fase soluble en ausencia de agente caotrópico obtenido a partir de una columna C12 con gradiente 20 – 60% acn.

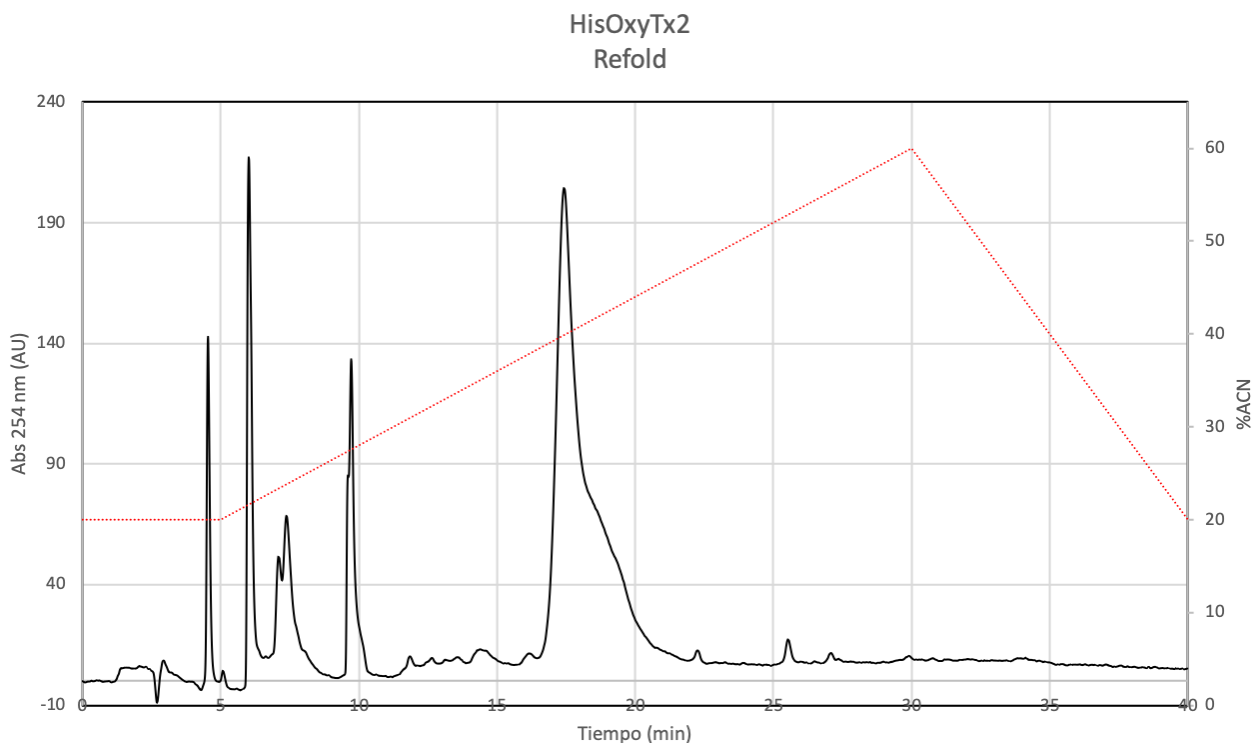


FIG 7.2.6. Cromatograma fase reversa de HisOxyTx2 purificado de cuerpos de inclusión y posterior replegamiento controlado obtenido a partir de una columna C12 con gradiente 20 – 60% acn.

Para la toxina HisOxyTx2 se observa la formación de una mayor cantidad de isoformas al purificar de la fracción soluble en comparación con el caso de HisOxyTx1. Existe un pico predominante, lo cual indicaría la presencia de una isoforma más estable que es favorecida (Fig 7.2.5). En el caso del cromatograma correspondiente a la purificación con posterior replegamiento (Fig 7.2.6), se observa que el pico predominante se mantiene alrededor del mismo tiempo de elución; sin embargo, existe un “hombro” que podría indicar la presencia de diferentes especies que eluyen a concentraciones ligeramente diferentes de acetonitrilo. Los picos correspondientes a diferentes isoformas disminuyen en intensidad, es decir que el proceso de replegamiento favorece la generación de una isoforma principal y reduce la incidencia de variaciones en las estructuras.

Tanto los picos predominantes como los picos minoritarios eluyen alrededor del mismo tiempo y concentración de acetonitrilo (Fig 7.2.7). Existe una disminución de la intensidad de todos los picos lo que indica que su concentración es menor. En el caso del pico predominante,

se asume que la disminución en intensidad se debe a la existencia del “hombro” que indica la presencia de al menos dos especies eluyendo alrededor del mismo tiempo.

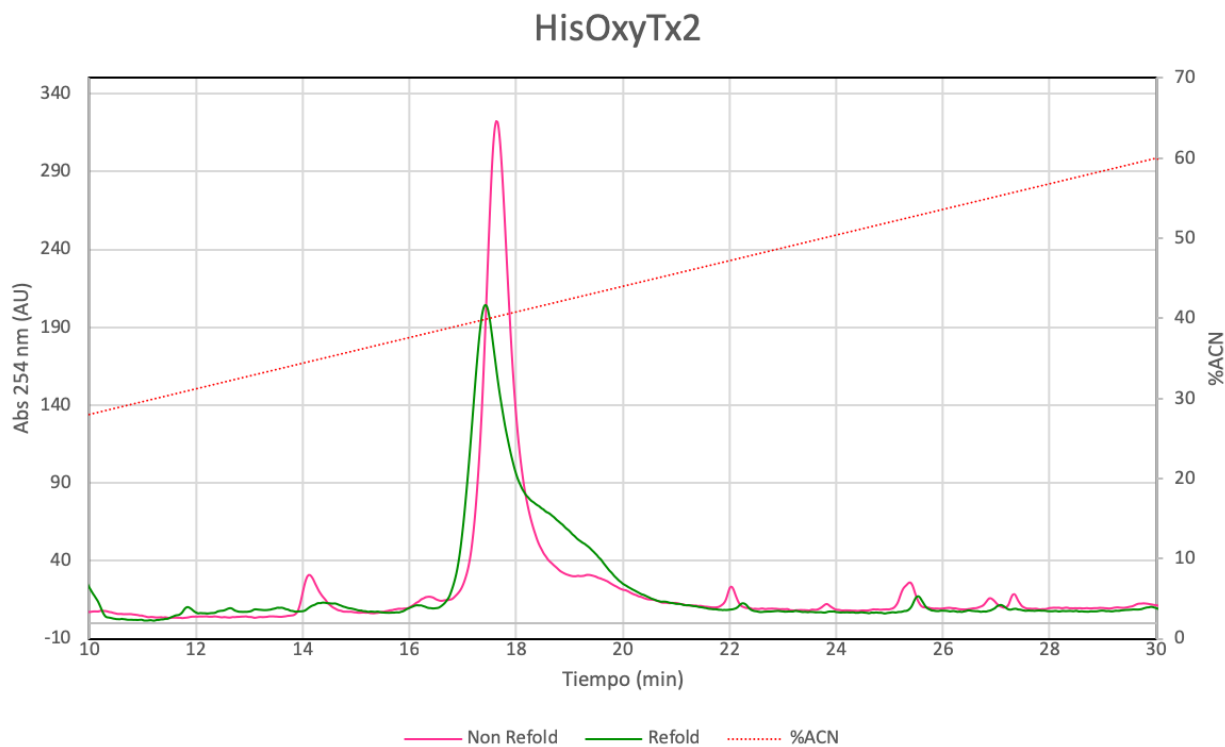


FIG 7.2.7. Comparación de los cromatogramas de HisOxyTx2. En rosa se muestran los picos correspondientes a la purificación de la toxina a partir de la fase soluble en ausencia de agente caotrópico y en verde los picos correspondientes a la purificación de la toxina sometida a un proceso de replegamiento controlado.

7.3. Escalamiento de la expresión.

Con el fin de obtener una mayor cantidad de toxina a partir de la fracción soluble (en la siguiente sección se presentan las cantidades obtenidas a partir de cada metodología), se realizó la expresión de ambos péptidos en biorreactores de 10 L. Todo el proceso de incubación se llevó a cabo a 20°C con el propósito de minimizar la precipitación de los cuerpos de inclusión. Se monitoreó la densidad óptica del cultivo (OD600), el nivel de pH y el porcentaje de oxígeno disuelto (%O₂).

Monitoreo de la expresión de HisOxyTx1 a 20°C en bioreactor de 10 L			
HisOxyTx1			
t (h)	OD	pH	%O2
0	0.068	6.86	100
2	0.136	6.91	96.7
4	0.251	6.82	97.5
6	0.334	6.78	87.6
7	0.391	6.79	86.4
8	0.462	7.00	80.5
9*	0.531	6.83	75.1
12	0.666	7.19	69.8
14	1.022	7.21	66.4
17	1.334	7.15	76.7
19	1.302	7.30	72.4
21	1.370	7.20	73.3

Tabla 7.3.1 El valor de tiempo marcado con * indica el punto en el que se adicionó el IPTG.

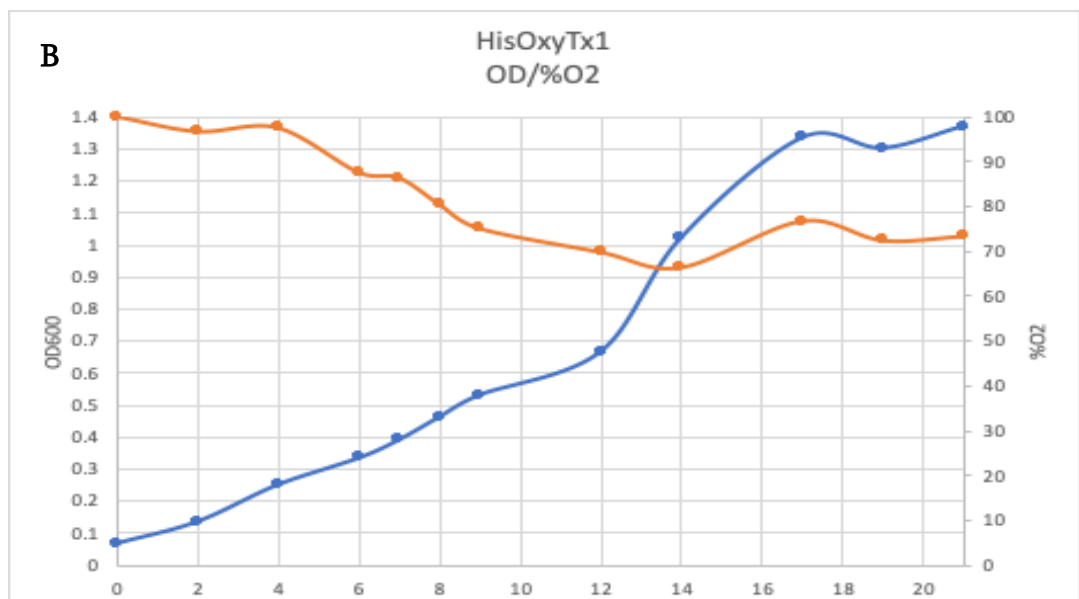
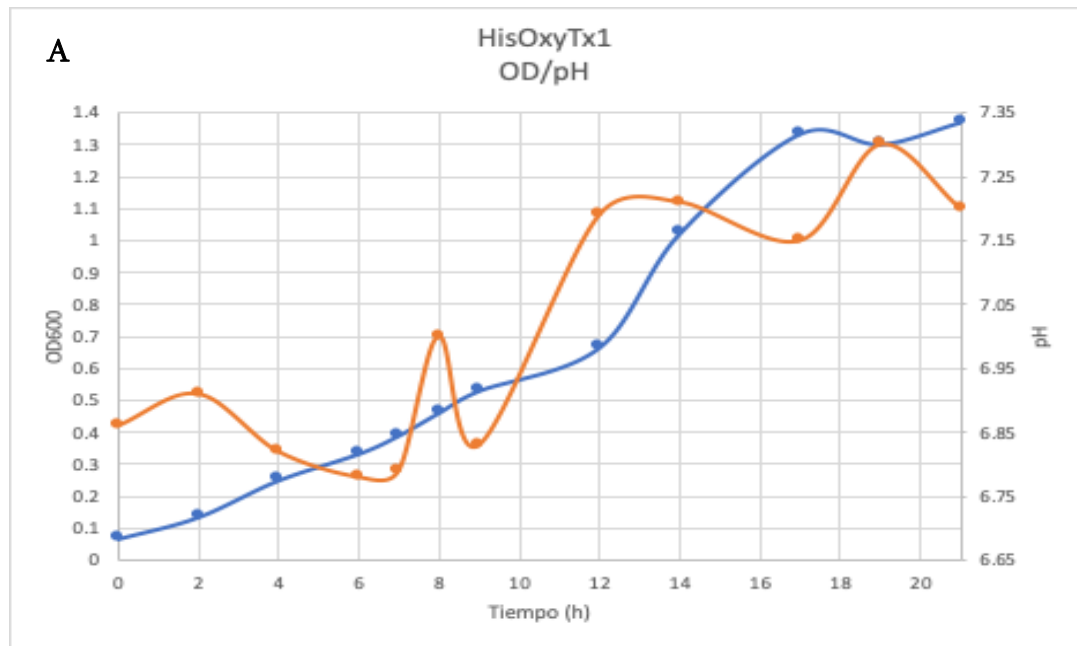
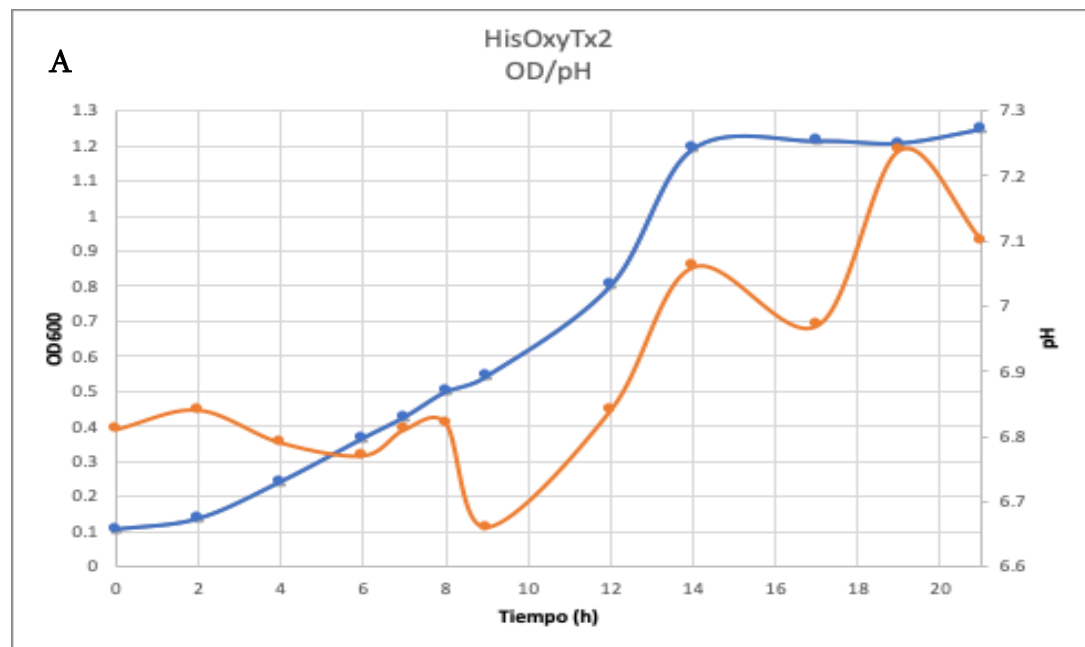


FIG 7.3.1. a) Cinética de crecimiento a 20°C de RG transformada con pQE30XaOxyTx1 mostrando OD600 y pH. b) Cinética de crecimiento a 20°C de RG transformada con pQE30XaOxyTx1 mostrando OD600 y %O₂. La línea en azul corresponde a OD600 y en naranja a pH y %O₂ respectivamente.

Monitoreo de la expresión de HisOxyTx2 a 20°C en bioreactor de 10 L				
HisOxyTx2				
t (h)	OD	pH	%O2	
0	0.105	6.81	100	
2	0.136	6.84	82	
4	0.241	6.79	81	
6	0.365	6.77	75.9	
7	0.425	6.81	71.3	
8	0.497	6.82	68.1	
9*	0.542	6.66	65.3	
12	0.8	6.84	64.3	
14	1.19	7.06	74.4	
17	1.212	6.97	85.1	
19	1.206	7.24	84.4	
21	1.246	7.1	83.1	

Tabla 7.3.2 El valor de tiempo marcado con * indica el punto en el que se adicionó el IPTG.



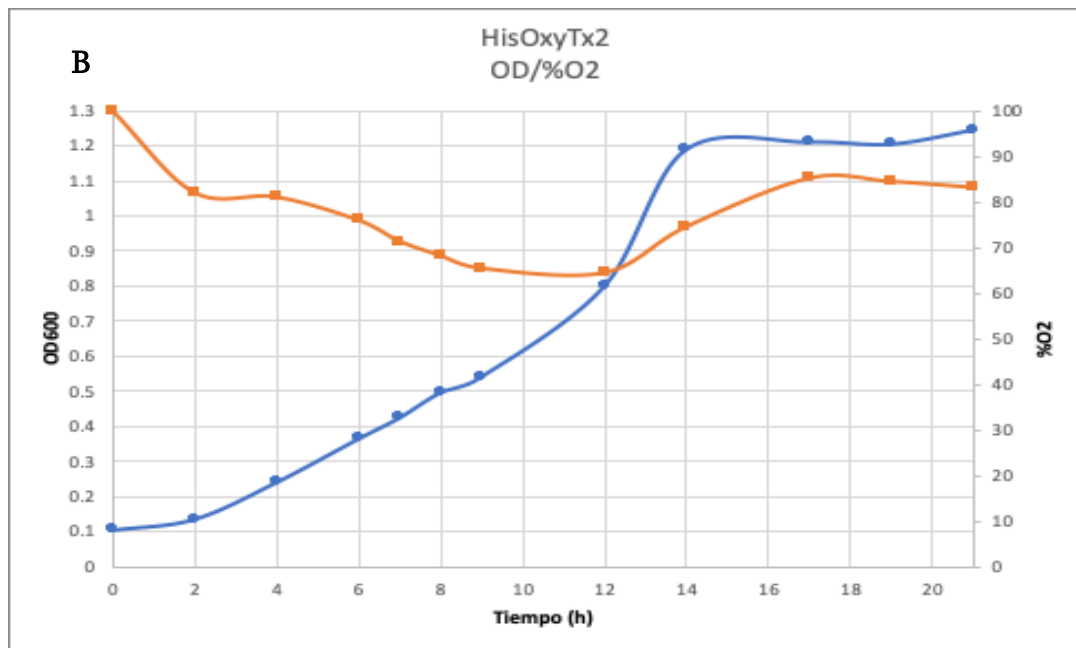


FIG 7.3.2. a) Cinética de crecimiento a 20°C de RG transformada con pQE30XaOxyTx2 mostrando OD600 y pH. b) Cinética de crecimiento a 20°C de RG transformada con pQE30XaOxyTx2 mostrando OD600 y %O2. La línea en azul corresponde a OD600 y en naranja a pH y %O2 respectivamente.

En ambos casos (Fig 7.3.1 y 7.3.2) se puede observar que no existe un cambio abrupto entre la fase de latencia y la fase exponencial; esto se debe a que la temperatura de incubación es menor a la temperatura óptima de crecimiento de *E. coli*. Para ambos cultivos la fase estacionaria se observa alrededor de 14 horas de incubación y 1.2 OD. No se observa una correlación clara entre el crecimiento celular y el nivel de pH del cultivo en ninguno de los casos. En ambos casos se puede observar una disminución de la concentración de oxígeno disuelto proporcional al crecimiento celular hasta el punto en el que se alcanza la fase estacionaria, cuando se observa una recuperación de la concentración de oxígeno.

7.4. Cuantificación de las toxinas

Una vez purificadas las toxinas, se procedió a su caracterización. En primer lugar, se cuantificó la toxina obtenida de cada una de las variaciones en metodología que se siguieron.

Cuantificación gravimétrica de las toxinas obtenidas por diferentes metodologías.						
Toxina	HisOxyTx1 NR	HisOxyTx1 Rep	HisOxyTx2 NR	HisOxyTx2 Rep	HisOxyTx1 Reactor	HisOxyTx2 Reactor
V cultivo (L)	1.5	1.5	1.5	1.5	5	5
m (μg)	200	<100	200	400	1600	1000
Rendimiento ($\mu\text{g/L}$)	133	66	133	267	320	200

Tabla 7.4.1 NR indica toxinas obtenidas de la fase soluble sin replegamiento posterior; Rep indica toxinas obtenidas de fase insoluble con replegamiento posterior y Reactor indica toxinas obtenidas del proceso escalado a biorreactores (también de fase soluble)

7.5. Espectrometría de masas MALDI-TOF

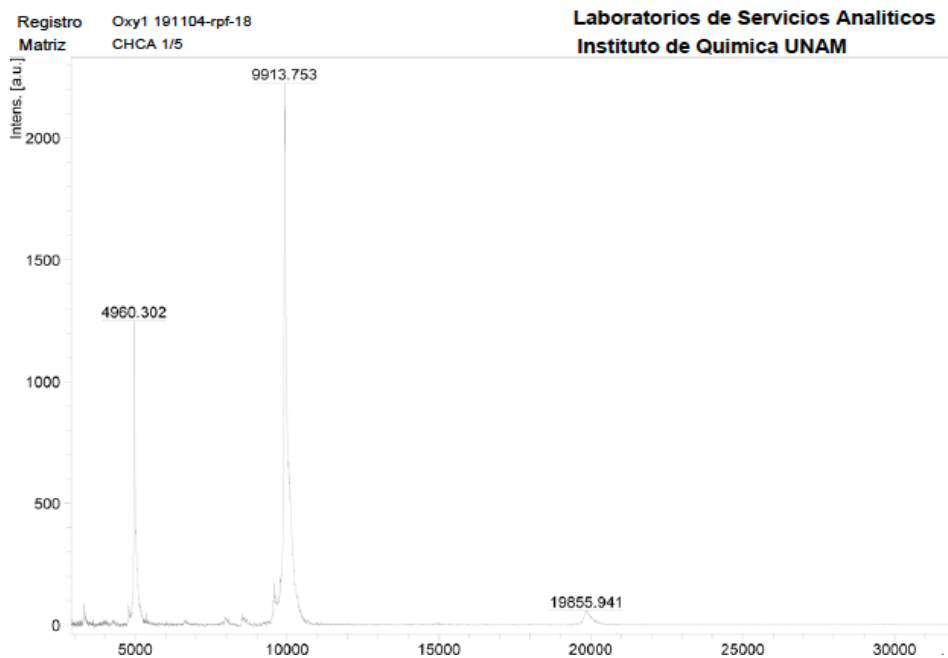


FIG 7 5 1 Espectro de masas MALDI-TOF de HisOxyTx1

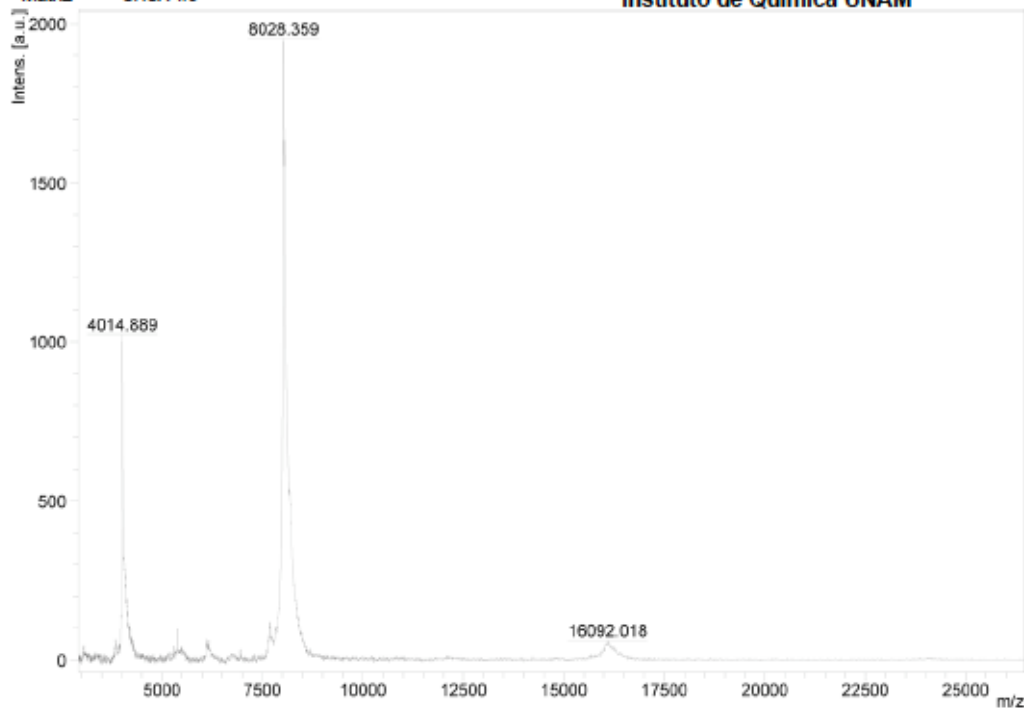


FIG 7.5.2. Espectro de masas MALDI-TOF de HisOxyTx2.

En la fig 7.5.1 se observa un pico mayoritario correspondiente al ion molecular de la proteína HisOxyTx1, el cual muestra un valor de masa molecular de 9913.753 Da cuyo valor de masa molecular esperada en su forma reducida es de 9923.1 Da. La masa observada experimentalmente es consistente con la masa que se esperaría si la molécula se encuentra en su forma oxidada con los cinco puentes disulfuro esperados formados

Para el caso de HisOxyTx2, en la fig 7.5.2 el pico mayoritario se observa en un valor de masa de 8028.359 Da. La masa esperada para el péptido en su forma reducida es de 8040.1 Da. Esta masa experimental es consistente con la que se esperaría para una molécula en su forma oxidada, con los cinco posibles puentes disulfuro formados.

7.6. Dicroísmo circular

Después de determinar que las toxinas obtenidas presentan la masa molecular y estado de oxidación esperados, se realizaron experimentos de dicroísmo circular para determinar los motivos de estructura secundaria presentes. Los resultados se analizaron utilizando el servidor DichroWeb.⁵⁶

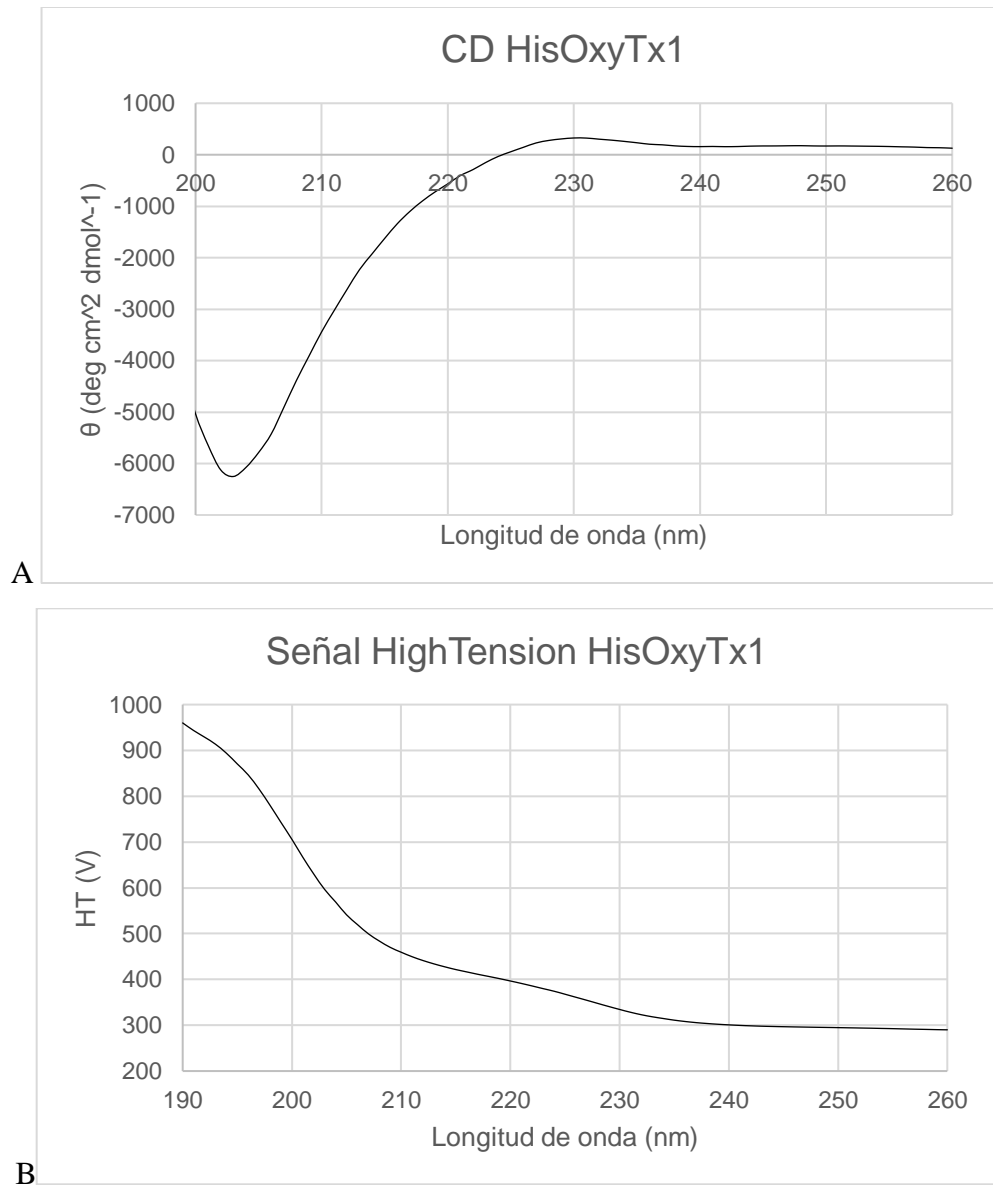


FIG 7.6.1 A) Espectro de dicroísmo circular reconstruido para HisOxyTx1 y B) su señal HighTension correspondiente.

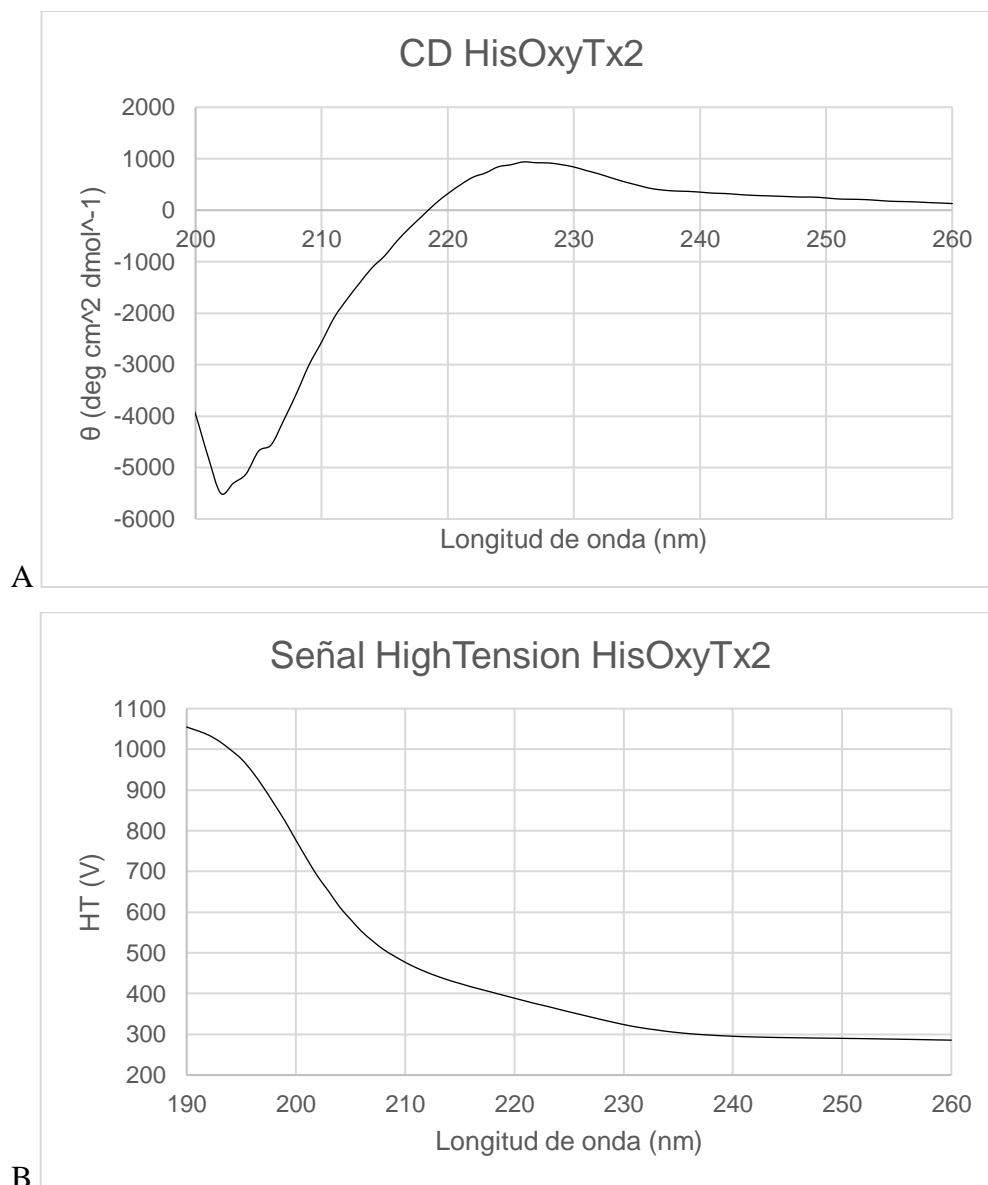


FIG 7.6.2 A) Espectro de dicroísmo circular reconstruido para HisOxyTx2 y B) su señal HighTension correspondiente.

El motivo estructural ICK está compuesto de hojas beta predominantemente, con el resto del péptido asumiendo una conformación tipo *random coil* (desestructurada). En un espectro de CD, se espera observar un máximo por encima del eje horizontal alrededor de 190 nm, así como un mínimo por debajo del eje horizontal entre 210 y 220 nm para una proteína cuya estructura secundaria es predominantemente de hoja beta. Para una proteína predominantemente desestructurada, se espera un mínimo por debajo del eje horizontal entre 190 y 200 nm. En la figura 7.6.3 se observan ejemplos representativos de las diferentes curvas que se presentan en un espectro de CD.^{57,58}

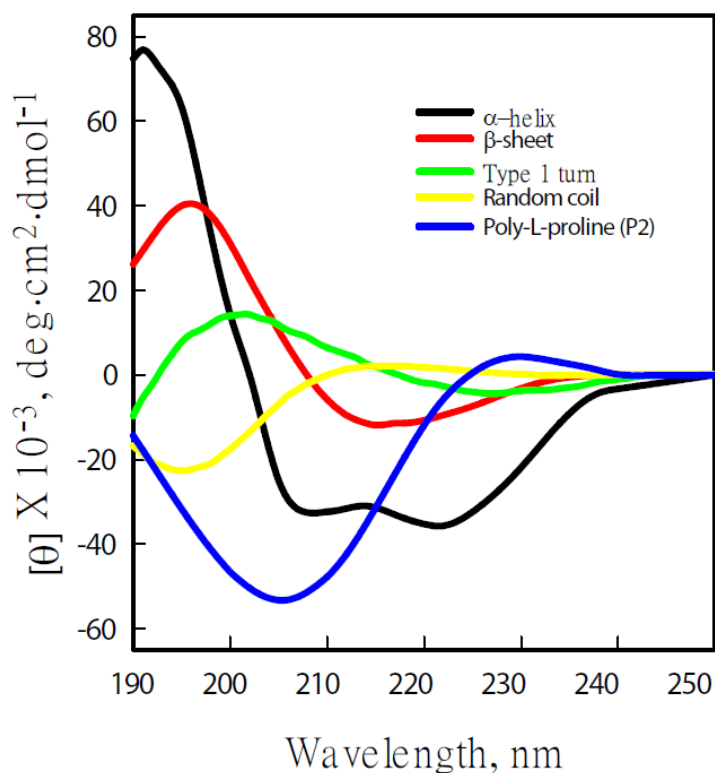


FIG. 7.6.3 Espectros de dicroísmo circular esperados para proteínas con distintas estructuras secundarias predominantes. De interés para este trabajo son las correspondientes a hoja beta (rojo) y proteínas desestructuradas (amarillo).

En los espectros de CD obtenidos para HisOxyTx1 y HisOxyTx2 (Fig 7.6.1A y 7.6.2A) se observa un mínimo debajo del eje horizontal alrededor de 205 nm. Esta señal se encuentra a valores de longitud de onda mayores a los que se esperaría para una proteína desestructurada pero menores a los que se esperarían para una proteína predominantemente conformada por hojas beta. Esto puede indicar que hay una contribución significativa de ambas estructuras secundarias y por lo tanto la señal esperada se desplaza a un punto medio entre las dos.

Para optimizar la calidad de datos obtenidos por CD, es importante que la absorbancia no sea demasiado alta de manera que impida la correcta penetración de la luz. Esto se puede monitorear utilizando la señal High Tension (HT).⁵⁷ El manual del fabricante para el equipo utilizado (Jasco J-1500) recomienda que la señal HT sea menor a 700 V, de lo contrario se observa demasiado ruido en la señal CD. Para ambos espectros obtenidos se presentan los valores de esta señal a lo largo de la adquisición (Fig 7.6.1B y 7.6.2B), y en ambos casos se

observa que esta señal supera los 700 V entre las longitudes de onda 190 y 200. Por lo tanto, los valores obtenidos en esta región no son confiables y no se pueden utilizar para interpretar la curva obtenida. Por esta razón se presentan los espectros de dicroísmo circular obtenidos a partir del valor de 200 nm de longitud de onda ya que estos son los valores que se pueden considerar confiables según la señal HT.

En la tabla 7.6.1 se presentan los valores de distribución de estructura secundaria predicha a partir de los datos de dicroísmo circular según el servidor Dichroweb. Es importante mencionar que los datos analizados corresponden a todo el rango de longitud de onda utilizado durante la adquisición (190 – 260 nm). Previamente se ha mencionado que los valores obtenidos entre 190 y 200 nm no son confiables y por lo tanto se espera que estos aumenten el error en estas predicciones. Los resultados igualmente se presentan como un punto de partida en el análisis de las estructuras a partir de los experimentos de CD; si se buscara continuar con el análisis se buscaría repetir estos experimentos para obtener resultados más confiables.

Distribución de estructura secundaria predicha por los resultados de CD.			
	Alfa hélice	Hoja Beta	Random Coil
HisOxyTx1	0.05	0.47	0.48
HisOxyTx2	0.05	0.48	0.48

Tabla 7.6.1

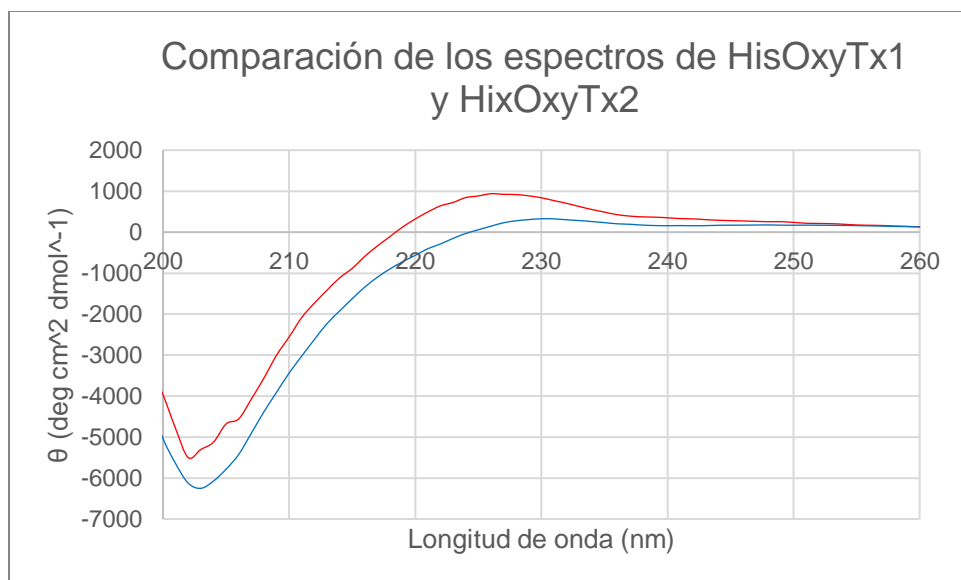


FIG. 7.6.4 Comparación de las curvas de diroísmo circular obtenidas para HisOxyTx1 (azul) y HisOxyTx2 (rojo).

En la figura 7.6.4 se observa que existe una alta similitud entre los espectros y los resultados predichos de porcentajes de estructura secundaria de ambas toxinas. Esto indica que las toxinas posiblemente presentan el mismo tipo de plegamiento. Se observa un porcentaje cercano a cero de alfa hélice, y porcentajes aproximadamente equivalentes de hoja beta y giro (tabla 7.6.1). Esto es consistente con el análisis previamente descrito de las curvas, y consistente también con el motivo estructural ICK esperado.

7.7. Resonancia magnética nuclear

En el caso de HisOxyTx1, no fue posible obtener cantidades suficientes de toxina pura para adquirir un espectro unidimensional a partir de cuerpos de inclusión, por lo tanto, se muestra sólo el espectro correspondiente a la toxina obtenida de la fase soluble.

Instituto de Química, UNAM
Elixa OxyTx1 No Refold
200 ug en 200 uL
24 de mayo de 2019
Pulse Sequence: water_f2

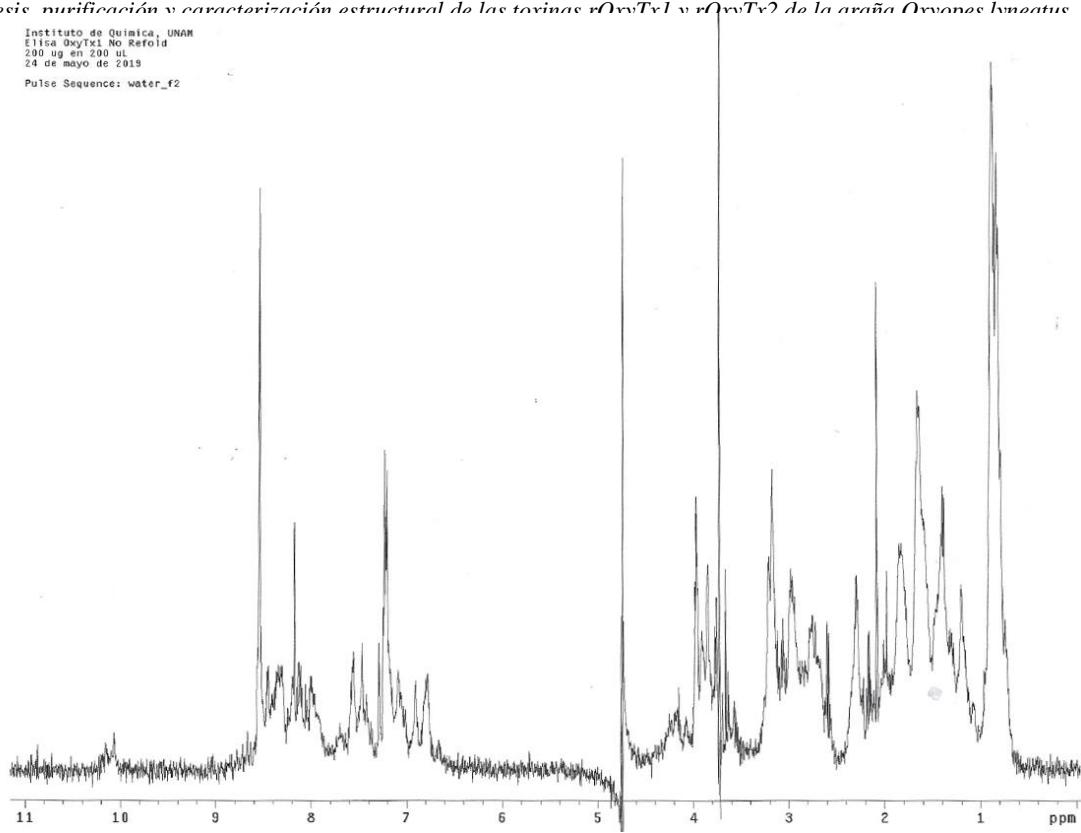


FIG 7.7.1. Espectro unidimensional de ^1H en $\text{H}_2\text{O}/\text{D}_2\text{O}$ correspondiente a HisOxyTx1 purificada de la fase soluble.

La finalidad del experimento unidimensional es determinar cualitativamente la pureza y estado de plegamiento de la proteína. Se puede observar, gracias a la dispersión de señales que existe entre 6.8 y 8.6 ppm en la figura 7.7.1, que la proteína HisOxyTx1 se encuentra en una conformación plegada. Estas señales corresponden a los protones de amidas en la cadena polipeptídica. Cuando una proteína se encuentra desplegada, el estado electrónico de estos protones es sumamente variado y por lo tanto no se puede obtener una sola señal al adquirir el espectro, resultando en señales no resueltas. La presencia de buena resolución en esta región es indicativa de un plegamiento particular de la toxina.

Para el caso de HisOxyTx2, se obtuvieron los espectros unidimensionales de ambas muestras.

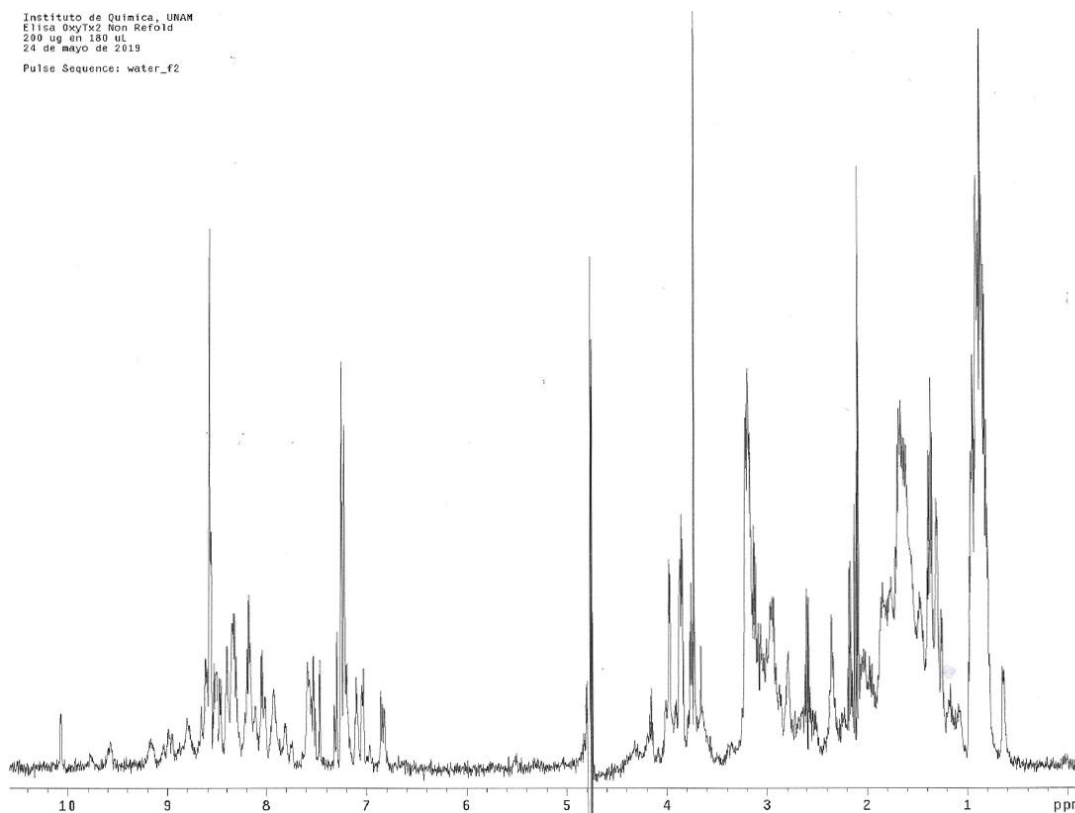


FIG 7.7.2. Espectro unidimensional de ^1H en $\text{H}_2\text{O}/\text{D}_2\text{O}$ correspondiente a HisOxyTx2 purificada de la fase soluble.

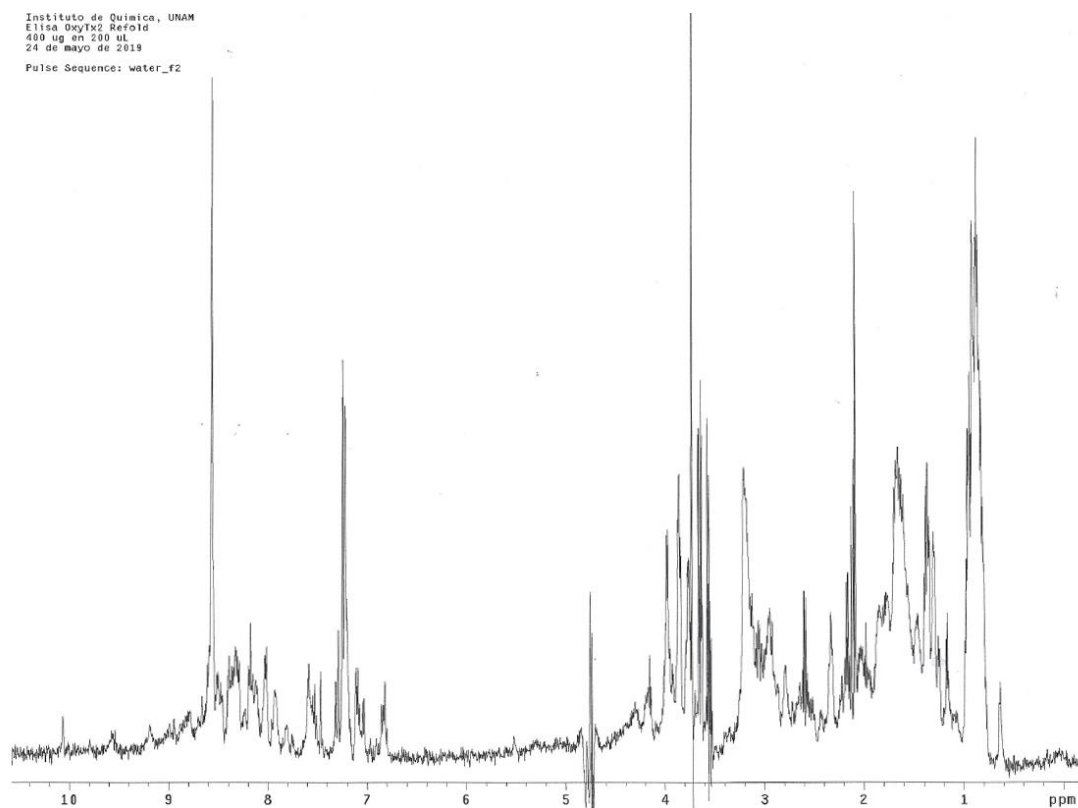


FIG 7.7.3. Espectro unidimensional de ^1H en $\text{H}_2\text{O}/\text{D}_2\text{O}$ correspondiente a HisOxyTx2 purificada en presencia de agente caotrópico y posterior replegamiento.

Ambos espectros presentan buena resolución de las señales correspondientes a los H de amidas, indicando que existe un plegamiento específico de las toxinas. Sin embargo, a pesar de que el espectro de la figura 7.7.3. se adquirió a partir de una muestra de mayor concentración, la intensidad de la señal alrededor de 1 ppm, correspondiente a H de metilos, es menor con respecto a las señales alrededor de 7.2 y 8.5, correspondientes a residuos de histidinas al comparar con la figura 7.7.2. Esta comparación se puede observar en la figura 7.7.3.

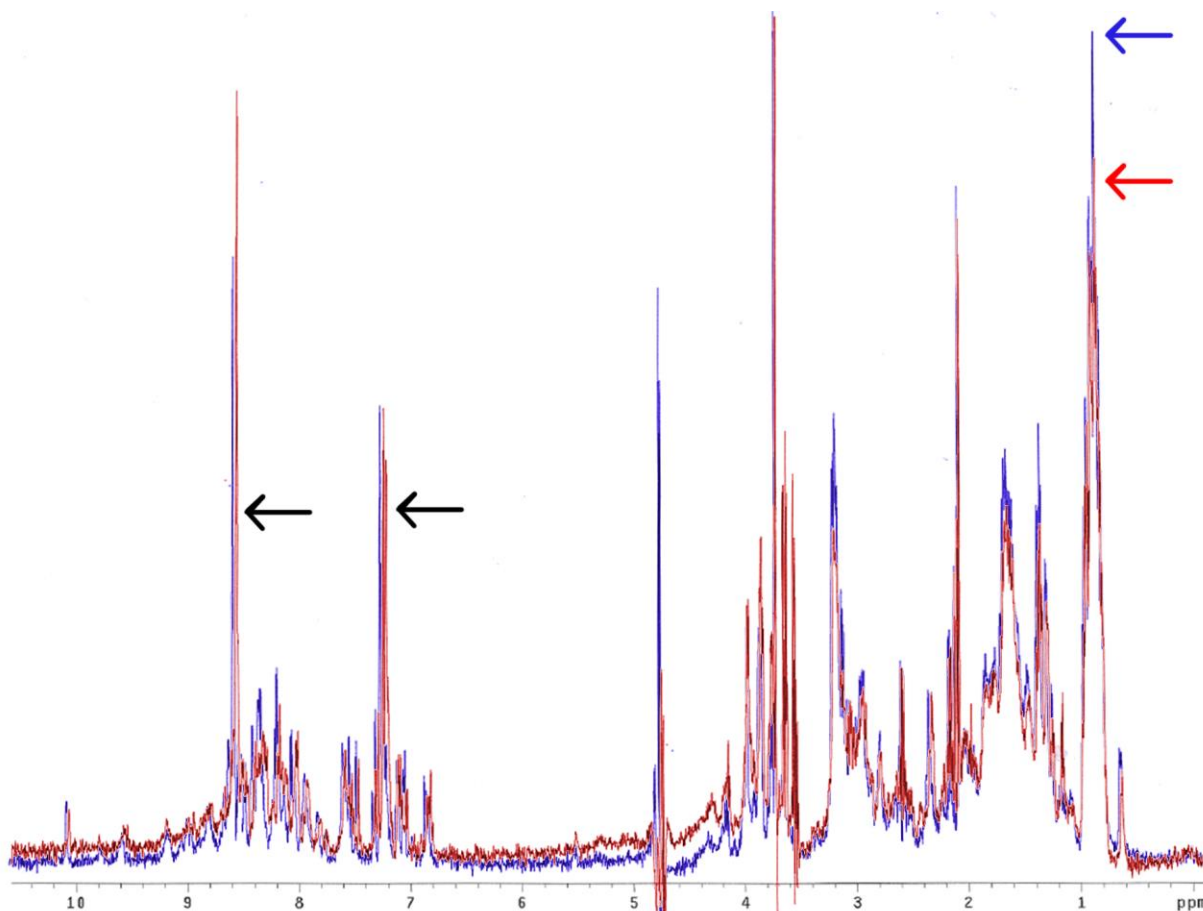


FIG. 7.7.4. Comparación de los espectros de ^1H -RMN correspondientes a HisOxyTx2 purificado de fase soluble (azul) y a partir de cuerpos de inclusión y replegamiento (rojo). Se muestran las señales correspondientes a metilos (flechas azul y roja respectivamente para cada espectro) y las señales correspondientes a His (flechas negras).

Como se pudo observar en la Fig. 7.2.6., al purificar esta toxina en RP-HPLC se generó un pico ancho y con un “hombro” con respecto a la toxina de fase soluble, lo que indicaría la

presencia de especies distintas a las que corresponden al pico mayoritario. Esto se muestra en el espectro unidimensional, ya que, a pesar de presentar señales dispersas, son de menor intensidad, lo que sugiere la presencia de múltiples especies con plegamientos diferentes.

Ambos espectros correspondientes a HisOxyTx2 presentan aproximadamente las mismas señales (Fig 7.7.4). Esto indica que la toxina purificada de la fase soluble y la que fue sometida a un proceso de replegamiento presentan un plegamiento definido, y que el plegamiento es similar.

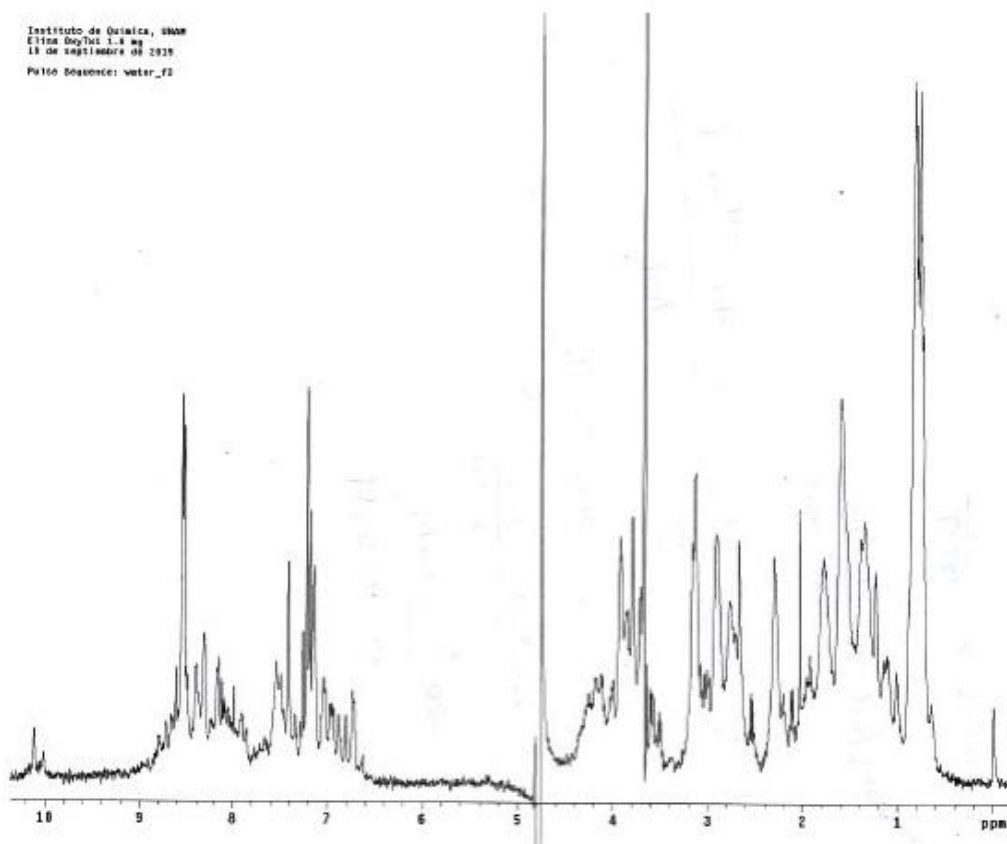


FIG 7.7.5. Espectro unidimensional de ^1H correspondiente a HisOxyTx1 purificado de fase soluble obtenida a partir del proceso de escalamiento de expresión en $\text{H}_2\text{O}/\text{D}_2\text{O}$.

Después de realizar el proceso de escalamiento de la expresión de las toxinas HisOxyTx1 y HisOxyTx2, se volvieron a realizar experimentos unidimensionales de ^1H de ambos péptidos

para observar el plegamiento de las toxinas recién purificadas a partir de una reacción de mayor volumen donde se obtuvo mejor rendimiento.

En el espectro de la figura 7.7.5, correspondiente a HisOxyTx1, se observa una buena distribución de las señales correspondientes a protones de amida, lo cual indica que la nueva muestra se encuentra plegada. Además, en comparación con los espectros anteriores, se observa que la intensidad de las señales correspondientes a histidinas es menor con respecto a las señales correspondientes a metilos. Esto puede ser efecto de la concentración de la muestra y puede indicar la presencia de un solo conformero del péptido.

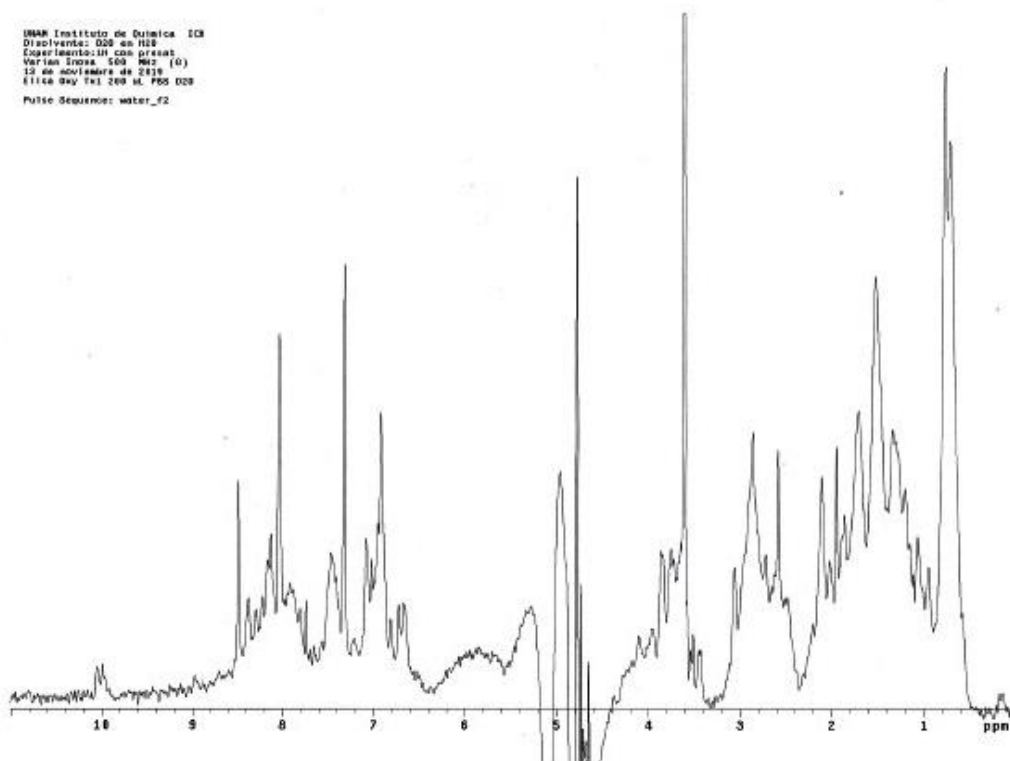


FIG 7.7.6. Espectro unidimensional de ^1H correspondiente a HisOxyTx1 en PBS/D₂O.

Posteriormente se repitió el experimento unidimensional de ^1H con la misma muestra de HisOxyTx1 utilizando el amortiguador PBS como disolvente para determinar si se presenta una variación en el plegamiento de la proteína por la presencia de sales en el medio (Fig 7.7.6).

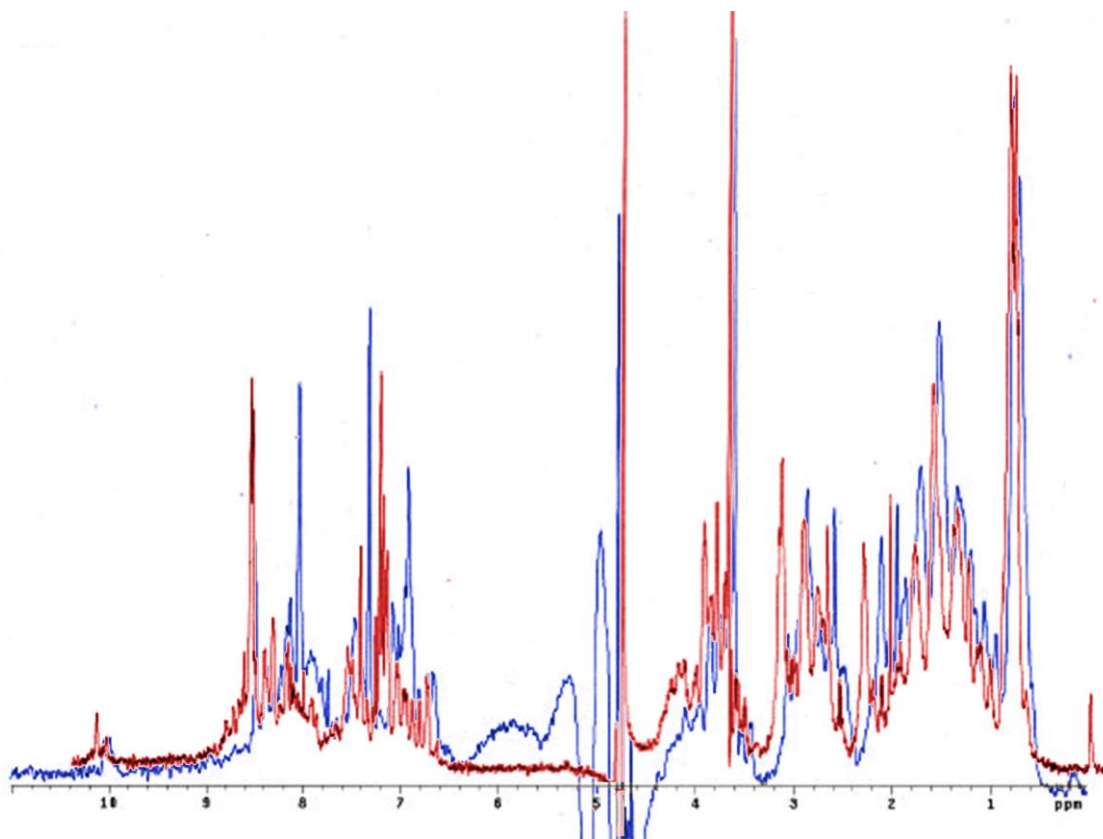


FIG. 7.7.7. Comparación de los espectros de HisOxyTx1 purificada a partir de fase soluble en H₂O/D₂O (rojo) y PBS (azul).

No se observa una mejor definición de las señales de amida (ni se mantiene, se observa una menor definición) cuando la muestra se disuelve en PBS (Fig 7.7.7). Este amortiguador tiene un pH de 6.7, lo cual es más bajo de lo que se esperaría para la disolución de la muestra en agua desionizada. Esta puede ser la razón por la baja definición de las señales lo cual indica que la toxina a este pH es menos estable y su plegamiento no está bien definido.

En el caso de la toxina HisOxyTx2 después de realizar el proceso de escalamiento de la expresión y purificar la parte soluble se observa menor definición de las señales de protones de amida en la figura 7.7.8 en H₂O/D₂O con respecto al espectro anterior (purificación realizada en matraces a 30°C aprox, fig 7.7.2). Sin embargo, las señales correspondientes a histidinas presentan menor intensidad con respecto a la señal correspondiente a metilos, lo cual puede indicar la presencia de una sola especie en lugar de una combinación de distintos conformeros. Esta comparación se puede apreciar en la figura 7.7.9.

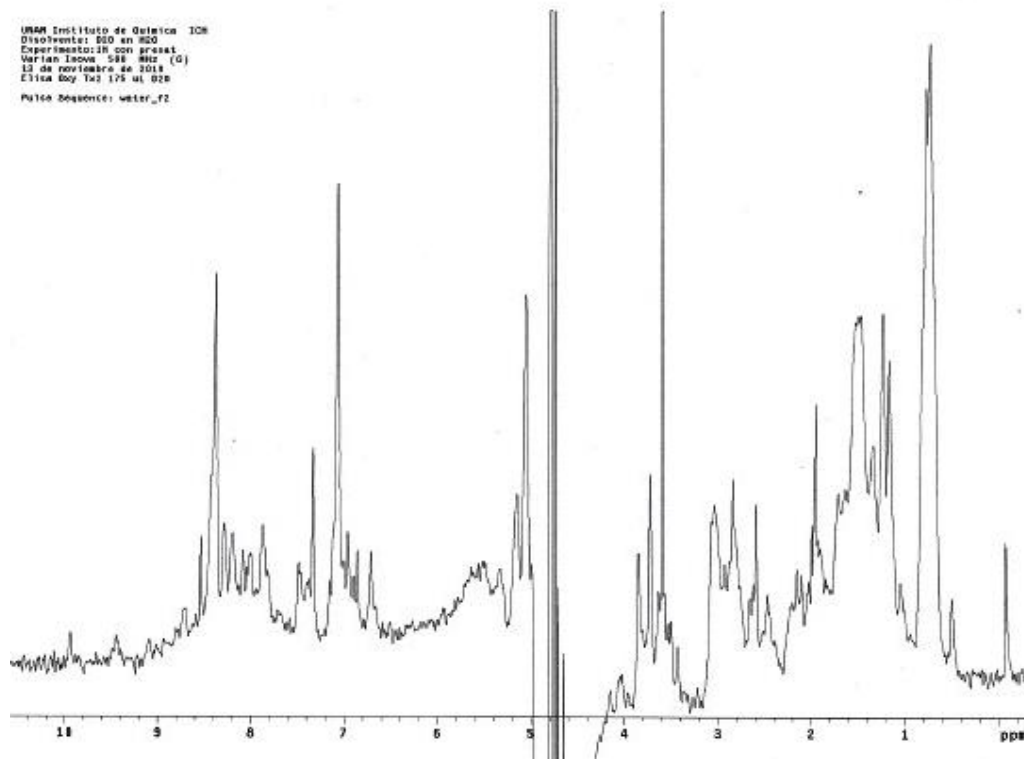


FIG 7.7.8. Espectro unidimensional de ^1H correspondiente a HisOxTx2 en $\text{H}_2\text{O}/\text{D}_2\text{O}$.

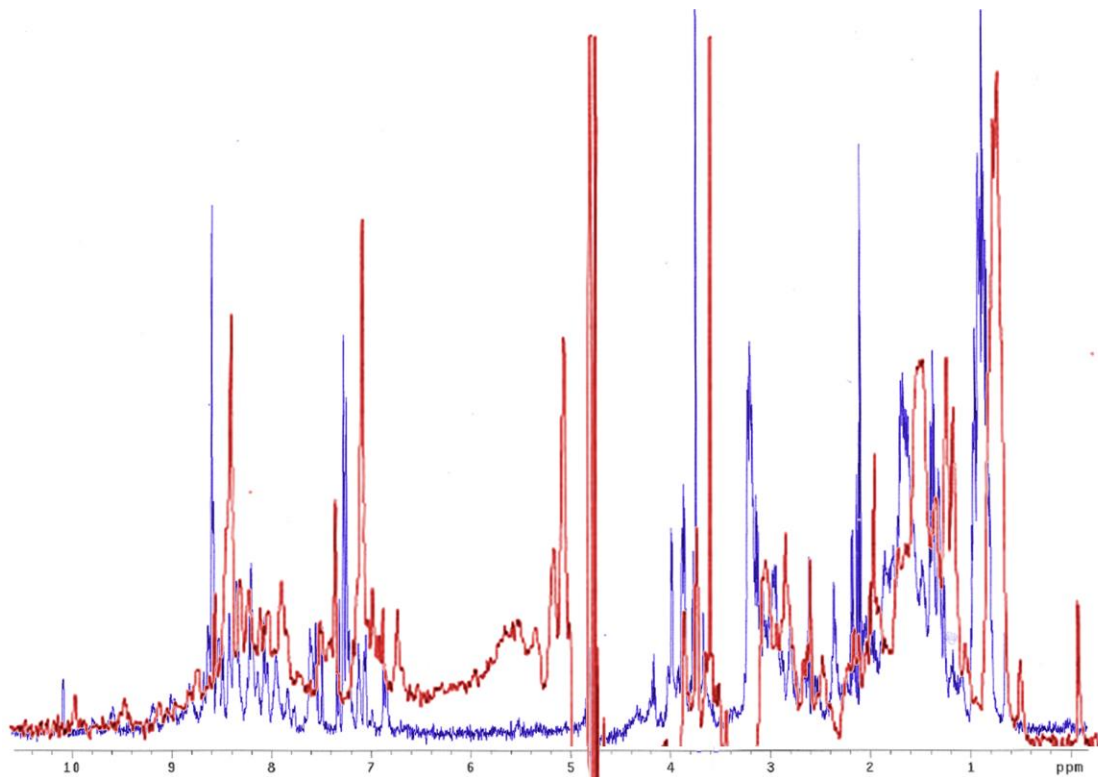


FIG. 7.7.9. Comparación de los espectros de ^1H -RMN correspondientes a HisOxTx2 obtenidos de fase soluble expresada en matraces (azul) y fase soluble a partir de escalamiento de expresión a reactores (rojo).

Para poder adquirir los experimentos de RMN de dos dimensiones, se requiere una concentración mínima de 50 mM según la literatura. Para lograr esta cantidad, se requieren alrededor de 2.5 mg de cada una de las toxinas de interés. Por causas relacionadas a la pandemia de Covid-19, no fue posible obtener mayor cantidad de estas toxinas y por lo tanto no se logró adquirir los experimentos bidimensionales necesarios para determinar la estructura experimental de las toxinas.

7.8. Predicción de estructura tridimensional por modelaje por homología

7.8.1 Determinación de la proteína plantilla.

Las toxinas de araña, cuya estructura se reporta en la literatura, se dividen en dos grupos: citotoxinas lineales (*Lachesana tarabaevi*⁵⁹, *Acanthoscurria gomesian*⁶⁰, *Lasiadora sp*⁶¹) y neurotoxinas con el motivo estructural ICK (nudo inhibitorio de cisteínas) (*Psalmopoeus cambridgei*⁶²). Al llevarse a cabo la búsqueda de proteínas con similitud a las proteínas de interés, se obtuvo el dominio C-terminal de la toxina OtTx1a de la araña *Oxyopes takobius*, también conocida como Spiderine1a.⁴⁴

OtTx1a es una proteína modular que presenta dos dominios: una secuencia lineal N-terminal con actividad citolítica (OtTx1a-AMP) y una secuencia C-terminal rica en cisteínas con un motivo estructural ICK (OtTx1a-ICK). Este segundo dominio es el que presenta alta similitud a las Oxitoxinas y será utilizado como proteína plantilla. Su estructura determinada por RMN está reportada en el PDB con el ID 2N86.⁶³

Score	Expect	Method	Identities	Positives	Gaps
44.3 bits	0.005	Compositional matix adjust	25/62 (40%)	33/62 (53%)	14/62 (22%)
OxyTx1	3 ECLPLHSSCDNDCVCCCKNHHCPCYSN VSKLEKWLPEWAKIPDALK RCSCQRNDKDKGKINTC 64				
	+C +++C +DC CC HHCHCP KWLP LRC+CQ +D K+N C				
2N86	8 KCWAIGTTCSDDCDCPEHHCHCPAG-----KWLP-----GLFRCTCQVTESD-KVNKC 55				

Tabla 7.8.1 Alineación de secuencias de OxyTx1 y OtTx1a

Score	Expect	Method	Identities	Positives	Gaps
52 bits	8e-11	Compositional matix adjust	25/49 (51%)	31/49 (63%)	1/49 (2%)
OxyTx2	3	KCLPKDSTCGDDCDCCEGLHCHCPLRNMLPAILRCSCQ-SKDDHINTCP	50		
		KC +TC DDCDCC HCHCP LP + RC+CQ ++ D +N CP			
2N86	8	KCWAIGTTCSDDCDCCEHHCHCPAGKWLPLGLFRCTCQVTESDKVNKCP	56		

Tabla 7.8.2 Alineación de secuencias de OxyTx2 y OtTx1a

OtTx1a-ICK presenta una identidad de 40 y 51% con OxyTx1 y OxyTx2 respectivamente. Otra similitud importante de notar es la presencia de 10 cisteínas en las tres secuencias. La estructura reportada de OtTx1a-ICK presenta 5 puentes disulfuro; se espera que este también sea el caso para las Oxitoxinas, según lo que se ha observado al obtener su masa molecular experimental por MALDI-TOF. Por esta razón, se esperaría que el motivo estructural de OxyTx1 y OxyTx2 sea similar al reportado para OtTx1a-ICK.

7.8.2 Evaluación de la proteína plantilla.

Para evaluar las dos estructuras, se utiliza el servidor MolProbity⁶⁴. MolProbity es un servidor de validación de estructuras que se puede utilizar para evaluar modelos estructurales de proteínas o ácidos nucleicos. Se basa en el análisis de todos los átomos de la molécula y la colocación optimizada de hidrógenos, junto con versiones actuales de los criterios de geometría covalente y ángulos de torsión. Detecta errores que pueden afectar la reconstrucción manual del modelo, tales como valores extremos de Ramachandran, cadenas laterales invertidas y posiciones incorrectas de grupos funcionales.

Todos los servidores, cuya propuesta se estudiará a continuación, utilizan la estructura de OtTx1a-ICK obtenida por RMN (PDB 2N86) como plantilla. Como primer paso, se realizó la evaluación de esta estructura por MolProbity para tomarlo como punto de comparación.

Evaluación MolProbity 2N86

Evaluación de PDB 2N86 por MolProbity.			
Contactos	Clashcore (todos los átomos)	2.41	
Geometría de la	Rotámeros pobres	12	23.53%
	Rotámeros favorecidos	28	54.90%

proteína	Valores extremos Ramachandran	3	5.26%
	Valores favorecidos Ramachandran	42	73.68%
	C β desviaciones >0.25Å	0	0.00%
	Enlaces erróneos:	0 / 450	0.00%
	Ángulos erróneos:	0 / 614	0.00%

Tabla 7.8.3

7.8.3 Modelos: OxyTx1

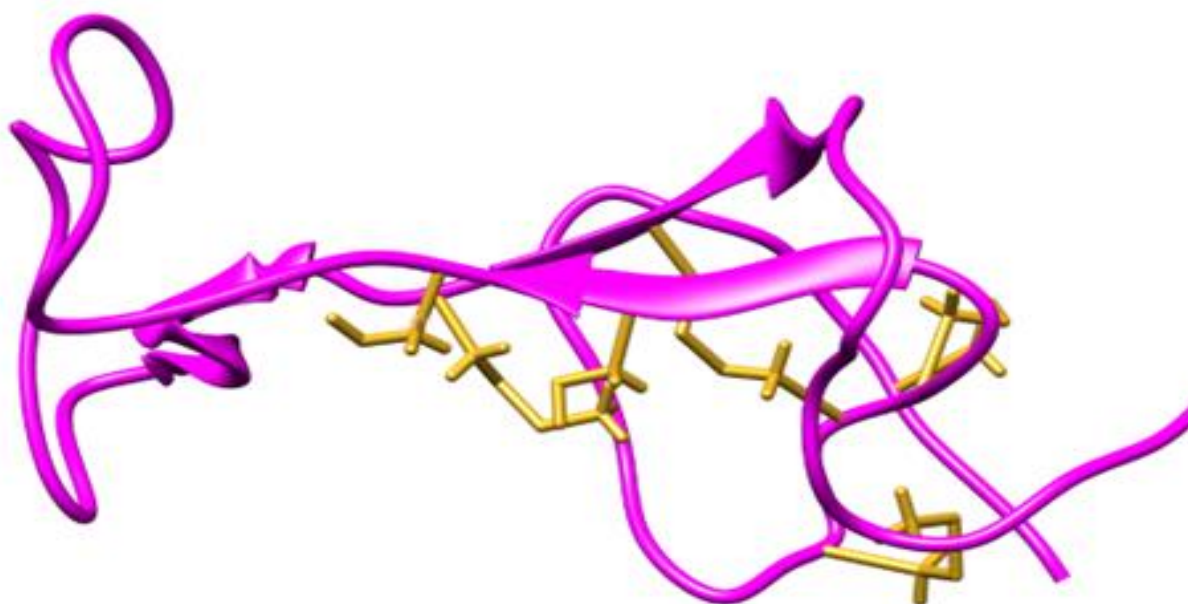
I-Tasser:

FIG 7.8.1 Modelo de OxyTx1 obtenido por I-Tasser. En amarillo se muestran los residuos cisteína, de los cuales 8 están formando enlaces disulfuro.

I-Tasser se considera una buena opción para casos como el de las Oxitoxinas (donde el porcentaje de identidad es relativamente bajo) ya que asume al iniciar el modelado que toda secuencia introducida corresponde a una proteína estructurada y en regiones largas desestructuradas, busca proponer una estructura por modelado *ab initio*. Sin embargo, la presencia de regiones significativas desestructuradas degrada la calidad del modelo, por lo que es importante evaluar cuidadosamente las propuestas de este servidor cuando se trabaja con proteínas con regiones de este tipo.

I-Tasser predice la presencia de regiones con estructura secundaria tipo hélice (Fig 7.8.1), en especial en la región de los aa 30-41. Esta región, como se mencionará más adelante, se encuentra dentro del bucle entre hebras beta en el modelo tridimensional que no tiene correspondencia en la molécula molde y por lo tanto se modela de maneras diferentes en cada uno de los servidores utilizados. En este caso, se puede observar en el modelo tridimensional la presencia de un fragmento tipo hélice. No se han reportado ejemplos de toxinas de araña que presenten el motivo estructural ICK y contengan hélices; además, los resultados de difracción circular son consistentes con una alta proporción de hoja beta y una proporción despreciable de hélice. Por estas razones, es razonable concluir que la hélice presente en el modelo obtenido por este servidor no existe en la estructura real y es una manera del algoritmo de estabilizar una zona poco estable de la secuencia modelada.

I-Tasser es el único de los tres modelos obtenidos que predice la existencia de cuatro enlaces disulfuro en lugar de cinco. La conectividad de tres de estos enlaces coincide con la obtenida por los otros modelos. Gracias a los resultados obtenidos por espectrometría de masas MALDI-TOF, se sabe que la estructura de la toxina OxyTx1 obtenida por métodos experimentales contiene todas sus cisteínas en la forma oxidada y por lo tanto cinco enlaces disulfuro formados

Evaluación de ITasser-OxyTx1 por MolProbity.			
Contactos	Clashcore (todos los átomos)	8.34	
Geometría de la proteína	Rotámeros pobres	15	22.73%
	Rotámeros favorecidos	40	60.61%
	Valores extremos Ramachandran	16	23.88%
	Valores favorecidos Ramachandran	35	52.24%
	C β desviaciones >0.25Å	11	16.18%
	Enlaces erróneos:	0 / 573	0.00%
	Ángulos erróneos:	26 / 774	3.36%

Tabla 7.8.4

Se utilizó también el servidor PIC (Protein Interaction Calculator)⁶⁵ para determinar interacciones hidrofóbicas, así como enlaces disulfuro basados en el modelo tridimensional. Se comparan las interacciones que se repiten entre los modelos de la siguiente manera:

- En rojo las que no se repiten en ningún modelo
- En amarillo las que se repiten entre dos modelos
- En verde las que se repiten entre los tres modelos.

Este código de colores se utilizará en todas las tablas que listen enlaces disulfuro o interacciones hidrofóbicas para todos los modelos de ambas toxinas.

Conectividad de puentes disulfuro en ITasser-OxyTx1.				
Posición	Residuo	Posición	Residuo	Distancia (Å)
11	CYS	50	CYS	2.06
15	CYS	64	CYS	2.03
17	CYS	52	CYS	2.04
4	CYS	18	CYS	2.05
2-8, 3-10, 4-9, 1-5 C23 (6) y C25 (7) no aparecen formando un enlace disulfuro.				

Tabla 7.8.5 A cada cisteína se le otorga un número del 1 - 10 según su posición en la secuencia primaria de la proteína. La conectividad de cisteínas se presenta usando estos números como referencia.

Interacciones hidrofóbicas

Evaluación de interacciones hidrofóbicas de ITasser-OxyTx1 por PIC			
Posición	Residuo	Posición	Residuo
2	TRP	67	TYR
5	LEU	16	VAL
5	LEU	6	PRO
26	PRO	47	LEU
27	TYR	40	TRP
27	TYR	46	ALA
27	TYR	47	LEU
33	LEU	36	TRP
33	LEU	47	LEU
37	LEU	38	PRO
37	LEU	43	ILE
37	LEU	47	LEU
40	TRP	41	ALA
43	ILE	44	PRO
46	ALA	47	LEU

Tabla 7.8.6.

A pesar de que el modelo obtenido por I-Tasser tiene el menor valor de Clashcore (tabla 7.8.5) y por lo tanto se puede considerar el más favorecido de los obtenidos, no se utilizó como punto de partida para el refinamiento de las estructuras por dinámica molecular ya que su

conectividad de enlaces disulfuro no corresponde con la que se espera según los resultados experimentales. Una dinámica molecular no permite la ruptura o formación de enlaces y por lo tanto no sería capaz de corregir este error de conectividad.

SwissModel

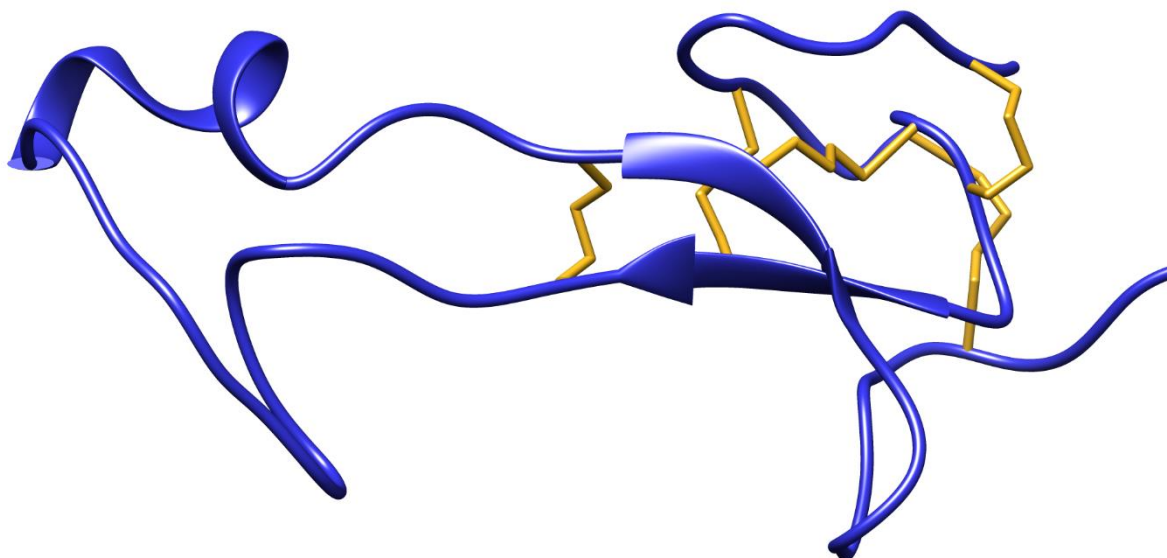


FIG 7.8.2 Modelo de OxyTx1 obtenido por SwissModel. En amarillo se muestran los residuos cisteína. Todas se encuentran formando enlaces disulfuro.

Evaluación de SwissModel-OxyTx1 por MolProbity.			
Contactos	Clashcore (todos los átomos)	12.96	
Geometría de la proteína	Rotámeros pobres	7	11.29%
	Rotámeros favorecidos	49	79.03%
	Valores extremos Ramachandran	3	4.76%
	Valores favorecidos Ramachandran	48	76.19%
	C β desviaciones >0.25Å	2	3.12%
	Enlaces erróneos:	0 / 533	0.00%
	Ángulos erróneos:	15 / 720	2.08%

Tabla 7.8.7

SwissModel ha recibido evaluaciones favorables de proteínas generadas a partir de plantillas cercanas en comparación con otros servidores según la evaluación CAMEO (Evaluación Continua de Modelos Automatizados). Sin embargo, para proteínas que no tienen

plantillas cercanas o modelaje de novo, no es el método mejor evaluado. En el caso de las Oxitoxinas, existen regiones de la secuencia con alta similitud y otras con baja, así que, se esperaría que el servidor genere mejores resultados para las de alta similitud y la calidad del modelo baje cuando se trate de regiones de baja similitud.

Para OxyTx1, SwissModel predice una estructura tipo ICK con cinco enlaces disulfuro formados. La región que tiene correspondencia con la plantilla es muy similar a la modelada por I-Tasser, pero se observa un cambio en la región desestructurada entre hebras beta. SwissModel sí presenta una estructura tipo hélice en esta región, tal como se predice en la estructura secundaria en I-Tasser incluso cuando no se observa en el modelo tridimensional. El resto de la región entre hebras se observa extendida, esto puede deberse a que el servidor busca alejar los átomos entre sí para evitar choques.

Enlaces disulfuro

Enlaces disulfuro predichos por el servidor SwissModel para OxyTx1				
Posición	Residuo	Posición	Residuo	Distancia (Å)
11	CYS	23	CYS	2.02
15	CYS	64	CYS	2.02
17	CYS	52	CYS	2.03
25	CYS	50	CYS	2.03
4	CYS	18	CYS	2.04
2-6, 3-10, 4-9, 7-8, 1-5				

Tabla 7.8.8. A cada cisteína se le otorga un número del 1 - 10 según su posición en la secuencia primaria de la proteína. La conectividad de cisteínas se presenta usando estos números como referencia. Esta conectividad de cisteínas se repite también para el modelo obtenido por RaptorX, por lo tanto, se presenta sólo en este caso.

Interacciones hidrofóbicas

Evaluación de interacciones hidrofóbicas de SwissModel-OxyTx1 por PIC.			
Posición	Residuo	Posición	Residuo
5	LEU	16	VAL
5	LEU	6	PRO
27	TYR	30	VAL
27	TYR	47	LEU
30	VAL	43	ILE
30	VAL	46	ALA
30	VAL	47	LEU

36	TRP	38	PRO
37	LEU	38	PRO
40	TRP	41	ALA
43	ILE	44	PRO
43	ILE	46	ALA
46	ALA	47	LEU

Tabla 7.8.9

RaptorX

FIG 7.8.3 Modelo de OxyTx1 obtenido por RaptorX. En amarillo se muestran los residuos cisteína. Todas se encuentran formando enlaces disulfuro.

Al ser evaluado en el proyecto CASP9 (Evaluación Crítica de Técnicas para Predicción de Estructura de Proteínas), RaptorX obtuvo la segunda calificación más alta (los servidores del grupo de Zhang (eg I-Tasser) obtuvieron la calificación más alta) a partir de la evaluación de alrededor de 10 programas de modelado. Sin embargo, para el caso de las proteínas consideradas más difíciles de modelar, RaptorX obtuvo la calificación más alta. RaptorX requiere por lo menos un modelo de similitud baja, ya que el algoritmo no produce buenos resultados en el caso de modelaje *de novo*.

De los modelos analizados, RaptorX es el único que no genera una estructura secundaria tipo hélice en la región correspondiente al giro entre hebras beta. Al tener entonces una región relativamente larga de la secuencia sin una estructura definida que puede generar una gran cantidad de choques con otras regiones de la proteína, el servidor parece alargar esta región de la estructura para minimizar las interacciones con el resto de la proteína. Es importante notar que, de los tres modelos obtenidos, RaptorX obtuvo la calificación más baja del criterio Clashcore

que mide los traslapes estéricos presentes en la proteína. Esto probablemente se debe a que esta región es desestructurada.

Evaluación de RaptorX-OxyTx1 por MolProbity			
Contactos	Clashcore (todos los átomos)	76.99	
Geometría de la proteína	Rotámeros pobres	3	4.55%
	Rotámeros favorecidos	57	86.36%
	Valores extremos Ramachandran	1	1.49%
	Valores favorecidos Ramachandran	60	89.55%
	C β desviaciones >0.25Å	2	2.94%
	Enlaces erróneos:	0 / 574	0.00%
	Ángulos erróneos:	15 / 776	1.93%

Tabla 7.8.10

Interacciones hidrofóbicas

Evaluación de interacciones hidrofóbicas de RaptorX-OxyTx1 por PIC.			
Posición	Residuo	Posición	Residuo
2	TRP	67	TYR
5	LEU	16	VAL
5	LEU	6	PRO
27	TYR	33	LEU
27	TYR	37	LEU
27	TYR	40	TRP
27	TYR	47	LEU
36	TRP	37	LEU
36	TRP	38	PRO
37	LEU	38	PRO
40	TRP	47	LEU
41	ALA	43	ILE
43	ILE	44	PRO

Tabla 7.8.11

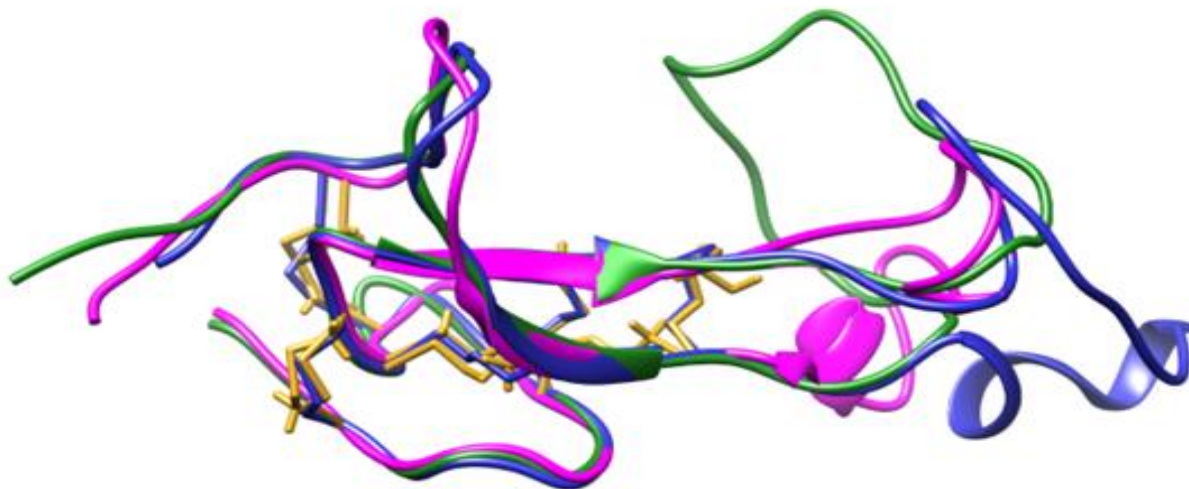


FIG 7.8.4 Comparación visual de los tres modelos obtenidos para la estructura de OxyTx1. ITasser: magenta, RaptorX: verde, SwissModel: azul.

La toxina OtTx1a es un péptido de 59 aminoácidos. Esta longitud es similar a la de OxyTx2, pero considerablemente más corta que OxyTx1. Por esta misma razón, hay una región de OxyTx1 correspondiente a los aminoácidos 26-49 que no está bien representada en la molécula molde.

Score	Expect	Method	Identities	Positives	Gaps
44.3 bits	0.005	Compositional matrix adjust	25/62 (40%)	33/62 (53%)	14/62 (22%)
OxyTx1	3 ECLPLHSSCDNDCVCCCKNHHCHCPYSN VSKLEKWLPEWAKIPDALKRCSCQRNDKDGKINTC 64 +C ++C +DC CC HHCHCP KWLP LRC+CQ +D K+N C				
2N86	8 KCWAIGTTCSDDCDCPEHHCHCPAG-----KWLP-----GLFRCTCQVTESD-KVNKC 55				

Tabla 7.8.12 Comparación de las secuencias de aminoácidos de OxyTx1 y OtTx1a. En rojo se marca la región correspondiente al giro entre hebras beta en los tres modelos.

Comparación de los valores de RMS entre los modelos de OxyTx1 obtenidos divididos por regiones de la toxina.			
Modelos	RMS molécula completa	RMS hebras	RMS lazo entre hebras
ITasser-SwissModel	4.937	1.587	6.059
ITasser-RaptorX	6.351	1.535	7.998
SwissModel-RaptorX	7.618	0.805	8.148

Tabla 7.8.13

En la tabla 7.8.13 se presentan valores de RMS entre los modelos por regiones de la toxina. En este caso se utiliza RMS ya que se realiza una comparación uno a uno entre cada par

de toxinas (a diferencia de un valor de RMSD, donde se compara una estructura con un promedio de estructuras). Al comparar los tres modelos obtenidos, se observa una alta identidad estructural en las regiones correspondientes a las hojas beta, así como los extremos (tabla 7.8.12). Sin embargo, el giro entre las dos hebras beta, al no tener una correspondencia directa con la plantilla, es estructurada de maneras diferentes. En el caso de RaptorX se observa que esta región se extiende, alejándose del resto de los aminoácidos. En el caso de I-Tasser, al contrario, existe una aparente contracción de la estructura. SwissModel predice la presencia de un motivo tipo hélice en esta región.

Las regiones hacia los extremos de la proteína, las cuales sí tienen una región correspondiente en el molde, se modelan de manera muy similar en los tres casos. Las regiones N-terminal son muy similares, mientras que las C-terminal muestran mayor variación. La región de las hebras beta es muy similar en los tres modelos. RaptorX y SwissModel presentan la misma conectividad de enlaces disulfuro, mientras que I-Tasser muestra dos cisteínas no enlazadas (C23 y C25) y un enlace disulfuro que no se muestra en las otras estructuras (C11-C50). Esto indica que el servidor probablemente predice la presencia de ángulos de rotación de las cisteínas que no existen en la molécula real, ya que los experimentos previamente realizados de MALDI-TOF muestran que la proteína contiene 5 pares de cisteínas enlazadas.

7.8.4 Modelos: OxyTx2

ITasser

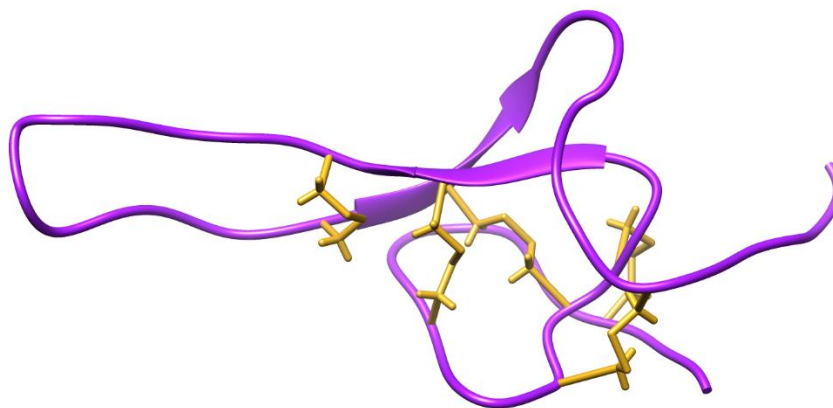


FIG 7.8.5 Modelo de OxyTx2 obtenido por I-Tasser. En amarillo se muestran los residuos cisteína. Todas se encuentran formando enlaces disulfuro.

Evaluación de ITasser-OxyTx2 por MolProbit			
Contactos	Clashcore (todos los átomos)	4.87	
Geometría de la proteína	Rotámeros pobres	8	15.69%
	Rotámeros favorecidos	36	70.59%
	Valores extremos Ramachandran	5	9.43%
	Valores favorecidos Ramachandran	35	66.04%
	C β desviaciones >0.25Å	3	5.66%
	Enlaces erróneos:	0 / 431	0.00%
	Ángulos erróneos:	8 / 580	1.38%

Tabla 7.8.14

Para OxyTx2, ITasser también predice la existencia de regiones tipo hélice en la predicción de estructura secundaria. Al igual que en el caso de OxyTx1, el α -hélice corresponde al giro entre hebras y confirma que es la manera en la que el algoritmo busca estabilizar una región desordenada. Sin embargo, no se observa la presencia de esta hélice en el modelo tridimensional, sino un motivo estructural tipo ICK similar al de la plantilla.

Enlaces disulfuro predichos por el servidor ITasser para OxyTx2.				
Posición	Residuo	Posición	Residuo	Distancia (Å)
11	CYS	23	CYS	2.04
15	CYS	49	CYS	2.01
17	CYS	39	CYS	1.99
25	CYS	37	CYS	2.05
4	CYS	18	CYS	2.04
2-6, 3-10, 4-9, 7-8, 1-5				

Tabla 7.8.15 A cada cisteína se le otorga un número del 1 - 10 según su posición en la secuencia primaria de la proteína. La conectividad de cisteínas se presenta usando estos números como referencia. Esta conectividad de cisteínas se repite para todos los modelos de OxyTx2 obtenidos por servidores, por lo tanto, solo se presenta en este primer caso.

Interacciones hidrofóbicas

Evaluación de interacciones hidrofóbicas de ITasser-OxyTx2 por PIC.			
Posición	Residuo	Posición	Residuo
1	ALA	2	TRP
5	LEU	6	PRO
26	PRO	35	LEU
27	LEU	34	ILE
30	MET	31	LEU

31	LEU	32	PRO
31	LEU	33	ALA

Tabla 7.8.16

SwissModel

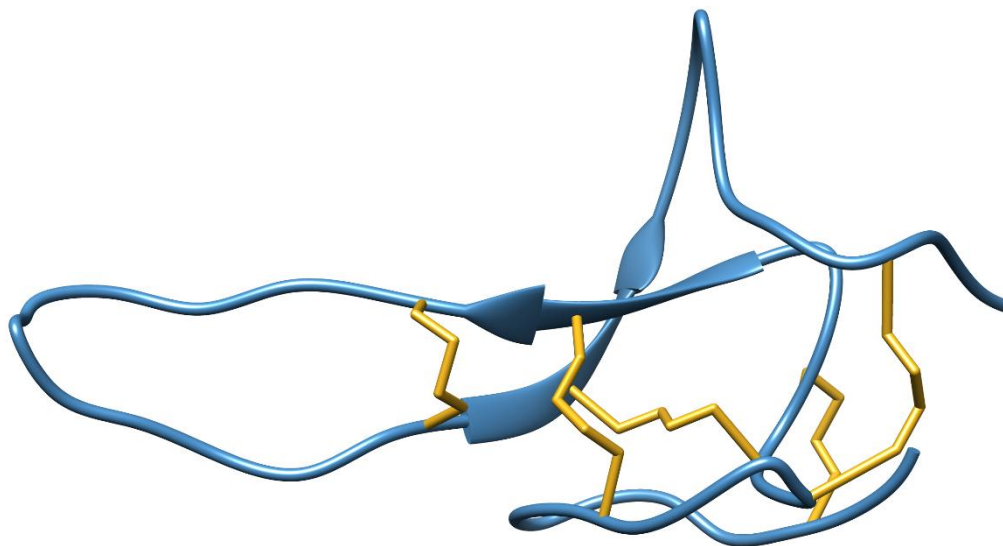


FIG 7.8.6 Modelo de OxyTx2 obtenido por SwissModel. En amarillo se muestran los residuos cisteína. Todas se encuentran formando enlaces disulfuro.

Evaluación de SwissModel-OxyTx2 por MolProbity			
Contactos	Clashcore (todos los átomos)	6.62	
Geometría de la proteína	Rotámeros pobres	4	8.33%
	Rotámeros favorecidos	38	79.17%
	Valores extremos Ramachandran	4	8.16%
	Valores favorecidos Ramachandran	34	69.39%
	C β desviaciones >0.25Å	5	10.20%
	Enlaces erróneos:	0 / 403	0.00%
	Ángulos erróneos:	11 / 545	2.02%

Tabla 7.8.17

Para OxyTx2 en el servidor SwissModel se observa nuevamente un motivo estructural ICK y 5 enlaces disulfuro formados con la misma conectividad que los otros dos modelos obtenidos.

Interacciones hidrofóbicas

Evaluación de interacciones hidrofóbicas de SwissModel-OxyTx2 por PIC			
Posición	Residuo	Posición	Residuo
2	TRP	52	TYR
5	LEU	6	PRO
27	LEU	34	ILE
31	LEU	32	PRO

Tabla 7.8.18

RaptorX

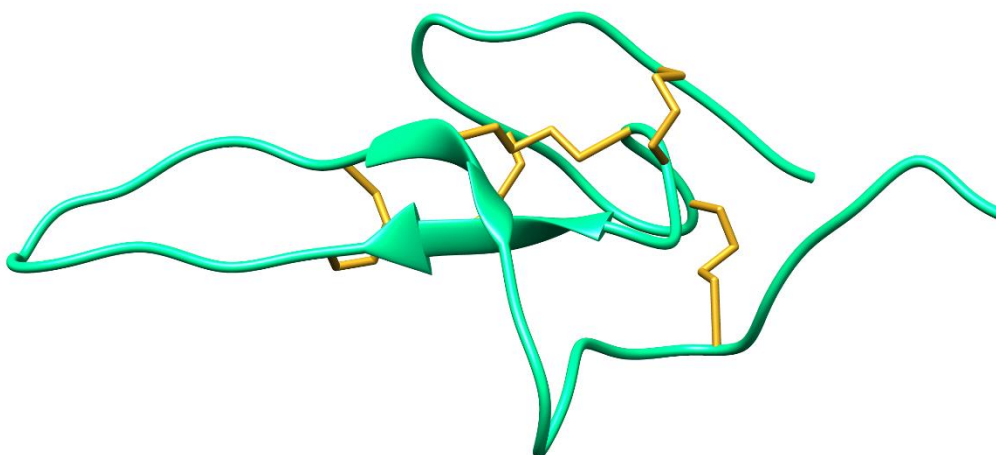


FIG 7.8.7 Modelo de OxyTx2 obtenido por RaptorX. En amarillo se muestran los residuos cisteína. Todas se encuentran formando enlaces disulfuro.

Evaluación de RaptorX-OxyTx2 por MolProbit			
Contactos	Clashcore (todos los átomos)	75.33	
Geometría de la proteína	Rotámetros pobres	2	3.92%
	Rotámetros favorecidos	43	84.31%
	Valores extremos Ramachandran	5	9.43%
	Valores favorecidos Ramachandran	46	86.79%
	C β desviaciones >0.25Å	0	0.00%
	Enlaces erróneos:	0 / 432	0.00%
	Ángulos erróneos:	13 / 582	2.23%

Tabla 7.8.19

La estructura tridimensional de OxyTx2 que predice RaptorX es muy similar a las dos analizadas anteriormente, pero su calificación de Clashcore (tabla 7.8.20) es considerablemente más alta (es decir, no es una buena estructura) que la de los otros modelos. No se observan mayores cambios en la cadena principal. Otro resultado notable es que el modelo de RaptorX, según la evaluación de PIC (tabla 7.8.21), solamente presenta interacciones hidrofóbicas entre aminoácidos cercanos, a diferencia de los otros dos modelos, los cuales presentan por lo menos una interacción hidrofóbica entre aminoácidos cercanos a los extremos. Es posible que existan rotaciones desfavorables de cadena secundaria en estos aminoácidos que no permitan que existan estas interacciones.

Interacciones hidrofóbicas

Evaluación de interacciones hidrofóbicas de RaptorX-OxyTx2 por PIC.			
Posición	Residuo	Posición	Residuo
1	ALA	52	TYR
2	TRP	52	TYR
5	LEU	6	PRO
27	LEU	34	ILE
31	LEU	32	PRO
50	PRO	52	TYR

Tabla 7.8.20

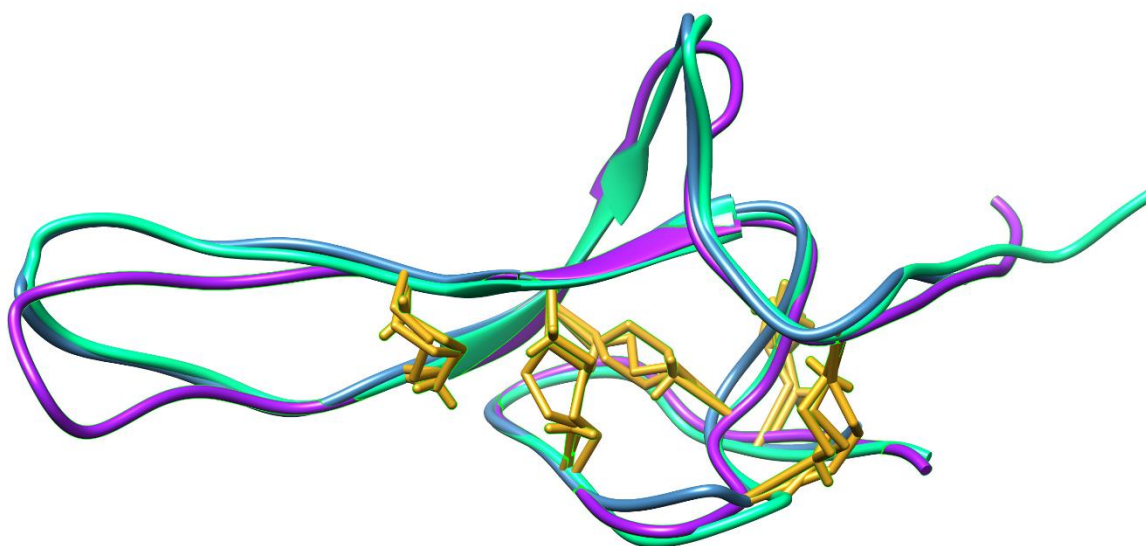


FIG 7.8.8 Comparación visual de los tres modelos obtenidos para la estructura de OxyTx2. ITasser: morado, RaptorX: turquesa, SwissModel: azul.

La longitud de la proteína molde OtTx1a es similar a la de OxyTx2, por lo que en el caso de estos modelos no existe una región que no tenga correspondencia en la plantilla. Esto provoca que exista una identidad mayor entre los modelos de los tres servidores. Sin embargo, la región del giro entre las hebras beta es la que presenta mayor variación. (Tabla 7.8.22) Esto se puede deber a que al ser una región sin una estructura secundaria definida puede presentar mayor variación y flexibilidad. Además, en esta región no se encuentran cisteínas en la secuencia. Todos los enlaces disulfuro, los cuales estabilizan el motivo estructural, se encuentran en la región de la proteína que está mejor conservada. La falta de un enlace de este tipo que le confiera mayor estabilidad a la estructura también contribuye a la variación en esta región, y podríamos esperar que en disolución esta región esté en constante movimiento. Al igual que en el caso de OxyTx1, se utilizan valores de RMS ya que se realizó una comparación uno a uno entre cada par de estructuras, no una comparación con un promedio.

Comparación de los valores de RMS entre los modelos obtenidos de OxyTx2 divididos entre regiones de la toxina.			
Modelos	RMS molécula completa	RMS hebras	RMS lazo entre hebras
ITasser-SwissModel	2.933	1.643	2.115
ITasser-RaptorX	3.374	1.255	1.854
SwissModel-RaptorX	2.177	1.242	1.636

Tabla 7.8.21

Luego de analizar los modelos obtenidos por los tres servidores para los dos péptidos se concluye que los modelos de SwissModel son en ambos casos los mejores puntos de partida para el refinamiento posterior por dinámica molecular. Esto se debe a que las calificaciones de Clashcore tanto para OxyTx1-SM y OxyTx2-SM son mejores comparadas con las obtenidas por RaptorX. Además, en el caso de OxyTx1, SwissModel predice que todos los residuos cisteína se encuentran formando enlaces disulfuro, que es consistente con los resultados previamente observados de espectrometría de masas. Para simplificar la comparación, se eligió trabajar con los modelos obtenidos por SwissModel en ambos casos.

7.9. Refinamiento de modelos por dinámica molecular.

Para ambas toxinas se eligió trabajar con los modelos obtenidos por SwissModel para el proceso de refinamiento por dinámica molecular. Esto es porque en ambos casos el modelo de

SM obtuvo una evaluación favorable de MolProbity que se espera pueda ser mejorada a partir de los experimentos de DM. Además, en ambos casos SM genera modelos que presentan una conectividad de enlaces disulfuro consistente con la esperada. Esto es importante ya que los experimentos de DM no permiten la ruptura de enlaces y por lo tanto no serían capaces de corregir una conectividad mal establecida.

7.9.1 OxyTx1

Se realizó un análisis por clústeres de los resultados obtenidos de la dinámica molecular para determinar cuáles son las estructuras representativas generadas a partir de ésta. Se generaron grupos de los valores ubicados alrededor de cada uno de los centroides (que corresponden una estructura representativa) y así se generaron grupos con números variantes de estructuras (*frames*, cada uno corresponde a 1 ns de la dinámica)

Análisis de clústeres de la dinámica de OxyTx1SM.						
#Cluster	Frames	Frac	AvgDist	Stdev	Centroid	AvgCDist
0	154	0.308	1.216	0.333	144	2.276
1	145	0.290	1.174	0.271	286	2.427
2	97	0.194	1.172	0.264	365	2.571
3	43	0.086	1.357	0.338	170	2.342
4	22	0.044	1.296	0.339	48	2.526
5	13	0.026	1.287	0.306	247	3.060
6	13	0.026	1.506	0.523	18	2.526
7	10	0.020	1.766	0.450	8	2.892
8	2	0.004	1.398	0.000	3	4.067
9	1	0.002	0.000	0.000	1	3.343

Tabla 7.9.1.

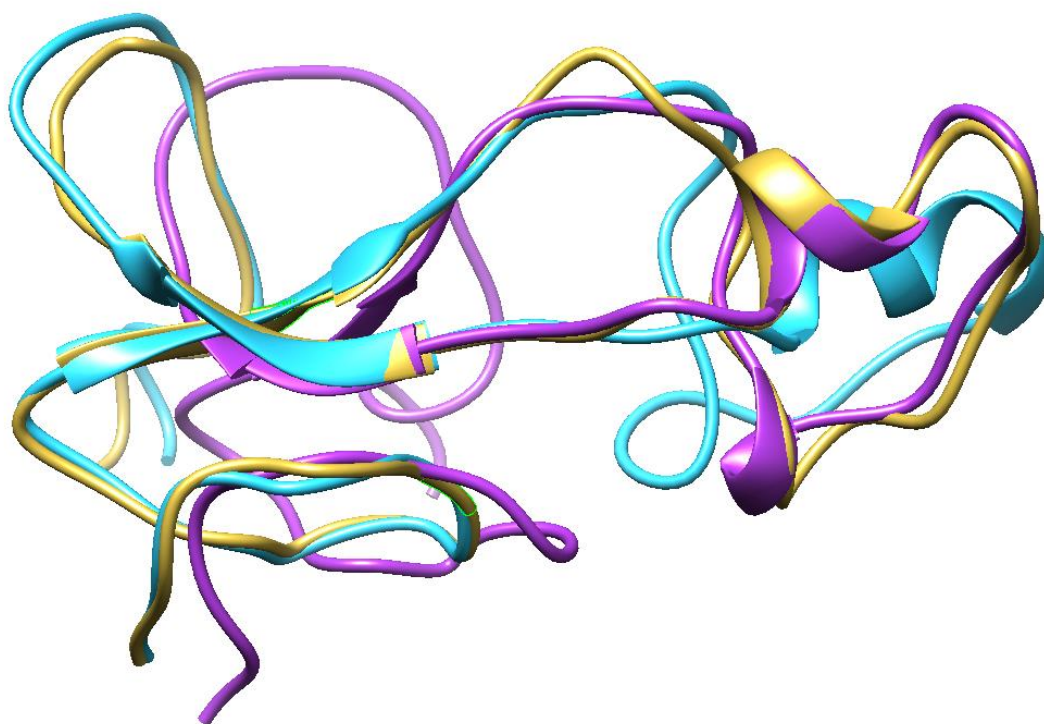


FIG. 7.9.1 Modelo tridimensional de los centroides de los tres clústeres más poblados de la DM de OxyTx1SM. Se puede observar que a lo largo de la simulación se mantuvo la presencia de una zona tipo hélice en la región entre hebras.

En la tabla 7.9.1 se observa que más del 79% de las estructuras generadas a lo largo de los 500 ns de simulación se encuentran en los tres primeros clústeres. Los centroides de estos tres clústeres se representan en la figura 7.9.1. La población de estos grupos representativos indica que la estructura de la toxina se mantuvo relativamente estable; esta observación se puede corroborar analizando los valores de RMSD del modelo a lo largo de la DM.

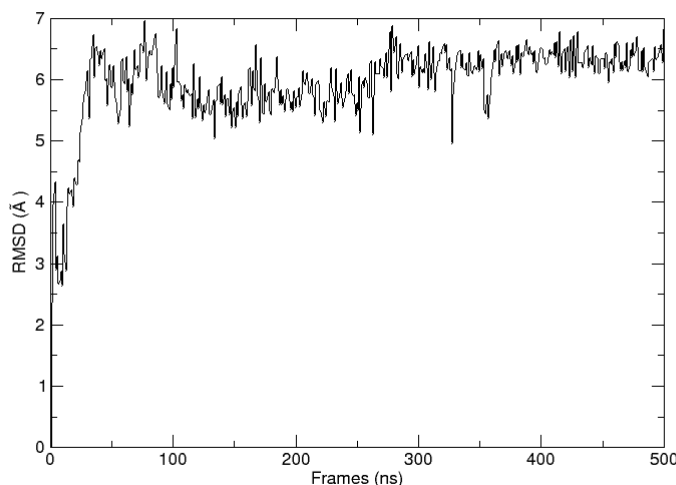


FIG 7.9.2 Valores de RMSD de la toxina a lo largo de la DM de OxyTx1SM.

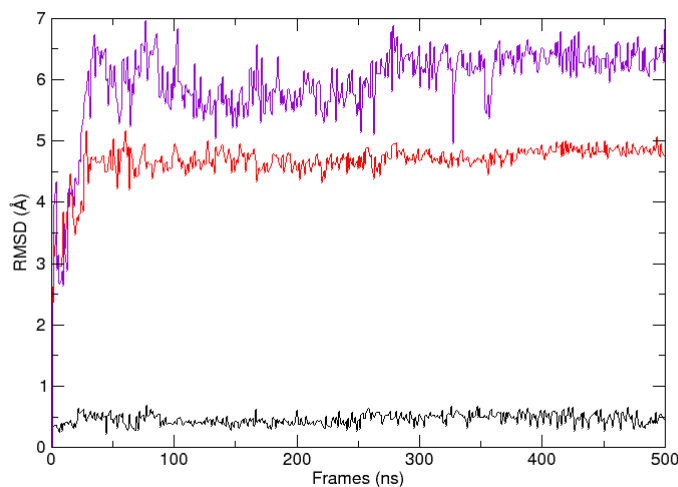


FIG 7.9.3 Comparación de los valores de RMSD por regiones a lo largo de la DM. En morado se muestran los valores correspondientes a la proteína completa, en negro se muestran los valores correspondientes a la región de hebra beta y en rojo el giro entre hebras.

De acuerdo a los valores de RMSD a lo largo de los 500 ns de simulación (Fig 7.9.2), la estructura de la toxina permaneció constante la mayor parte del tiempo. Sin embargo, durante los primeros 100 ns los valores de RMSD aumentaron; al ser el RMSD un valor calculado con respecto a la estructura introducida a la simulación, esto indica que la propuesta inicial de

estructura no es favorecida y la proteína tuvo un proceso de relajación que es significativamente diferente a la inicial. Una vez que la toxina complete este proceso de relajación, los valores de RMSD se mantuvieron relativamente estables; la variación media fue menor de 1 Å.

En la figura 7.9.3 se observan los valores de RMSD de dos diferentes regiones de la toxina así como la toxina completa. El valor de RMSD de la hebra beta permaneció constante con un valor promedio de 0.462 Å. El valor correspondiente a el giro entre hebras se observa alrededor de 4.5 Å y presenta el mismo patrón que el valor correspondiente a la toxina completa. Esto confirma que esta región contribuye significativamente más a la variación en la toxina comparado con la región de hebra que se encuentra estabilizada por la presencia de enlaces disulfuro.

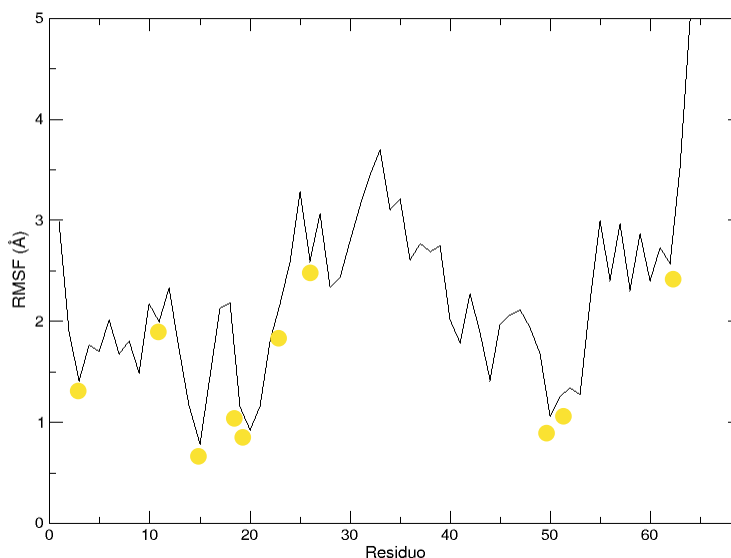


FIG 7.9.4 Valores de RMSF por residuo durante la DM de OxyTx1SM. En amarillo se marca la posición de los residuos cisteína.

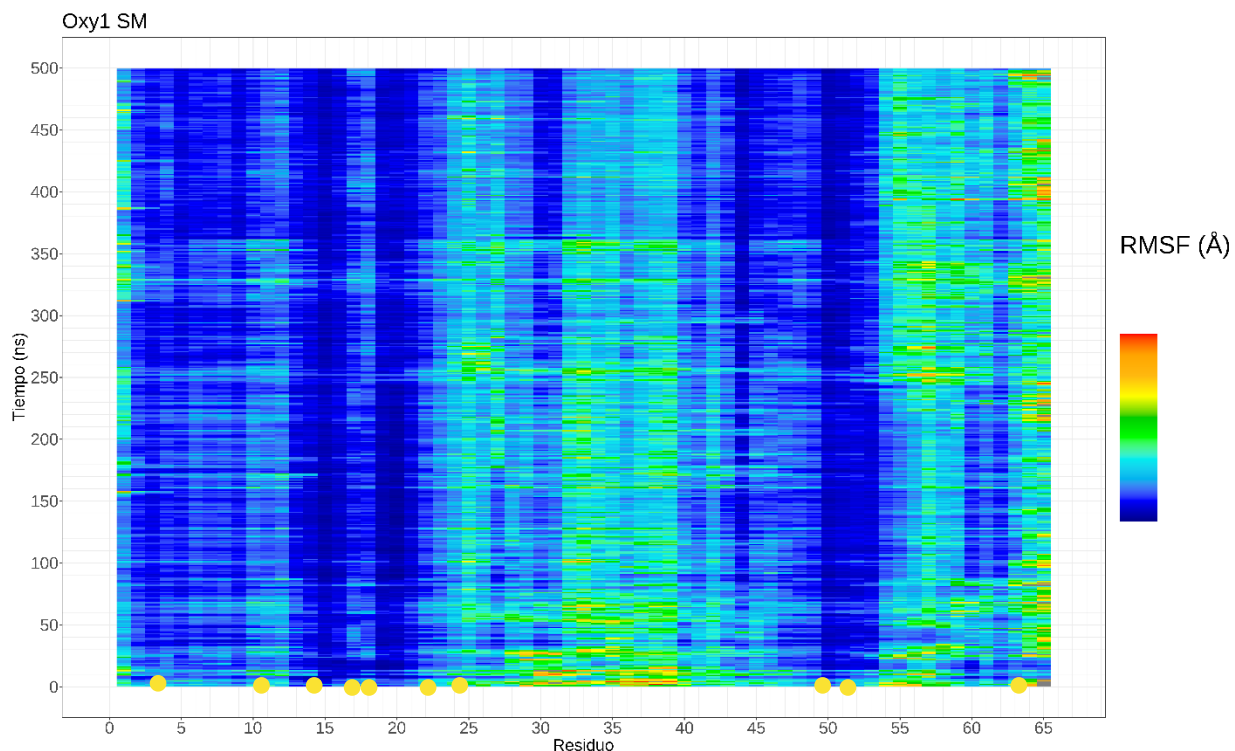


FIG 7.9.5 Fluctuación de los residuos a lo largo de la DM indicada por el valor de RMSF para cada residuo a lo largo de los 500 ns de la trayectoria de OxyTx1SM. En amarillo se marca la posición de los residuos cisteína.

Los valores de RMSF por residuo son indicadores de la fluctuación de cada aminoácido a lo largo del tiempo de simulación. En la figura 7.9.4 se observa que esta fluctuación tiende a ser mayor en la región correspondiente al giro entre hebras y menor en la región correspondiente a la hebra beta. Además, se observa que los residuos cisteína tienden a corresponder a valores mínimos de RMSF. Esto es consistente con una estabilización de la estructura provocada por la presencia de enlaces disulfuro, y puede explicar los valores bajos de RMSD al estudiar la región de hebra beta. En la figura 7.9.5 se observa que la fluctuación de la molécula es considerablemente mayor en la región entre hebras, en la cual no hay residuos cisteína para estabilizar la estructura.

Comparación de las evaluaciones de MolProbity de OxyTx1SM antes y después del proceso de refinamiento por DM.					
		OxyTx1SM sin refinar		OxyTx1SM refinado	
Contactos	Clashcore (todos los átomos)	12.96		1.06	
Geometría de la proteína	Rotámeros pobres	7	11.29%	4	8.33%
	Rotámeros favorecidos	49	79.03%	51	85.42%
	Valores extremos Ramachandran	3	4.76%	0	0.00%
	Valores favorecidos Ramachandran	48	76.19%	31	91.18%
	C β desviaciones >0.25Å	2	3.12%	4	6.25%
	Enlaces erróneos:	0 / 533	0.00%	0 / 533	0.00%
	Ángulos erróneos:	15 / 720	2.08%	4 / 720	0.55%

Tabla. 7.9.2

Se realizó la evaluación por MolProbity de las tres estructuras representativas de los clústeres 0, 1 y 2. En la tabla 7.9.2 se observa esta evaluación en comparación con la evaluación previamente realizada del modelo generado por SwissModel. El parámetro clashcore mejoró significativamente, indicando que existen menos traslapes estéricos en la estructura obtenida a partir de la DM. La geometría de la molécula mejoró también, al generar un menor número de rotámeros pobres, así como un menor número de enlaces y ángulos erróneos.

7.9.2 OxyTx2

Análisis de clústeres de la dinámica de OxyTx2SM.						
#Cluster	Frames	Frac	AvgDist	Stdev	Centroid	AvgCDist
0	127	0.254	1.015	0.236	457	1.438
1	71	0.142	0.948	0.224	24	1.533
2	60	0.120	1.198	0.246	263	1.530
3	58	0.116	1.162	0.234	467	1.553
4	40	0.080	1.129	0.239	388	1.805
5	38	0.076	0.932	0.221	197	1.627
6	38	0.076	1.217	0.285	51	1.678
7	35	0.070	1.145	0.252	303	1.908
8	21	0.042	0.935	0.195	64	2.326
9	12	0.024	1.141	0.225	12	1.851

Tabla 7.9.3.

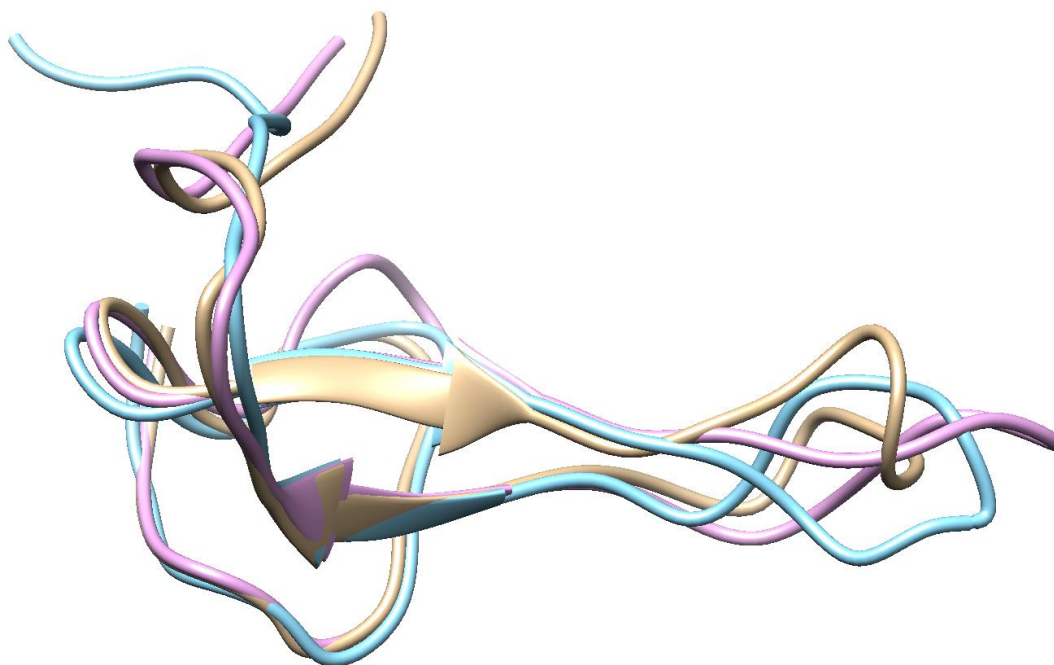


FIG. 7.9.6 Modelo tridimensional de los centroides de los tres clústeres más poblados de la DM de OxyTx2SM.

A diferencia de lo observado en los clústeres en OxyTx1SM, los clústeres generados a partir de la DM de OxyTx2SM están distribuidos de manera más equitativa (tabla 7.9.3); sin embargo, la mayor parte de las estructuras se mantienen dentro de los tres clústeres más poblados (51.6%), de los cuales su estructura representativa se ilustra en la figura 7.9.6. Además, cuatro de los grupos menos poblados contienen la misma cantidad de estructuras (o números muy cercanos), y la diferencia entre estos y los grupos menos poblados es menor que en el caso de OxyTx1SM. Esta distribución indica que probablemente existe mayor variación entre las estructuras generadas a partir de la DM de OxyTx2SM en comparación con la de OxyTx1SM. Los valores de RMSD a lo largo de la trayectoria ayudan a corroborar esta observación.

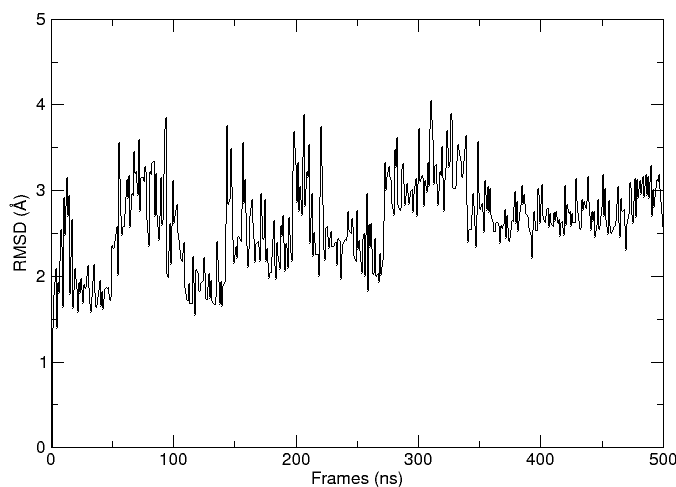


FIG 7.9.7 Valores de RMSD de la toxina a lo largo de la DM de OxyTx2SM.

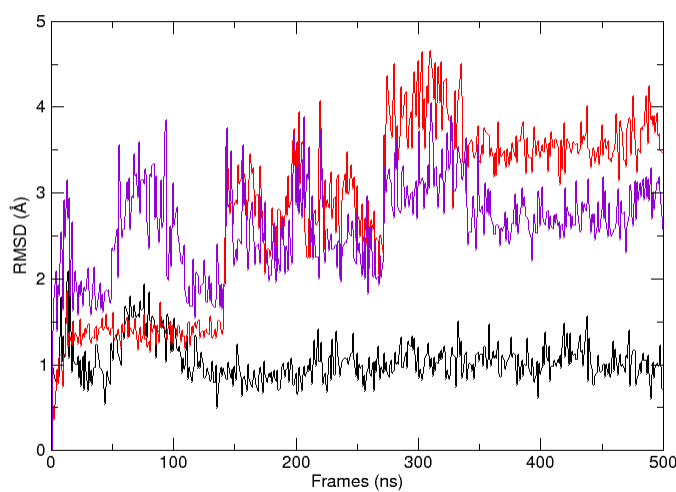


FIG 7.9.8 Comparación de los valores de RMSD por regiones a lo largo de la DM de OxyTx2SM. En morado se muestran los valores correspondientes a la proteína completa, en negro se muestran los valores correspondientes a la región de hebra beta y en rojo el giro entre hebras.

Según los valores de RMSD a lo largo de la simulación que se observan en la figura 7.9.7, la variación máxima con respecto al promedio fue de 1.84 Å. A diferencia de lo observado con OxyTx1, no se observa un cambio drástico que se pueda atribuir a una relajación de la estructura, aunque sí se puede ver que a partir del ns 350 los valores de RMSD variaron menos

con respecto a los primeros 350 ns. El valor promedio de RMSD se mantuvo constante alrededor de 2.8 Å. A partir de esta información se puede concluir que la propuesta inicial de estructura arrojada por el servidor SwissModel es favorecida en un ambiente de disolvente explícito, en especial al compararla con los resultados obtenidos de OxyTx1.

Al realizar la comparación de valores de RMSD en la figura 7.9.8, se observa que para el caso de la región de giro entre hebras sí hay un salto en los valores de RMSD alrededor de 150 ns. A partir de este punto de la simulación la variación entre valores aumenta considerablemente, y en la gráfica se observa una tendencia de similar a la generada por la proteína completa. En el caso de la hebra beta, la variación entre los valores de RMSD es menor, pero a diferencia de la región de giro, la mayor variabilidad se observa antes de los 150 ns. Al comparar estas dos gráficas con la que corresponde a la de la proteína completa se observa que cada una de las regiones proporciona la contribución mayoritaria para los valores de RMSD en la proteína en diferentes puntos de la dinámica (antes de 150 ns la contribución mayor es de la región de hebra y después de 150 ns la contribución mayor es de la región de giro). La región de hebra se encuentra más estabilizada (lo que se espera por la presencia de enlaces disulfuro) con respecto a la región de giro, sin embargo, en comparación con OxyTx1 la variación de la región de giro es menor y esto es consistente con lo que se ha observado en todos los experimentos *in silico* hasta ahora, ya que al ser una región más corta hay menos variabilidad en su estructura.

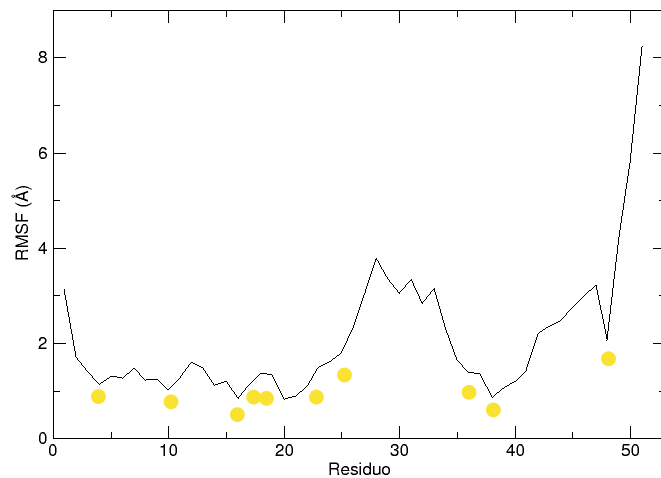


FIG. 7.9.9 Valores de RMSF por residuo durante la DM de OxyTx2SM. En amarillo se marca la posición de los residuos cisteína

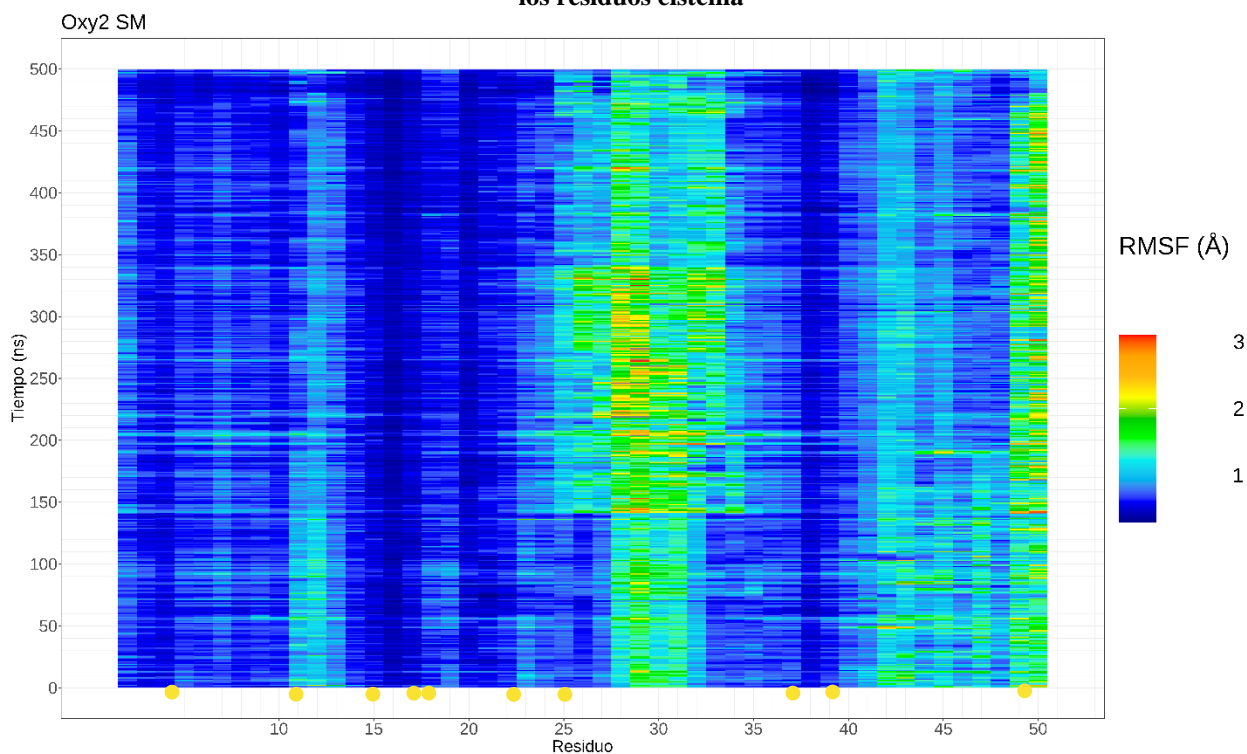


FIG. 7.9.10 Fluctuación de los residuos a lo largo de la DM indicada por el valor de RMSF para cada residuo a lo largo de los 500 ns de la trayectoria de OxyTx2SM. En amarillo se marca la posición de los residuos cisteína.

La fluctuación de los residuos de la toxina (Fig 7.9.9 y 7.9.10) sigue un patrón similar al que se observó en el caso de OxyTx1SM. Los valores de RMSF son menores en los residuos que

corresponden a la región de hebra y se presentan valores máximos en la región que corresponde al giro entre hebras. Los valores más altos de RMSF corresponden a la región de carboxilo terminal (como en el caso de OxyTx1SM). Esto es consistente con la movilidad del carboxilo terminal y es de esperarse en especial por la ausencia de puentes disulfuro en esta región, lo que evita que se estabilice. Se observa también que en las regiones de la toxina correspondientes a los residuos cisteína se generan puntos mínimos de RMSF lo que indica que la fluctuación de la molécula a lo largo de la simulación es menor con respecto a otras regiones; esto confirma la estabilización de la molécula que se genera a partir de la presencia de enlaces disulfuro.

Comparación de las evaluaciones de MolProbity de OxyTx2SM antes y después del proceso de refinamiento por DM.					
		OxyTx2SM sin refinar		OxyTx2SM refinado	
Contactos	Clashcore (todos los átomos)	6.62		1.43	
Geometría de la proteína	Rotámeros pobres	4	8.33%	0	0.00%
	Rotámeros favorecidos	38	79.17%	34	97.14%
	Valores extremos Ramachandran	4	8.16%	1	5.00%
	Valores favorecidos Ramachandran	34	69.39%	15	75.00%
	C β desviaciones >0.25Å	5	10.20%	3	6.12%
	Enlaces erróneos:	0 / 403	0.00%	0 / 533	0.00%
	Ángulos erróneos:	11 / 545	2.02%	0 / 720	0.00%

Tabla. 7.9.2

Después de realizar la evaluación de las tres estructuras representativas de los clústeres más poblados y comparar con la evaluación obtenida por el modelo OxyTx2SM antes de la DM, se observa que la calidad del modelo mejoró considerablemente gracias al proceso de refinamiento. El valor de clashcore y número de rotámeros pobres disminuyó, lo que indica que hay un número menor de traslapes estéricos. Además, el modelo no presenta enlaces ni ángulos erróneos, lo que indica que la geometría de la molécula es más estable y favorecida.

8. Conclusiones

Después de llevar a cabo la purificación de ambas toxinas tanto de la fase soluble como de cuerpos de inclusión, los datos obtenidos de los cromatogramas y los espectros unidimensionales de ^1H para el caso de HisOxyTx1 y HisOxyTx2, se concluye que las toxinas

obtenidas tienen un plegamiento definido. La principal diferencia se encuentra en la cantidad de proteína que se obtuvo del material lisado. El proceso de replegamiento en el caso de HisOxyTx2 permite obtener una mayor cantidad de toxina, pero favorece la cantidad de isoformas generadas y por lo tanto disminuye el rendimiento final de toxina pura con el plegamiento más estable. Para el caso de HisOxyTx1, el proceso de replegamiento disminuyó el rendimiento del proceso de biosíntesis, por lo que no se pudo estudiar su plegamiento por RMN. Esta disminución en el rendimiento puede indicar que el péptido es menos estable incluso en su forma plegada en comparación con HisOxyTx2 y por lo tanto no se puede regresar a su forma plegada una vez llevado a cabo el proceso de desnaturalización necesario para realizar el replegamiento.

La caracterización de ambas toxinas por espectrometría de masas MALDI-TOF y dicroísmo circular indica que los péptidos obtenidos se encuentran en su forma oxidada formando todos los enlaces disulfuro esperados y que el plegamiento que presentan es consistente con el motivo estructural ICK esperado.

Los modelos obtenidos a partir de los experimentos *in silico* de OxyTx1 presentan variaciones significativas con respecto a la plantilla utilizada, así como entre los modelos obtenidos a partir de diferentes servidores y posterior refinamiento por DM. En el caso de OxyTx2 la variación es menor debido a la similitud entre la longitud de la plantilla y la secuencia de interés. En ambos casos se puede confirmar la importancia de la presencia de enlaces disulfuro en la estabilización de las moléculas. La obtención de estos modelos a partir de la estructura tridimensional de OtTx1a respalda la propuesta de una familia de toxinas de araña estabilizadas con motivos ICK estabilizadas por cinco puentes disulfuro. Es necesario obtener modelos experimentales de las Oxitoxinas para confirmar esta propuesta.

9. Perspectivas

Los modelos obtenidos a partir de este trabajo sirven como punto de partida para realizar una comparación con los resultados experimentales que se pueden conseguir más adelante. Habiendo ya refinado el proceso de obtención de los péptidos de interés a partir de sistemas heterólogos, se buscará obtener cantidades suficientes para realizar los experimentos de RMN de

dos dimensiones. Los experimentos a realizar son COSY, TOCSY y NOESY. Una vez obtenidos estos, se deben realizar las asignaciones de señales que permitan generar las restricciones esperadas en la estructura terciaria. Así, partiendo de los modelos in silico ya generados, se podrá realizar nuevamente un proceso de dinámica molecular para determinar la estructura experimental a proponer.

Por otro lado, para refinar los modelos obtenidos se pueden realizar experimentos de predicción de estructura por algoritmos de *machine learning*, tales como AlphaFold, para tener más información sobre el plegamiento de los péptidos y refinar aún más las propuestas de estructuras tridimensionales que aquí se presentan.

10. Referencias

1. Pineda, S. S. *et al.* ArachnoServer 3.0: an online resource for automated discovery, analysis and annotation of spider toxins. *Bioinformatics* **34**, 1074–1076 (2018).
2. King, G. F. Tying pest insects in knots: the deployment of spider-venom-derived knottins as bioinsecticides. *Pest Management Science* **75**, 2437–2445 (2019).
3. Escoubas, P. Molecular diversification in spider venoms: A web of combinatorial peptide libraries. *Molecular Diversity* **10**, 545–554 (2006).
4. Cardoso, F. C. & Lewis, R. J. Structure-function and therapeutic potential of spider venom-derived cysteine knot peptides targeting sodium channels. *Frontiers in Pharmacology* **10**, 366 (2019).
5. Alberts, B. *Molecular Biology of the Cell*. (Garland Science, 2015).
6. Olivera, B. M., Miljanich, G. P., Ramachandran, J. & Adams, M. E. CALCIUM CHANNEL DIVERSITY AND NEUROTRANSMITTER RELEASE: The ω -Conotoxins and ω -Agatoxins. <https://doi.org/10.1146/annurev.bi.63.070194.004135> **63**, 823–867 (2003).
7. Hughes, A. Calcium Channel Blockers. in *Hypertension: A Companion to Braunwald's Heart Disease E-Book* (eds. Bakris, G. L. & Sorrentino, M.) (Elsevier, 2018).
8. Lipscombe, D., Helton, T. D. & Xu, W. L-type calcium channels: The low down. *Journal of Neurophysiology* **92**, 2633–2641 (2004).
9. Perez-Reyes, E. Molecular physiology of low-voltage-activated T-type calcium channels. *Physiological Reviews* **83**, 117–161 (2003).
10. Catterall, W. A. Voltage-Gated Calcium Channels. *Cold Spring Harbor Perspectives in Biology* **3**, a003947 (2011).
11. Wolf, M., Eberhart, A., Glossmann, H., Striessnig, J. & Grigorieff, N. Visualization of the Domain Structure of an L-type Ca²⁺ Channel Using Electron Cryo-microscopy. *Journal of Molecular Biology* **332**, 171–182 (2003).
12. Hall, L. M. *et al.* Calcium Channel as a New Potential Target for Insecticides. 162–172 (1995) doi:10.1021/BK-1995-0591.CH010.

13. Amorim, F. G., Cordeiro, F. A., Pinheiro-Júnior, E. L., Boldrini-França, J. & Arantes, E. C. Microbial production of toxins from the scorpion venom: properties and applications. *Applied Microbiology and Biotechnology* 2018 102:15 **102**, 6319–6331 (2018).
14. Jonasson, P., Liljeqvist, S., Nygren, P.-A. & Ståhl, S. Genetic design for facilitated production and recovery of recombinant proteins in *Escherichia coli*. *Biotechnology and Applied Biochemistry* **35**, 91–105 (2002).
15. Rosano, G. L., Morales, E. S. & Ceccarelli, E. A. New tools for recombinant protein production in *Escherichia coli*: A 5-year update. *Protein Science* **28**, 1412–1422 (2019).
16. Arnau, J., Lauritzen, C., Petersen, G. E. & Pedersen, J. Current strategies for the use of affinity tags and tag removal for the purification of recombinant proteins. *Protein Expression and Purification* **48**, 1–13 (2006).
17. Betton, J. M. & Chaffotte, A. [Recombinant protein folding and production]. *Medecine Sciences : M/S* **21**, 613–617 (2005).
18. Uhoraningoga, A., Kinsella, G. K., Henehan, G. T. & Ryan, B. J. The Goldilocks Approach: A Review of Employing Design of Experiments in Prokaryotic Recombinant Protein Production. *Bioengineering* 2018, Vol. 5, Page 89 **5**, 89 (2018).
19. Diehl, B. Principles in NMR Spectroscopy. *NMR Spectroscopy in Pharmaceutical Analysis* 1–41 (2008) doi:10.1016/B978-0-444-53173-5.00001-9.
20. Rule, G. S. & Hitchens, T. K. *Fundamentals of protein NMR spectroscopy*. vol. 5 (Springer Science and Business Media, 2006).
21. Nelson, J. *Nuclear magnetic resonance spectroscopy*. (Upper Saddle River: Prentice Hall, 2003).
22. Keeler, J. *Understanding NMR Spectroscopy*. (Wiley, 2002).
23. Kleckner, I. R. & Foster, M. P. An introduction to NMR-based approaches for measuring protein dynamics. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1814**, 942–968 (2011).
24. Göbl, C., Madl, T., Simon, B. & Sattler, M. NMR approaches for structural analysis of multidomain proteins and complexes in solution. *Progress in Nuclear Magnetic Resonance Spectroscopy* **80**, 26–63 (2014).
25. Bax, A. Two-dimensional NMR and protein structure. *Annu Rev Biochem* **58**, 223–256 (1989).
26. Rehm, T., Huber, R. & Holak, T. A. Application of NMR in Structural Proteomics: Screening for Proteins Amenable to Structural Analysis. *Structure* **10**, 1613–1618 (2002).
27. Wüthrich, K. Protein Structure Determination in Solution by Nuclear Magnetic Resonance Spectroscopy. *Science (1979)* **243**, 45–50 (1989).
28. Ginalski, K. Comparative modeling for protein structure prediction. *Current Opinion in Structural Biology* **16**, 172–177 (2006).
29. Kuhlman, B. & Bradley, P. Advances in protein structure prediction and design. *Nature Reviews Molecular Cell Biology* 2019 20:11 **20**, 681–697 (2019).
30. Hall, B. *Homology: The Hierarchical Basis of Comparative Biology*. (Academic Press, 2013).
31. Pearson, W. R. An Introduction to Sequence Similarity (“Homology”) Searching. *Current Protocols in Bioinformatics* **42**, 3.1.1-3.1.8 (2013).
32. Gupta, S. & Varadwaj, P. K. A brief overview on molecular dynamics simulation of biomolecular system: Procedure, algorithms and applications. *International Journal of Pharmaceutical Sciences and Research* **9**, 1333–1350 (2018).

33. Puligheddu, M., Xia, Y., Chan, M. & Galli, G. Computational prediction of lattice thermal conductivity: A comparison of molecular dynamics and Boltzmann transport approaches. *Physical Review Materials* **3**, 085401 (2019).
34. Donadio, D. & Galli, G. Thermal conductivity of isolated and interacting carbon nanotubes: Comparing results from molecular dynamics and the Boltzmann transport equation. *Physical Review Letters* **99**, 255502 (2007).
35. Ruiz-Franco, J., Rovigatti, L. & Zaccarelli, E. On the effect of the thermostat in non-equilibrium molecular dynamics simulations. *The European Physical Journal E* **2018** 41:7 **41**, 1–13 (2018).
36. Anandkrishnan, R., Drozdetski, A., Walker, R. C. & Onufriev, A. v. Speed of Conformational Change: Comparing Explicit and Implicit Solvent Molecular Dynamics Simulations. *Biophysical Journal* **108**, 1153–1164 (2015).
37. Onufriev, A. Implicit Solvent Models in Molecular Dynamics Simulations: A Brief Overview. *Annual Reports in Computational Chemistry* **4**, 125–137 (2008).
38. Geng, H., Chen, F., Ye, J. & Jiang, F. Applications of Molecular Dynamics Simulation in Structure Prediction of Peptides and Proteins. *Computational and Structural Biotechnology Journal* **17**, 1162–1170 (2019).
39. Feig, M. & Mirjalili, V. Protein structure refinement via molecular-dynamics simulations: What works and what does not? *Proteins: Structure, Function, and Bioinformatics* **84**, 282–292 (2016).
40. Nurisso, A., Daina, A. & Walker, R. C. A Practical Introduction to Molecular Dynamics Simulations: Applications to Homology Modeling. *Life Sciences* **857**, 137–173 (2012).
41. Villegas, E. *et al.* Biochemical characterization of cysteine-rich peptides from *Oxyopes* sp. venom that block calcium ion channels. *Toxicon* **52**, 228–236 (2008).
42. Mintz, I. M. *et al.* P-type calcium channels blocked by the spider toxin ω -Aga-IVA. *Nature* **1992** 355:6363 **355**, 827–829 (1992).
43. Estrada, G. *et al.* Heterologous expression of five disulfide-bonded insecticidal spider peptides. *Toxicon* **119**, 152–158 (2016).
44. Nadezhdin, K. D. *et al.* Modular toxin from the lynx spider *Oxyopes takobius*: Structure of spiderine domains in solution and membrane-mimicking environment. *Protein Science* **26**, 611–616 (2017).
45. Yang, J. & Zhang, Y. Protein Structure and Function Prediction Using I-TASSER. *Current Protocols in Bioinformatics* **52**, 5.8.1–5.8.15 (2015).
46. Yang, J. *et al.* The I-TASSER Suite: protein structure and function prediction. *Nature Methods* **2015** 12:1 **12**, 7–8 (2014).
47. Waterhouse, A. *et al.* SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Research* **46**, W296–W303 (2018).
48. Källberg, M. *et al.* Template-based protein structure modeling using the RaptorX web server. *Nature Protocols* **2012** 7:8 **7**, 1511–1522 (2012).
49. Case, D. A., Aktulga, H. M., Belfon, K. & Ben-Shalom, I. Y. Amber 2021. *University of California, San Francisco.* (2021).
50. Tian, C. *et al.* Ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution. *Journal of Chemical Theory and Computation* **16**, 528–552 (2020).

51. Anandkrishnan, R., Aguilar, B. & Onufriev, A. v. H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Research* **40**, W537–W541 (2012).
52. Mark, P. & Nilsson, L. Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K. *Journal of Physical Chemistry A* **105**, 9954–9960 (2001).
53. Roe, D. R. & Cheatham, T. E. PTRAJ and CPPTRAJ: Software for processing and analysis of molecular dynamics trajectory data. *Journal of Chemical Theory and Computation* **9**, 3084–3095 (2013).
54. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *Journal of Molecular Graphics* **14**, 33–38 (1996).
55. Pettersen, E. F. *et al.* UCSF Chimera—A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **25**, 1605–1612 (2004).
56. Miles, A. J., Ramalli, S. G. & Wallace, B. A. DichroWeb, a website for calculating protein secondary structure from circular dichroism spectroscopic data. *Protein Science* **31**, 37–46 (2022).
57. Miles, A. J., Janes, R. W. & Wallace, B. A. Tools and methods for circular dichroism spectroscopy of proteins: a tutorial review. *Chemical Society Reviews* **50**, 8400–8413 (2021).
58. Kelly, S. M., Jess, T. J. & Price, N. C. How to study proteins by circular dichroism. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1751**, 119–139 (2005).
59. Kozlov, S. A. *et al.* Latarcins, Antimicrobial and Cytolytic Peptides from the Venom of the Spider *Lachesana tarabaei* (Zodariidae) That Exemplify Biomolecular Diversity. *Journal of Biological Chemistry* **281**, 20983–20992 (2006).
60. Tanner, J., Deplazes, E. & Mancera, R. The Biological and Biophysical Properties of the Spider Peptide Gomesin. *Molecules* **23**, 1733 (2018).
61. Ferreira, F. R. B. *et al.* Evaluation of antimicrobial, cytotoxic, and hemolytic activities from venom of the spider *Lasiadora* sp. *Toxicon* **122**, 119–126 (2016).
62. Pimentel, C. *et al.* Solution structure of PcFK1, a spider peptide active against *Plasmodium falciparum*. *Protein Science* **15**, 628–634 (2006).
63. Nadezhdin, K. *et al.* NMR structure of OtTx1a - ICK. *Protein Data Bank* (2016).
64. Williams, C. J. *et al.* MolProbity: More and better reference data for improved all-atom structure validation. *Protein Science* **27**, 293–315 (2018).
65. Tina, K. G., Bhadra, R. & Srinivasan, N. PIC: Protein Interactions Calculator. *Nucleic Acids Research* **35**, W473–W476 (2007).