



**UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS**

UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS
INSTITUTO DE INVESTIGACIÓN EN CIENCIAS BÁSICAS Y APLICADAS
CENTRO DE INVESTIGACIONES QUÍMICAS

***“Caracterización espectroscópica de la interacción proteína-ligando:
AtLEA4-5 - iones metálicos, LAOBP - aminoácidos y CreCBD - ácidos
nucleicos”***

TESIS

QUE PARA OBTENER EL GRADO DE

DOCTORA EN CIENCIAS

PRESENTA

LEIDYS FRENCH PACHECO

DIRECTOR DE TESIS

DR. CARLOS DANIEL AMERO TELLO

Co-DIRECTOR DE TESIS

DRA. LINA ANDREA RIVILLAS ACEVEDO



**UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS**

UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS
INSTITUTO DE INVESTIGACIÓN EN CIENCIAS BÁSICAS Y APLICADAS
CENTRO DE INVESTIGACIONES QUÍMICAS

***“Caracterización espectroscópica de la interacción proteína-ligando:
AtLEA4-5 - iones metálicos, LAOBP- aminoácidos y CreCBD - ácidos
nucleicos”***

TESIS

QUE PARA OBTENER EL GRADO DE

DOCTORA EN CIENCIAS

PRESENTA

LEIDYS FRENCH PACHECO

Comité Tutorial

Dra. Carmen Nina Pastor Colón
Dra. Rosario Adelaida Muñoz Clares
Dr. Carlos Daniel Amero Tello
Dra. Lina Andrea Rivillas Acevedo



**UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS**

UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MORELOS
INSTITUTO DE INVESTIGACIÓN EN CIENCIAS BÁSICAS Y APLICADAS
CENTRO DE INVESTIGACIONES QUÍMICAS

***“Caracterización espectroscópica de la interacción proteína-ligando:
AtLEA4-5 - iones metálicos, LAOBP - aminoácidos y CreCBD - ácidos
nucleicos”***

TESIS

QUE PARA OBTENER EL GRADO DE

DOCTORA EN CIENCIAS

PRESENTA

LEIDYS FRENCH PACHECO

Jurado Revisor de Tesis

Dr. Enrique Rudiño Piñera
Dra. Carmen Nina Pastor Colón
Dra. Rosario Adelaida Muñoz Clares
Dr. Daniel Alejandro Fernández Velasco
Dra. Claudia Díaz Camino
Dr. Carlos Daniel Amero Tello
Dra. Lina Andrea Rivillas Acevedo

A mi familia

Listado de publicaciones

- ◆ Maya-Martinez R, **French-Pacheco L**, Valdés-García G, Pastor N, Amero C. Different dynamics in 6aJL2 proteins associated with AL Amyloidosis, a conformational disease. Int J Mol Sci. 2019. 21;20(17): E4078.
- ◆ Tejera B, López RE, Hidalgo P, Cárdenas R, Ballesteros G, Rivillas L, **French L**, Amero C, Pastor N, Santiago Á, Groitl P, Dobner T, Gonzalez RA. The human adenovirus type 5 E1B 55kDa protein interacts with RNA promoting timely DNA replication and viral late mRNA metabolism. PLoS One. 2019. 3;14(4):e0214882.
- ◆ **French-Pacheco L**, Cuevas-Velazquez CL, Rivillas-Acevedo L, Covarrubias AA, Amero C. Metal-binding polymorphism in late embryogenesis abundant protein AtLEA4-5, an intrinsically disordered protein. PeerJ. 2018. 7;6:e4930.
- ◆ Sánchez-Carbente MDR, Batista-García RA, Sánchez-Reyes A, Escudero-García A, Morales-Herrera C, Cuervo-Soto LI, **French-Pacheco L**, Fernández-Silva A, Amero C, Castillo E, Folch-Mallol JL. The first description of a hormone-sensitive lipase from a basidiomycete: Structural insights and biochemical characterization revealed Bjerkandera adusta BaEstB as a novel esterase. Microbiologyopen. 2017. 6(4).
- ◆ Pelaez-Aguilar AE, Rivillas-Acevedo L, **French-Pacheco L**, Valdes-Garcia G, Maya-Martinez R, Pastor N, Amero C, “Inhibition of light chain 6aJL2-R24G amyloid fiber formation associated with AL amyloidosis”, Biochemistry, 2015 54(32):4978-4986.

Artículos en preparación (proyectos de colaboración)

- ◆ Site-specific interactions with copper promote lambda 6aJL2 amyloid fibril formation. Pelaez-Aguilar Valdés-García G, **French-Pacheco L**, Pastor N, Amero C, Rivillas-Acevedo L (*enviado*).
- ◆ Novel aggregation mechanism of human γ D crystallin protein induced by metal ions based in NMR studies. Fernández-Silva A, **French-Pacheco L**, Rivillas-Acevedo L, Amero C (*enviado*).
- ◆ A small heat shock protein of common bean forms functional high molecular weight oligomers. Lledías F, Fernández-Silva F, Rodríguez-López J, Olivares JE, **French-Pacheco L**, Corkidi G, Amero C, Saab Rincón G, Díaz-Camino C. (*En preparación*).
- ◆ Analysis of structural disorder in the human Adenovirus proteome. Cárdenas-Mena R, **French-Pacheco L**, Gonzalez RA, Amero C (*En preparación*).

Proyectos asociados durante el período de tesis doctoral

- Caracterización espectroscópica del efecto de la temperatura en el desplegamiento de la cadena ligera 6aJL2-R24G (Co-dirección de Tesis de Licenciatura).

- Estudio de la interacción del extremo amino de la recombinasa Cre del fago T4 con DNA mediante RMN (Estancia de investigación, OSU, USA, supervisión Prof. Mark Foster).
- Caracterización de la interacción de LAOBP con arginina mediante RMN (Colaboración con Dr. Alejandro Sosa, Facultad de Medicina, UNAM).

AGRADECIMIENTOS

A mi tutor, el Dr. Carlos Amero por permitirme ser parte de su grupo de investigación, por ser genuino, sincero y siempre defender la ética profesional. Porque en estos 4 años me enseñó no solo de ciencia, sino a ser mejor persona, pues me inculcó ver siempre lo positivo de los que me rodean. Utilizando una de sus frases, de verdad, de verdad, de verdad, muchas gracias.

A mi co-tutora, la Dra. Lina Rivillas por sus enseñanzas en los temas de espectroscopía.

A mi comité tutorial, en particular a las Dras. Nina Pastor y Rosario Muñoz por ser ejemplo para mi del paradigma de una mujer científica. Por su apoyo constante y sus valiosas sugerencias en cada uno de los exámenes tutorales durante estos años.

A mi jurado revisor de la tesis (comité tutorial y los doctores Enrique Rudiño, Claudia Díaz y Alejandro Fernández) por sus aportes y comentarios que colaboraron al avance y mejora del escrito.

Al Dr. Enrique Rudiño por permitirme la accesibilidad a su laboratorio, apoyarnos con el uso del HPLC y la determinación de las masas de los péptidos.

Al Dr. Alejandro Sosa por su apoyo y colaboración en el proyecto de LAOBP.

Al Dr. Mark Foster y su grupo, por permitirme realizar una estancia de investigación en su laboratorio de OSU brindándome una visión distinta de cómo funciona el mundo de la ciencia.

A la Dra. Alejandra Covarrubias por todos sus aportes académicos, desde mi maestría, examen de candidatura y durante mi doctorado.

A mis compañeros del lab por ser parte cada día de mi vida en el laboratorio.

A Conacyt por la beca otorgada (No becario 305284)

A México y sus instituciones (UAEM, CIQ, CIDC, IICBA, LANEM, UNAM, ISSSTE) por permitirme empezar una nueva etapa de mi vida y cumplir con el sueño de ser doctora en ciencias.

A todos mis amigos, en particular a Arline, Gilberto, Viri, Eya, Reinier G, Ruth y mis amigas de la preparatoria con los que he compartido lindos y difíciles momentos.

A Carlos René, que siempre ha sido incondicional a mi, por ser el hombro donde me apoyo constantemente, por hacerme reír y quererme como soy.

A mi familia, en especial a mi madre y a mi hermana, pues todo lo que soy y he logrado en mi vida ha sido gracias a sus enseñanzas y cuidados para conmigo.

Gracias a todos los que han contribuido con la realización y culminación de la tesis.

PREFACIO

La tesis doctoral que se presenta incluye la descripción de los resultados de los estudios de caracterización estructural realizados a tres sistemas proteicos AtLEA4-5, LAOBP y CreCBD. Se discute la importancia del reconocimiento molecular en la interacción proteína – ligando y la utilización de distintas técnicas espectroscópicas y biofísicas para su estudio (dicroísmo circular, resonancia paramagnética electrónica, resonancia magnética nuclear, dispersión dinámica de luz y calorimetría de titulación isotérmica).

La tesis está dividida en 4 capítulos. En cada capítulo se analizan los siguientes aspectos:

En el Capítulo 1 se describe el concepto y algunas de las metodologías de estudio del reconocimiento molecular.

En el Capítulo 2 se describen estudios realizados a la proteína AtLEA4-5 de *Arabidopsis thaliana*, una proteína intrínsecamente desordenada que interacciona con iones metálicos.

En el Capítulo 3 se describen estudios realizados a la proteína LAOBP de *Salmonella*, una proteína periplasmática que une y transporta aminoácidos.

En el Capítulo 4 se describen estudios realizados al dominio amino de la proteína Cre del bacteriofago T4, una recombinasa de ADN.

En particular los capítulos 2, 3 y 4 contienen las secciones: Introducción, Hipótesis y Objetivos, Métodos, Resultados, Discusión, Conclusiones, Recomendaciones y Bibliografía.

RESUMEN

El reconocimiento molecular (RM) en las interacciones proteínas-ligando, proteína-proteína o proteína-DNA gobierna la afinidad y la especificidad para la formación de estos complejos. El RM es el proceso mediante el cual macromoléculas biológicas interactúan con pequeñas moléculas a través de interacciones no covalentes. Está involucrado en procesos esenciales para la vida como la replicación, el metabolismo o la señalización. Los estudios de biología estructural no sólo tienen como objetivos ampliar los conocimientos documentados hasta el momento, sino también tienen una amplia aplicación, que va desde la terapéutica, mejora del diagnóstico molecular hasta la ingeniería de biosensores, entre otros. Comprender el RM en cuanto a las fuerzas físicas que subyacen a las interacciones proteína-ligando tienen una gran relevancia científica y práctica.

Esta tesis se basa en el uso de distintas técnicas espectroscópicas para caracterizar tres sistemas proteicos cuyos ligandos difieren en sus propiedades fisicoquímicas, AtLEA4-5 – iones metálicos; LAOBP – aminoácidos; CreCBD – ácidos nucleicos. Se emplearon las técnicas espectroscópicas de absorción electrónica en la región UV-vis, dicroísmo circular (DC) y resonancia magnética nuclear (RMN), así como la calorimetría de titulación isotérmica (ITC). En esta tesis se discuten los resultados obtenidos con cada metodología empleada con la finalidad de determinar aspectos relevantes como los residuos involucrados en la interacción o la termodinámica de la unión, así como las estrategias a seguir para completar la caracterización biofísica de la interacción proteína-ligando, todo lo cual pudiera contribuir al entendimiento del funcionamiento de estas proteínas.

Índice

Capítulo I.....	1
I. Reconocimiento molecular entre proteínas y ligandos.....	1
I.1 Introducción.....	1
I.1.1 Mecanismos fisicoquímicos de la interacción proteína - ligando.....	1
I.1.1.1 Cinética de la interacción proteína – ligado.....	1
I.1.1.2 Conceptos básicos de las relaciones termodinámicas.....	2
I.1.1.3 Compensación termodinámica: entropía vs entalpía.....	5
I.2 Paisaje conformacional de las proteínas.....	6
I.2.1 Proteínas intrínsecamente desordenadas (IDPs).....	7
I.3 Técnicas para el estudio de la interacción proteínas-ligando.....	9
I.3.1 Absorción electrónica en la región UV/visible.....	9
I.3.2 Dicroísmo circular (DC).....	10
I.3.2.1 Aspectos espectroscópicos generales de la interacción proteína-metal.....	11
I.3.3 Resonancia paramagnética electrónica (EPR).....	13
I.3.4 Resonancia magnética nuclear en solución (RMN).....	15
I.3.5 Calorimetría de titulación isotérmica (ITC).....	18
I.4 Bibliografía.....	21
Capítulo II.....	24
II. Interacción AtLEA4-5 – iones metálicos.....	24
II.1 Proteínas LEA (Late Embryogenesis Abundant), un grupo de IDPs.....	24
II.1.1 Estudios de la interacción de las proteínas LEA con metales.....	26
II.2 Justificación del proyecto AtLEA4-5 – iones metálicos.....	30
II.3 Hipótesis.....	31
II.4 Objetivos.....	31
II.5 Métodos.....	32
II.5.1 Caracterización espectroscópica de la interacción AtLEA4-5 – Cu(II).....	32
II.5.1.1 Absorción UV-Vis, dicroísmo circular y EPR.....	32
II.5.1.2 Determinación de la constante de unión utilizando curvas de EPR.....	32
II.5.2 Caracterización del estado oligomérico de AtLEA4-5 en ausencia y presencia de Cu(II).....	32
II.5.3 <i>Diseño, síntesis y purificación de los péptidos de AtLEA4-5</i>	33
II.5.3.1 Purificación de los péptidos de AtLEA4-5.....	36
II.5.3.2 Espectrometría de masas ESI.....	36
II.5.4 Caracterización de la interacción péptido – iones metálicos mediante absorción UV-Vis y dicroísmo circular.....	36
II.5.5 Determinación de las zonas de interacción de los péptidos mediante RMN.....	36
II.5.6 Determinación de la termodinámica de la interacción de péptidos – iones metálicos mediante ITC.....	37
II.6 Resultados: Proyecto interacción LEA4-5 – iones metálicos.....	38
II.6.1 Caracterización estructural de AtLEA4-5.....	38
II.6.2 Caracterización espectroscópica de la interacción de AtLEA4-5 con Cu(II).....	40

II.6.3 Caracterización del estado oligomérico de AtLEA4-5. Mecanismo de agregación inducida por Cu(II).....	41
II.6.4 Síntesis y purificación de los péptidos de AtLEA4-5.....	46
II.6.5 <i>Determinación de los modos de coordinación del Cu(II) en la proteína y fragmentos peptídicos de AtLEA4-5</i>	46
II.6.5.1 Caracterización del péptido LEA61-69 – iones metálicos.....	46
II.6.5.2 Caracterización de los péptidos LEA134-146 – iones metálicos.....	49
II.7 Conclusiones.....	56
II.8 Recomendaciones.....	57
II.9 Bibliografía.....	58
Capítulo III.....	61
III. Interacción LAOBP – aminoácidos.....	61
III.1 Proteína Lysine-Arginine-Ornithine Binding Protein (LAOBP), una proteína transportadora de aminoácidos.....	61
III.1.1 Características estructurales de PBP.....	61
III.1.2 Estudios de caracterización de LAOBP.....	62
III.2 Justificación del proyecto: <i>interacción LAOBP - aminoácidos</i>	66
III.3 <i>Hipótesis</i>	67
III.4 Objetivos.....	67
III.5 Métodos.....	68
III.5.1 Expresión de la proteína recombinante LAOBP.....	68
III.5.2 Preparación de la muestra y purificación de LAOBP mediante cromatografía de intercambio aniónico.....	68
III.5.3 Asignación de los desplazamientos químicos de los átomos de LAOBP y de su complejo con arginina.....	69
III.5.4 Determinación de las zonas de interacción de LAOBP con ligandos mediante RMN.....	69
III.5.6 Estudio del paisaje conformacional de LAOBP mediante RMN.....	69
III.6 Resultados.....	71
III.6.1 Asignación de los desplazamientos químicos de los átomos de LAOBP.....	71
III.6.2 Determinación de residuos involucrados en la interacción LAOBP y arginina.....	72
III.6.3 Estudio de las conformaciones de LAOBP.....	75
III.7 Conclusiones.....	78
III.8 Recomendaciones.....	79
III.9 Bibliografía.....	80
Capítulo IV.....	83
IV. Interacción CreCBD – ácidos nucleicos.....	83
IV.1 Dominio amino de la proteína Cre, una recombinasa de ADN.....	83
IV.1.1 Recombinasas específicas de sitio: tirosinas recombinasas.....	83
IV.1.2 <i>Recombinasa Cre del bacteriofago P1</i>	84
IV.2 <i>Justificación del proyecto: interacción CreCBD - ácidos nucleicos</i>	87
IV.3 <i>Hipótesis</i>	88
IV.4 <i>Objetivos</i>	88
IV.5 <i>Métodos</i>	89

IV.5.1	Expresión de la proteína recombinante CBD	89
IV.5.1.1	Preparación de la muestra y purificación de CBD mediante cromatografía de intercambio iónico y exclusión molecular	89
IV. 5.2	Ensayos de calorimetría de titulación isotérmica	90
IV.5.3	Estudio de la interacción proteína – ligando mediante RMN en solución.	90
IV.5.3.2	Determinación de las zonas de interacción de CBD y lox A mediante RMN91	
IV.5.4	Determinación de la composición de estructura secundaria por DC.....	91
IV.5.5	Estabilidad térmica de CBD y del complejo CBD-loxP.	91
IV.6	<i>Resultados</i>	92
IV.6.1	Caracterización del estado nativo de CBD.....	92
IV.6.2	Caracterización termodinámica de la interacción CBD – loxP	93
IV.6.3	Asignación de los desplazamientos químicos de los átomos de CBD	94
IV.6.4	Determinación de contenido de estructura secundaria de CBD y del complejo	96
IV.6.5	Ensayos de estabilidad térmica por CD.....	98
IV.6.6	Estudios de interacción de CBD con loxP mediante RMN.....	99
IV.6.7	Cambios en la estructura de loxP asociados a la unión a CBD	100
IV.7	<i>Conclusiones</i>	102
IV.8	<i>Recomendaciones</i>	103
IV.9	<i>Bibliografía</i>	104
IX	Anexos	106
IX.1.	Capítulo II	106
IX.1.1	<i>Determinación de la masa molecular de péptidos sintéticos de AtLEA4-5Análisis espectrométrico</i>	106

Lista de siglas y abreviaturas

IDR/IDP	intrIntrinsically disordered region/protein
DC	Dicroísmo circular
RMN	Resonancia magnética nuclear
EPR	Electron paramagnetic resonance
UV	ultravioleta
ITC	Isothermal titration calorimetry
LMCT	Ligand to metal charge transfer
°C	grados centígrados
mM	milimolar
nM	nanomolar
RPM	revoluciones por minuto
DOSY	Diffusion ordered spectroscopy
TOCSY	Totally correlated spectroscopy
HSQC	Heteronuclear single quantum coherence
NOESY	Nuclear overhauser enhancement spectroscopy
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
LEA	Late embryogenesis abundant
LAOBP	Lysine arginine ornithine binding protein
Cre	Cause recombination
CBD	Core binding domain of Cre recombinase
ADN	ácido desoxirribonucleico

Lista de Figuras

Figura I.1: Representación de una banda de absorción electrónica	10
Figura I.2: Variables importantes en RMN.....	17
Figura II.1: Alineamiento de secuencias de las proteínas AtLEA4.....	25
Figura II.2: Diagrama de flujo de la síntesis de péptido	35
Figura II.3: Evaluación del desorden estructural de AtLEA4-5.	39
Figura II.4: Estimación de Kd usando datos de EPR.....	41
Figura II.5: Análisis de datos de DLS de AtLEA4-5.....	42
FiguraII 6: Oligomerización de AtLEA4-5 inducida por iones Cu(II).....	44
FiguraII 7: Modelo de interacción de AtLEA4-5 con iones metálicos.....	44
FiguraII 8: Caracterización del péptido LEA61-69.	46
Figura II.9: Ensayo de turbidez del péptido LEA61-69	48
FiguraII 10: Caracterización del péptido LEA134-146	48
FiguraII 11: Caracterización del péptido LEA134-146H139A.....	49
FiguraII 12: Caracterización del péptido LEA134-146H141A.....	50
FiguraII 13: Ensayo de turbidez del péptido LEA134-146.....	52
FiguraII 14: Ensayo de turbidez de los péptidos LEA134-146H139A y LEA134-146H141A.....	53
Figura III.1: Estructura terciaria de apo LAOBP.....	62
FiguraIII 2: Asignación de los espectros ^1H - ^{15}N de LAOBP y su complejo con Arg.	70
FiguraIII 3: Diferencia de los desplazamientos químicos del complejo de LAOBP	71
FiguraIII 4: Comparación de espectros ^1H - ^{15}N HSQC de LAOBP a diferentes temperaturas.....	74
Figura IV.1: Estructura primaria y terciaria de Cre.	83
FiguraIV 2 Esquema de loxP	84
FiguraIV 3: Caracterización de CBD por MALDI-TOF y DC.....	91
Figura IV.4: Ensayo de ITC entre CBD y loxP.	93
FiguraIV 5: Superposición de espectros ^1H - ^{15}N HSQC de Cre y CBD.....	94
Figura IV.6: Asignación de ^1H - ^{15}N HSQC de CBD.....	95
Figura IV.7: Predicción de estructura secundaria mediante DISOPRED y CSI.....	96
FiguraIV 8: Espectros de DC de CBD, lox A y del complejo	97
FiguraIV 9: Superposición de espectros ^1H - ^{15}N HSQC de Cre y CBD.....	98
FiguraIV 10: Superposición de ^1H - ^{15}N HSQC de CBD y su complejo con LoxP.....	99
Figura IV.11: Espectro ^1H de Lox A, CBD y de su complejo.....	100

CAPÍTULO I

Introducción al reconocimiento molecular

Capítulo I

I. Reconocimiento molecular entre proteínas y ligandos

I.1 Introducción

El reconocimiento molecular (RM) se refiere al proceso en el cual las macromoléculas biológicas interactúan entre sí o con moléculas pequeñas a través de interacciones no covalentes para formar un complejo específico (Persch et al., 2015)□. Este proceso tiene dos características importantes: (i) especificidad, que permite que las proteínas distingan una molécula diana entre otras, (ii) afinidad, que determina que la formación del complejo tenga lugar aún a una baja concentración de las moléculas interactuantes (Du et al., 2016)□. El RM forma parte de varios procesos esenciales de la vida como la replicación, el metabolismo o el ciclo celular, entre otros.

Una de las clases de macromoléculas que utilizan el RM para llevar a cabo sus funciones son las proteínas. Las proteínas participan en una gran variedad de actividades celulares, incluyendo estructural (citoesqueleto), mecánica (músculo), bioquímica (enzimas), y de señalización celular (hormonas), entre otras (Martin & Clements, 2013)□. Esencialmente, las proteínas realizan sus funciones biológicas a través de su interacción física con otras moléculas, incluidas proteínas, péptidos, ácidos nucleicos, membranas, sustratos, pequeñas moléculas como oxígeno, solventes o iones metálicos. Un ligando ha sido definido como cualquier molécula capaz de unirse a una proteína con una alta especificidad y afinidad (Du et al., 2016)□.

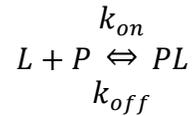
Para una mejor comprensión de las funciones de las proteínas es importante entender los mecanismos responsables de las interacciones proteína - ligando para lo que se requiere la caracterización y cuantificación de las fuerzas que gobiernan e impulsan la formación de los complejos.

I.1.1 Mecanismos fisicoquímicos de la interacción proteína - ligando

Un análisis detallado del reconocimiento molecular entre una proteína y su ligando, requiere de la identificación de los mecanismos fisicoquímicos así como las fuerzas impulsoras relevantes para que la interacción tenga lugar.

I.1.1.1 Cinética de la interacción proteína – ligando

La cinética de unión entre una proteína y su ligando describe la velocidad a la que estas dos especies interactúan (Martin & Clements, 2013)□. En un ejemplo simple, cuando se mezclan una molécula de proteína P y una molécula de ligando L con una afinidad determinada, la asociación entre ellas dependiente del tiempo puede formularse como:



donde: PL representa el complejo proteína - ligando, k_{on} y k_{off} son las constantes cinéticas de velocidad de orden 2 y orden 1 para la reacción de unión directa e inversa (o disociación), respectivamente. Las unidades de k_{on} y de k_{off} son $M^{-1}s^{-1}$ y s^{-1} , respectivamente.

En el equilibrio, se puede describir la relación entre la reacción de unión (directa) e inversa como:

$$k_{on} \cdot [P] \cdot [L] = k_{off} \cdot [PL]$$

donde: [P] y [L] representan la concentración en el equilibrio de cada especie molecular.

La constante de unión K_b (con unidades de M^{-1}) se define por:

$$K_b = \frac{k_{on}}{k_{off}} = \frac{[PL]}{[P] \cdot [L]} = \frac{1}{K_d}$$

donde: K_d se conoce como constante de disociación (con unidades M).

Por lo tanto, la combinación de una k_{on} rápida y una k_{off} lenta resultará en una K_b alta y una K_d baja, es decir en una alta afinidad de unión.

I.1.1.2 Conceptos básicos de las relaciones termodinámicas

Un sistema proteína-ligando-solvente es un sistema termodinámico compuesto por el soluto (moléculas de proteína y ligando) y el disolvente (agua líquida e iones del tampón). Las relaciones entre los componentes del sistema y la transferencia de calor entre ellas están asociadas a diversos cambios de energía. Todos estos procesos están dictados por las leyes de la termodinámica (Du et al., 2016; Khrapunov, 2018)□.

Entonces, las fuerzas motrices que dictan la asociación entre las proteínas y los ligandos resultan de la combinación de las distintas interacciones e intercambios de energía entre los componentes del sistema.

Una de las propiedades dinámicas más importantes para la caracterización de las fuerzas motrices es la variación de la energía libre de Gibbs (ΔG). El ΔG determina la capacidad de un sistema termodinámico para realizar trabajo de forma reversible a una temperatura y presión constante. Cuando $\Delta G < 0$, la unión entre la proteína y su ligando estará favorecida. La magnitud del ΔG negativo se relaciona con la estabilidad del complejo o lo que es igual a la afinidad de unión. La energía libre es una función de estado, por lo tanto, el ΔG considera los estados termodinámicos inicial y final, independientemente de la vía que los conecte (Gilson & Zhou, 2007)□.

La energía libre de unión estándar (ΔG°), se refiere al cambio de energía libre medido en condiciones de 1 atm de presión, una temperatura de 298 K y concentraciones de proteína y ligando de 1 M. El ΔG° está relacionado con la constante de unión K_b por la relación:

$$\Delta G^\circ = -R \cdot T \cdot \ln K_b$$

donde: R es la constante universal de los gases ($1.987 \text{ cal K}^{-1} \text{ mol}^{-1}$) y T es la temperatura (K).

De acuerdo a esta relación cuanto más alta es la constante de unión K_b , más negativo es el ΔG° . Los parámetros cinéticos (k_{on} , k_{off}) y K_b determinan las propiedades termodinámicas del complejo, es decir, la estabilidad y la afinidad del complejo proteína-ligando. La energía libre de unión (ΔG) en cualquier momento durante la asociación, en condiciones no estándares es dada por:

$$\Delta G = \Delta G^\circ + R \cdot T \cdot \ln Q$$

donde: Q es el cociente de reacción, que se define como la relación entre la concentración del complejo y el producto de las concentraciones de proteína y ligando libres en cualquier momento en el tiempo.

El ΔG también se puede analizar en términos de las contribuciones entálpicas y entrópicas al proceso de unión a través de la siguiente relación:

$$\Delta G = \Delta H - T\Delta S$$

donde ΔH y ΔS son cambios en la entalpía y entropía del sistema al unirse al ligando, respectivamente, y T es la temperatura.

Durante un proceso de unión el ΔH , o entalpía de enlace, refleja el cambio de energía calórica del sistema cuando el ligando se une a la proteína. El ΔH puede ser negativo (exotérmico) o positivo (endotérmico). Un ΔH negativo indica la formación de interacciones no covalentes energéticamente favorables entre los átomos de la interfaz de la interacción (contactos de van der Waals, enlaces de hidrógeno, enlaces iónicos o cualquier otra interacción polar). Si el ΔH es positivo sucede lo contrario (Li et al., 2014)□.□

La entropía es una medida de la distribución uniforme de la energía térmica en un sistema termodinámico y está asociada al desorden o aleatoriedad de los átomos y moléculas dentro de este (Bronowska, 2011)□. El cambio en la entropía (ΔS) es una propiedad del sistema, su signo positivo o negativo indica el aumento general o la disminución en los grados de libertad del sistema, respectivamente. El cambio neto en el calor global del sistema depende de la contribución favorable o desfavorable de algunos procesos como: formación y ruptura de muchas interacciones individuales entre la proteína y el solvente, entre el ligando y el solvente o la reorganización del solvente cerca de las superficies del complejo (Liu et al. 2012).

El ΔS total asociado a la unión se puede dividir en varios términos entrópicos:

$$\Delta S = \Delta S_{solv} + \Delta S_{conf}^{ligando} + \Delta S_{conf}^{proteína} + \Delta S_{r/t} + \Delta S_{otros}$$

donde ΔS_{solv} representa el cambio de la entropía del solvente, ΔS_{conf} representa el cambio de entropía conformacional del ligando y de la proteína respectivamente, $\Delta S_{r/t}$ representa el cambio de entropía de traslación y rotación de la proteína y el ligando y ΔS_{otros} representa el cambio de entropía asociado a una fuente no conocida.

El ΔS_{solv} está asociado principalmente a la liberación de moléculas de solvente de la superficie de interacción o del ligando lo que resulta en una contribución favorable a la entropía de unión, la cual alcanza valores positivos. También se ha asociado a cambios en el área superficial disponible tras la formación del complejo. El ΔS_{conf} refleja los cambios

en la libertad conformacional de la proteína y del ligando tras la unión, lo que puede contribuir favorable o desfavorablemente a la entropía de unión porque el grado de libertad del complejo puede ser mayor o menor que los de la proteína y el ligando libres. El $\Delta S_{r/t}$ representa la pérdida de grados de libertad traslacional y rotacional de la proteína y el ligando durante la formación del complejo, lo que contribuye desfavorablemente a la entropía del sistema de unión (Bronowska 2011; Liu et al. 2012, Caro et al., 2017)□.

Los procesos de unión estarán entrópicamente favorecidos si las penalizaciones entrópicas ($\Delta S_{r/t}$ negativa) son superadas con una ganancia entrópica (S_{solv} positivo).

I.1.1.3 Compensación termodinámica: entropía vs entalpía

El cambio en la entalpía y entropía, es un factor determinante de la unión pues determina el signo y la magnitud de la energía libre de unión. Las contribuciones de ambos, ΔH y ΔS , al ΔG están estrechamente relacionadas. Por ejemplo, un proceso de unión resultante de múltiples interacciones no covalentes favorables entre el ligando y la proteína (ΔH negativo) suele ir acompañado de un cambio de entropía negativo debido a la restricción de la movilidad de las moléculas que interactúan. Del mismo modo, una alta ganancia entrópica asociada al proceso de unión tiene como contraparte una penalización entálpica (ΔH positivo) debido al cambio de energía producto de las interacciones no favorables. Este fenómeno, en el que el cambio en magnitud del ΔG es causado por cambios complementarios entre entalpía y entropía, se denomina compensación termodinámica o compensación entalpía - entropía (Khrapunov, 2018)□.

El comportamiento de la compensación entalpía - entropía depende de varios factores estructurales y termodinámicos como son: propiedades del disolvente (efecto hidrofóbico, solvatación - desolvatación y estructura local del agua), la flexibilidad del sitio de unión al ligando y sus alrededores, la estructura molecular del ligando y los cambios en las fuerzas intermoleculares durante el proceso de unión.

El fenómeno de compensación termodinámica también es resultado de otros factores que deben tenerse en cuenta al analizar la contribución de este fenómeno al proceso de unión. Algunos de los descritos son: interpretación incorrecta de los datos, obtención de datos en un intervalo de temperatura relativamente estrecho, resultados producto de errores

aleatorios experimentales y sistemáticos o sesgados por algún elemento (Khrapunov, 2018)□.

En el estudio del reconocimiento molecular, con alguna finalidad aplicada, durante la interacción de la proteína con el ligando se busca lograr una disminución de la energía libre minimizando la penalización entrópica o entálpica y evitar así los efectos desfavorables de la compensación entalpía - entropía en el nivel termodinámico.

1.2 Paisaje conformacional de las proteínas

Las propiedades termodinámicas entalpía y entropía también contribuyen al plegamiento y estructura terciaria de las proteínas. La estructura tridimensional depende en gran medida de la secuencia de la proteína, y por otra parte del balance de las energías de interacción proteína-proteína, proteína-agua y agua-agua. Las proteínas requieren para la estabilización de su estructura terciaria que se establezcan interacciones no covalentes y en ocasiones de la formación de puentes disulfuros (Liu et al., 2012)□. La información requerida para que la proteína pueda adoptar su estructura tridimensional está contenida en su secuencia de aminoácidos, siendo la configuración final el estado de más baja energía libre de Gibbs del sistema (Bicout & Szabo, 2008)□.

La teoría o modelo del “embudo” o “cono entrópico” describe cómo las proteínas se pliegan moviéndose dentro de un embudo, cuyo diámetro representa la entropía de la cadena polipeptídica y la profundidad determina la parte energética del sistema. Hacia el fondo del embudo hay menos entropía y estados accesibles, por lo que las conformaciones adquiridas serán energéticamente más estables. En el final del cono se encuentra la forma plegada, donde la energía interna será óptima y el valor entrópico mínimo.

Las proteínas pueden presentar conformaciones transitorias estables las cuales podrían participar en su interacción con otras moléculas, en su oligomerización o constituir estados iniciales para la agregación.

Mediante estudios de plegamiento de proteínas se ha establecido la existencia de intermediarios de plegamiento parcialmente estructurados o estructuras estables conocidas como glóbulos fundidos y pre-glóbulos fundidos (Uversky & Dunker, 2010). Los glóbulos

fundidos conservan algunos elementos de estructura secundaria nativa y la posición tridimensional de éstos, sin embargo difieren de una estructura globular en cuanto a que presentan una mayor movilidad de sus asas, extremos amino y carboxilo y una menor compactación de sus cadenas laterales. Los pre-glóbulos fundidos son mucho más compactos que las estructuras en forma desplegada, pero son menos compactos y tienen menos elementos de estructura secundaria que los glóbulos fundidos.

De acuerdo a la teoría del cono entrópico las proteínas desplegadas se localizan en el borde superior del cono, con un valor entrópico alto. El paisaje energético de las proteínas desplegadas puede representarse como un cono casi plano o una bandera ondulante. En esta región del embudo, la proteína puede existir en múltiples estados energéticamente desordenados, siendo las barreras energéticas entre ellos pequeñas. Aquellas proteínas cuyo estado nativo es el desplegado, se conocen como proteínas intrínsecamente desplegadas (IDP).

I.2.1 Proteínas intrínsecamente desordenadas (IDPs)

Las proteínas o regiones de éstas que carecen de una estructura tridimensional definida en su estado nativo se han agrupado bajo el término de proteínas o regiones intrínsecamente desplegadas, desordenadas o sin estructura (IDPs/IDRs, de sus siglas en inglés) (Wright & Dyson, 1999)□. El término “intrínsecamente desordenado” se refiere a que la proteína existe como un ensamble estructural, en el que las posiciones de los átomos y los ángulos de la cadena principal en las gráficas de Ramachandran varían significativamente en el tiempo. Este término agrupa estructuras que van desde estructuras extendidas o “random coil” hasta dominios colapsados en forma de glóbulos parcialmente plegados, o pre-glóbulos fundidos.

Las IDPs/IDRs presentan poca variabilidad en sus secuencias, teniendo una composición sesgada a un número reducido y específico de aminoácidos. Principalmente están enriquecidas en determinados residuos (M, K, R, S, Q, P y E), mientras que a menudo carecen o contienen pocos residuos de cisteínas y de aminoácidos que favorecen la formación de parches hidrofóbicos (W, Y, F, I, V, y L) (Dunker et al., 2008; Wright & Dyson, 1999)□.

A diferencia de las proteínas plegadas, que tienen una conformación predominante, las IDPs no tienen la capacidad de formar suficientes interacciones inter-residuo y existen en el estado nativo como un conjunto de conformaciones (Mészáros et al., 2009)□. Se han propuesto varias ventajas funcionales del “desorden” estructural como son: actuar como cadenas entrópicas (conectores entre dominios), facilitar el reconocimiento molecular (pequeños motivos o dominios desordenados) (Tompa et al., 2005)□, desacoplamiento de la especificidad y la fuerza de unión (Wright & Dyson, 2009)□ y aumento de la velocidad de las interacciones (Dunker et al., 2008)□, entre otras (Pancsa & Tompa, 2012)□.

La funcionalidad inherente al “desorden” estructural se ha asociado al reconocimiento y ensamblaje molecular (Dunker et al., 2008)□.

Aunque las IDPs carezcan de una estructura tridimensional definida, algunas regiones de la secuencia pueden formar elementos de estructura secundaria vinculados al proceso de reconocimiento y se conocen como elementos de reconocimiento molecular. Estos motivos de reconocimiento pueden plegarse en hélice α , hebras β , o formar una estructura irregular en el proceso de unión a una molécula diana (Wright & Dyson, 2009)□. Algunas regiones pueden permanecer desplegadas dentro del complejo, lo que contribuye a minimizar la pérdida entrópica resultante de la unión y al ajuste fino de la afinidad de las interacciones (Tompa & Fuxreiter, 2008)□.

Se han propuesto dos posibles mecanismos extremos para explicar la vinculación entre el plegamiento y la unión en sí: plegado inducido o selección de la conformación (Tompa & Fuxreiter, 2008)□. En el primer mecanismo, la IDP se une a su ligando desde un estado completamente desordenado y posteriormente se pliega en el complejo con la proteína diana. En el mecanismo de selección de la conformación, la proteína diana “selecciona” la conformación del conjunto de conformaciones posibles que adopta la IDP (Wright & Dyson, 2009)□.

Desde el punto de vista estructural las IDPs pueden presentar “desorden” estático o dinámico (Tompa & Fuxreiter, 2008)□. El desorden estático se refiere a que una región puede adoptar varias conformaciones estables. El desorden dinámico se refiere a una región que constantemente está fluctuando en un conjunto de conformaciones. Tras la unión de la

IDP a su molécula blanco, el desorden o “ambigüedad” estructural presente en la interacción IDP-proteína se conoce como “*fuzziness*”, el cual puede ser estático o dinámico (Tompa & Fuxreiter, 2008)□. Si la IDP adquiere un estado plegado pero en distintas conformaciones en el estado unido se considera *fuzziness* estático. Si la IDP no sufre una transición de desplegado a plegado tras la unión, sino que permanece en un conjunto de conformaciones en equilibrio rápido se conoce como *fuzziness* dinámico.

1.3 Técnicas para el estudio de la interacción proteínas-ligando

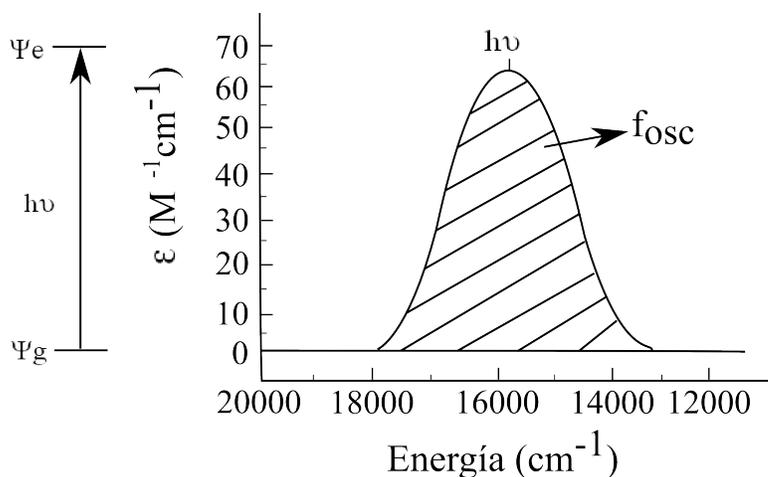
Para la caracterización de la interacción proteína-ligando se han combinado varias técnicas espectroscópicas, que incluyen: absorción electrónica en la región UV-vis y dicroísmo circular (DC) (Boucher et al., 2010)□, resonancia magnética nuclear (RMN) (Ágoston et al., 2011)□ y resonancia paramagnética electrónica (EPR, de sus siglas en inglés) (Bund et al., 2010), así como la calorimetría de titulación isotérmica (Freyer & Lewis, 2008)□.

1.3.1 Absorción electrónica en la región UV/visible

La técnica de espectroscopía de absorción electrónica en la región UV/visible ha sido empleada en la cuantificación de proteínas y ácidos nucleicos (Parson, 2008). La mayoría de las proteínas absorben en la región UV cercano (250-300nm) debido a sus residuos aromáticos y alrededor de los 200 nm por el grupo amino del enlace peptídico. La técnica también es útil en el estudio de interacciones proteína-ligando o en la detección del efecto del entorno, pues pueden ser detectadas variaciones en las bandas y energías de absorción características de una proteína (Parson, 2008)□.

La espectroscopía de absorción electrónica es una técnica que permite estudiar u observar las transiciones electrónicas entre el estado basal (ψ_g) y el excitado (ψ_e) de la molécula en estudio, como por ejemplo el complejo proteína-metal (Dickerson et al., 1979; Quintanar & Rivillas-Acevedo, 2013)□. La probabilidad de transición debido a la interacción de un fotón con un electrón de un complejo metálico se expresa teóricamente a través del operador del momento dipolo (M), el cual depende del dipolo eléctrico, del dipolo magnético y del cuadrupolo eléctrico. Para la espectroscopía de absorción en la región UV-vis el término que predomina para que la transición ocurra es el dipolo eléctrico. La

intensidad de una banda de transición electrónica depende de la fuerza del oscilador (F_{osc}). La F_{osc} asociado a la transición electrónica corresponde a la integral del área bajo la curva



de la banda de absorción (Figura I.1).

Figura I.1: Diagrama esquemático de una transición electrónica que se produce entre el estado fundamental (Ψ_g) y el estado excitado (Ψ_e) de una molécula, y ϵ conduce a una Gaussiana (banda de absorción formada). (Modificado de Quintanar & Rivillas-Acevedo, 2013).

Si la integral es diferente de cero, la transición electrónica está permitida y se observa en el espectro de absorción con alta intensidad. Si la integral es cero la transición electrónica está prohibida. Esto no necesariamente significa que la transición no se observará, pues existen varios mecanismos que pueden incrementar su intensidad (acoplamiento espín-orbita o acoplamiento vibrónico, entre otros) (Parson, 2008; Quintanar & Rivillas-Acevedo, 2013)□. Entonces, una transición electrónica prohibida por dipolo eléctrico pudiera ser observada en el espectro de absorción pero con menor intensidad que las transiciones permitidas.

I.3.2 Dicroísmo circular (DC)

La técnica espectroscópica de dicroísmo circular (DC) ha sido muy utilizada en los estudios de la estructura secundaria de las proteínas y ácidos nucleicos (Parson, 2008).

La región del UV lejana brinda información sobre los elementos de estructura secundaria, hélices α , hojas β y *random coil* (Parson, 2008). Si la proteína contiene principalmente hélices se observará una banda positiva cerca de 195 nm y dos mínimos

característicos en 208 y 222 nm. Las proteínas desplegadas presentan un único mínimo alrededor de los 200 nm y una banda positiva a 222 nm. Por su parte, las hojas β presentan un mínimo característico a 215-218 nm y una señal positiva cerca de los 200 nm (Martin & Schilstra, 2008)□. La región UV cercana brinda información sobre la estructura terciaria y sobre el ambiente de los grupos prostéticos o cofactores (Goyal et al., 2003)□.

La conformación B del ADN es la más frecuentemente encontrada. Su espectro de DC típico muestra una banda positiva a aproximadamente entre los 260–280 nm y una banda envolvente negativa alrededor de los 245 nm. Una horquilla es una variante intramolecular de la forma B que además de la doble hélice, contiene un bucle monocatenario. La posición y amplitudes de las bandas de DC de ADN difieren notablemente con la secuencia debido a que los cromóforos y propiedades conformacionales son distintos (Kypr et al., 2009).

La técnica DC se basa en la medición de la absorción diferencial de la luz circularmente polarizada (CP) a la derecha (ϵ_L) y a la izquierda (ϵ_R), de moléculas que son quirales ($\Delta\epsilon = \epsilon_L - \epsilon_R$). Particularmente, en el estudio de proteínas se mide el momento dipolo magnético del enlace peptídico. Experimentalmente, una señal de DC se caracteriza por la diferencia de absorción de los componentes de la luz ($\Delta A = A_L - A_R$): algunos de los instrumentos de DC cuantifica la diferencia en términos de elipticidad ($\theta = 32.98 \Delta A$) en grados. Entonces $\theta = 32.98 \Delta\epsilon Cl$, donde C es la concentración de la molécula quiral que absorbe, $\Delta\epsilon$ es la diferencia en el coeficiente de extinción molar para la luz CP absorbida a izquierda y la derecha, y l es la longitud del paso óptico (Martin & Schilstra, 2008; Parson, 2008)□.

I.3.2.1 Aspectos espectroscópicos generales de la interacción proteína-metal

Para explicar la base molecular de la coordinación de los metales de transición con sus ligandos se han desarrollado varias teorías de enlace (electrostática, enlace de valencia, campo cristalino y campo ligando), cada una de las cuales fue enriquecida teóricamente por la anterior (Dickerson et al., 1979)□.

La teoría de campo-ligando se basa en que cuando la interacción entre los orbitales moleculares de la proteína y los orbitales d del metal toma lugar, todos los orbitales

adquieren una energía diferente (Quintanar & Rivillas-Acevedo, 2013)□. Esta interacción se conoce como campo-ligando. La magnitud y características de este efecto campo-ligando depende de la geometría del complejo metálico y de la fortaleza de la interacción proteína-metal. Los orbitales del ligando que interactúan con el metal quedan estabilizados por la interacción. Cuando los orbitales *d* del metal no están llenos (tienen menos de 10 electrones) es posible detectar mediante espectroscopía de absorción electrónica las transiciones campo-ligando y las de transferencia de carga que involucran a la proteína y a los orbitales *d* del metal.

Las transiciones campo-ligando se refieren a las transiciones electrónicas que involucran orbitales *d*, y son conocidas como bandas d-d. Las transiciones transferencia de carga conocidas como *ligand to metal charge transfer* (LMCT, por sus siglas en inglés) se refieren a las transiciones electrónicas que se originan en un orbital del ligando y terminan en un orbital *d* del metal.

Ambas transiciones electrónicas (campo-ligando y transferencia de carga) pueden determinarse mediante espectroscopía de absorción electrónica y DC siempre que el metal que se une a la proteína en estudio sea paramagnético, es decir, tenga uno o más electrones desapareados. Estos complejos también pueden estudiarse mediante EPR (Quintanar & Rivillas-Acevedo, 2013)□. La mayoría de los complejos proteína-metal son quirales y resulta muy útil coleccionar los espectros de absorción y de DC.

En los complejos proteína-metal las transiciones d-d están prohibidas por dipolo eléctrico, por lo que tienen muy poca o ninguna intensidad. Por otra parte, las LMCT están permitidas por dipolo eléctrico y muestran una alta intensidad en los espectros de absorción. En un espectro del complejo proteína-metal las transiciones LMTC aparecen a mayor energía, mientras que las transiciones d-d aparecen a energías menores (Quintanar & Rivillas-Acevedo, 2013)□.

En la espectroscopía DC los términos dipolo magnético y dipolo eléctrico son dominantes en el operador, por lo que tanto las transiciones d-d como las LMCT se pueden observar por DC (Parson, 2008)□. Dado que las señales de DC pueden ser positivas o

negativas, las transiciones solapadas en el espectro de absorción pueden distinguirse en el espectro DC.

I.3.3 Resonancia paramagnética electrónica (EPR).

La resonancia paramagnética electrónica (EPR, siglas en inglés) es una técnica espectroscópica que permite el estudio de moléculas con al menos un electrón desapareado (Lancaster, 1967; Quintanar & Rivillas-Acevedo 2013). Mediante un marcaje sitio específico con un compuesto paramagnético se pueden estudiar cambios conformacionales, movilidad, distancias inter e intramoleculares y accesibilidad de proteínas (Drescher, 2012). Además, permite el estudio de complejos proteína-metal, siempre que el metal sea paramagnético, así como la formación de radicales libres.

La interacción del espín electrónico de un electrón aislado con un campo magnético externo causa su desdoblamiento en dos estados de diferente energía, lo que se conoce como efecto Zeeman (Weckhuysen et al. 2004). La energía relativa de cada subnivel de espín depende de la magnitud del campo magnético aplicado a la muestra (B_0) a través de la relación $E = m_s g_e \beta B_0$, donde m_s es el espín electrónico ($\pm 1/2$), g_e es un factor espectroscópico (2.0023192778 para el electrón desapareado) y β es el magnetón de Bohr.

Debido a su carga y a su momento angular (S) el electrón actúa como un dipolo magnético con un momento magnético μ . En el contexto de una molécula, los electrones cambian su momento angular total pues este tendrá la contribución del momento angular intrínseco y el momento angular orbital (L) (Hagen, 2006). La magnitud de la contribución del acoplamiento espín-orbital depende del tamaño del núcleo donde se encuentre el electrón desapareado, entonces el factor g de elementos como los iones metálicos difieren significativamente de g_e . El factor g ($g_e + g_g$) contiene información química de la naturaleza del enlace entre el electrón y la molécula y de la estructura electrónica de la molécula. Resulta evidente la importancia de la determinación del factor g mediante un experimento de EPR para los estudios de caracterización de complejos proteína – metal.

Además del acoplamiento espín-orbital, los espines electrónicos de los electrones desapareados en las moléculas se puede acoplar a otros momentos magnéticos nucleares angulares ocasionando interacciones hiperfinas nucleares o dipolo-dipolo (con otros espines

electrónicos). Las interacciones hiperfinas involucran al propio núcleo de donde proviene el electrón y las superhiperfinas a núcleos distintos a éste.

A partir del espectro se puede determinar la distancia entre dos señales medida de la constante de acoplamiento hiperfino (A). El factor g y la constante A brindan información sobre el ambiente químico alrededor del espín electrónico, y con ello se pueden conocer la orientación y el modo de coordinación del metal unido a la proteína (Quintanar & Rivillas-Acevedo 2013; Weckhuysen et al. 2004).

En un equipo de EPR la muestra paramagnética es colocada en un campo magnético uniforme. En un experimento típico la muestra es irradiada con una frecuencia de microondas (GHz) fija y sometida a un barrido del campo magnético (Gauss) (Jeschke, 2017). Como resultado el estado energético basal desdobra en niveles de energía ($\Delta E = g \beta B_0 = h\nu$). Dado que h y β ($h = 6.626 \times 10^{-34}$ J s; $\beta = 9.274 \times 10^{-28}$ J G⁻¹) son constantes y los valores de B_0 son conocidos, el valor g puede determinarse como:

$$g = \frac{h\nu}{\beta \cdot B_0} = 0.7145 \frac{\nu(\text{MHz})}{B_0(\text{Gauss})}$$

Otro factor importante en los estudios por EPR es la anisotropía, la cual resulta de la orientación de los orbitales de la molécula. Los espectros de metal-proteína, en general, son adquiridos en tubos de cuarzo congelados y en consecuencia los factores g (descritos mediante un tensor g), están distribuidos en un sistema de eje principal a lo largo de los ejes x, y y z (g_x , g_y y g_z) y variarán en ese sentido ($g = g_e + \delta g_z$; $g' = g_e + \delta g_{xy}$) (Hagen, 2006). Los valores de g_x , g_y y g_z brindan información sobre la simetría de la distribución de la densidad del espín electrónico. Se han descritos tres clases de anisotropía: isotrópica ($g_x = g_y = g_z$); axial ($g_x = g_y < g_z$ o $g_x = g_y > g_z$) y rómbica ($g_x \neq g_y \neq g_z$). Si el eje de simetría es único, g tendrá simetría axial (g_z se ubica a lo largo del eje $g_z = g_{||}$, mientras los otros dos g serán perpendiculares ($g_{xy} = g_{\perp}$)). Si la simetría es tetrahédrica o cúbica, g será isotrópico.

Particularmente, en los complejos Cu(II)-proteína, se han encontrado dos tipos de coordinación. En la tipo I, la coordinación es trigonal planar con varios ligandos axiales (Jeschke, 2017). En la tipo II el cobre prefiere una coordinación cuadrado plana (plano ecuatorial), con uno o dos ligandos axiales. El tensor g muestra simetría axial.

Existen distintas fuentes de frecuencia de microondas (L, S, X, Q, W, J) (Jeschke, 2017). En los estudios de metaloproteínas usualmente se usan equipos de EPR equipados con banda X, pues se obtienen mejor resolución de las interacciones hiperfinas. Aun cuando podrían variarse el campo magnético o la frecuencia, la fuente de radiación de un equipo de EPR (por ejemplo *X band klystron*) tiene un ancho de banda espectral de 8.8-9.6 Ghz, lo cual dificulta hacer un barrido de frecuencia. Debido a esto es más frecuente variar B_0 .

Los espectros de absorción de EPR se analizan como la primera derivada con la finalidad de mejorar la sensibilidad y la resolución.

Habitualmente para conocer los átomos involucrados en la coordinación ecuatorial se interpolan los valores experimentales de g y $A_{||}$ en las gráficas de Peisach and Blumberg. En estas se correlacionan valores de g y $A_{||}$ obtenidos para complejos que incluyen la participación de átomos de nitrógeno, oxígeno y azufre, combinados o no (Peisach & Blumberg, 1974).

I.3.4 Resonancia magnética nuclear en solución (RMN)

La RMN ha emergido como una técnica única que permite el estudio con resolución atómica de las proteínas en disolución, no perturba la muestra y tiene un amplio rango de aplicaciones, como determinar zonas de interacción, estudio de la dinámica de proteínas o la elucidación de la estructura terciaria (Amero, 2014)□.

En forma similar a otras técnicas espectroscópicas, en RMN la absorción de un fotón promueve una transición de un estado basal a un estado excitado, en particular de un espín nuclear. Esta transición depende de la existencia de un campo magnético externo, como en EPR. El estado excitado de RMN tiene un tiempo de vida 10^9 veces mayor que el estado excitado electrónico, lo cual genera líneas espectrales angostas, facilita estudios multidimensionales y permite la medición de la dinámica molecular en una amplia escala de tiempo.

La energía de los estados basal y excitado surge de la interacción del momento dipolo magnético nuclear con un campo magnético externo. La diferencia entre dos niveles de

energía es baja (1 parte en 10^6), por lo que su excitación se logra con ondas electromagnéticas de baja frecuencia (radiofrecuencias) (Rule & Hitchens, 2006)□.

El momento angular del espín (S) está cuantizado y se describe a través de su número cuántico I ($S = \hbar\sqrt{I(I + 1)}$), donde \hbar es la constante de Dirac. Un spin nuclear con valor 1/2 toma dos posibles valores de S en la componente Z, $+1/2\hbar$ o $-1/2\hbar$, lo que se corresponde con el número cuántico $m_z = +1/2$ y $m_z = -1/2$ respectivamente. Las transiciones de átomos con espines nucleares con estas características están permitidas pues cumplen con las reglas de selección. Los núcleos más frecuentemente utilizados en RMN son ^1H , ^{15}N , ^{13}C y ^{31}P , los que tienen espín $1/2$ y el ^2H que tiene espín 1 (Keeler, 2010; Kleckner & Foster, 2011)□.

La diferencia de energía entre dos estados puede ser obtenida de la relación:

$$\Delta E = \gamma H B_o / 4 \pi$$

donde γ es la constante giromagnética y B_o es el campo magnético externo.

De esta relación se deduce la conocida ecuación de Larmor $\omega = \gamma B$, donde ω se refiere a la frecuencia de resonancia de un núcleo.

La señal es medida como función del tiempo y su pérdida se refiere como *Free Induction Decay* (FID), (Figura I.2A). La pérdida de coherencia de la FID en el tiempo (I(t)) se evidencia por un decaimiento exponencial de la intensidad de la señal.

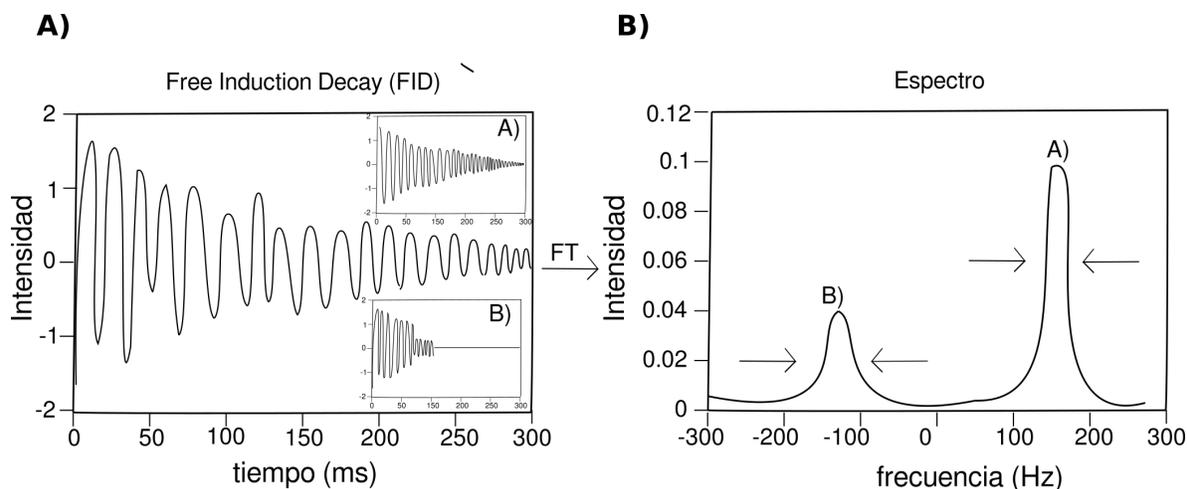


Figura 1.2: Tres variables importantes en RMN. FID, principal parámetro de la RMN. Brinda información de las señales individuales de la molécula. A la señal FID dependiente del tiempo (izquierda) se le aplica una transformada de Fourier, y se obtiene el cambio en la intensidad de la señal en el dominio de frecuencia (derecha). La señal que se relaja más rápidamente (B) tiene menos altura y es más ancha, mientras que la señal que se relaja más lentamente (A) tiene un pico alto y angosto. El área bajo la curva es igual para los dos picos, porque simulan dos poblaciones idénticas (Modificado de Kleckner and Foster, 2011).

La velocidad de decaimiento de la intensidad puede cuantificarse con la determinación de la constante de velocidad de la relajación transversal (R_2) mediante la relación $I(t) = \exp(-R_2 t)$. R_2 también es referido como el inverso de T_2 . La frecuencia de la señal inducida es exactamente igual a la frecuencia de la transición del espín nuclear.

A la FID se le aplica una transformada de Fourier para hacer una transformación de los datos del dominio tiempo al de frecuencia con lo que se obtiene un espectro característico de RMN. En este espectro se pueden identificar tres variables importantes: desplazamiento químico (δ), intensidad (I) y ancho del pico (λ) (Kleckner & Foster, 2011). □

El desplazamiento químico (δ) es la posición de la señal en el espectro de frecuencia. Brinda información sobre el entorno químico local del núcleo y es útil para identificar ángulos de torsión, enlaces de hidrógeno, elementos de estructura secundaria, entre otras. La intensidad (I) se cuantifica por la altura o volumen del pico y está asociada al número de núcleos que resuenan a una frecuencia dada o que experimentan ese entorno químico en particular. El ancho del pico a la mitad de la altura de éste (λ) reporta sobre las propiedades dinámicas de las moléculas en solución. Está relacionado con R_2 mediante la ecuación R_2 (Hz) = $1/T_2$ (s^{-1}) = $\lambda/2$ (Hz).

Los experimentos de RMN pueden ser en una o más dimensiones y se pueden lograr correlaciones entre núcleos iguales o distintos. Los multidimensionales consisten en la repetición de un experimento en la dimensión menor donde se varía el intervalo de tiempo entre ellos, lo que proporciona la frecuencia de la nueva dimensión. En un experimento tipo ^1H - ^{15}N HSQC (*Heteronuclear single quantum coherence spectroscopy*) se correlacionan los protones unidos a nitrógenos de la cadena principal (Amero, 2014).

Un requisito indispensable en el estudio de proteínas por RMN es asignar la proteína. En este proceso se hace corresponder cada señal del espectro de HSQC con un residuo de la proteína.

I.3.5 Calorimetría de titulación isotérmica (ITC)

La técnica calorimetría de titulación isotérmica (ITC, de sus siglas en inglés) permite la caracterización termodinámica de la interacción entre biomoléculas (Freyer & Lewis, 2008)□. La técnica se basa en la medición del calor absorbido o liberado durante el proceso de asociación entre moléculas a una temperatura determinada. En un solo experimento se pueden determinar la constante de unión, la estequiometría de la reacción y la entalpía de la asociación (Freyer & Lewis, 2008)□. El equipo consta de dos celdas, nombradas celda de referencia (CR) y celda de muestra (CM). La diferencia de temperatura entre ambas es mantenida constante mediante un controlador de temperatura, el cual utiliza un valor de potencia de referencia.

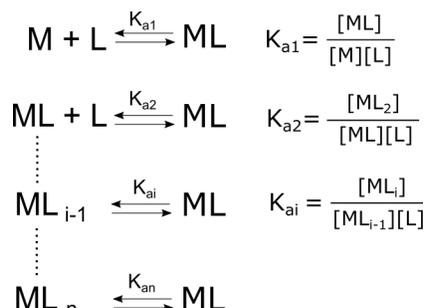
Durante el procedimiento experimental se realizan normalmente dos titulaciones independientes. En cada una se colocan en la celda de reacción la muestra de proteínas (P) o el tampón, respectivamente. En cada experimento se adicionan cantidades constantes del ligando (L) o sustrato de la reacción. Se registra el calor necesario para mantener constante la temperatura de la celda después de cada inyección de la molécula en estudio.

El calor absorbido o liberado durante la reacción de asociación molécula-proteína se obtiene como la diferencia entre el calor de la reacción molécula-proteína y el correspondiente a la reacción molécula-tampón (Nall et al., 1999)□. Los datos experimentales se ajustan a una ecuación que describe la unión de un ligando a una macromolécula.

El caso más sencillo es la unión de un ligando a un sitio de la proteína. El modelo para la descripción de la unión de un L a un sitio de una proteína puede describirse de la siguiente manera:



La unión puede también involucrar varios sitios de la macromolécula, pudiendo ser éstos idénticos o no. La unión secuencial a n sitios independientes y diferentes involucra las constantes $K_{a1}, K_{a2} \dots K_n$, las cuales se describen como:



Los parámetros se determinan resolviendo un polinomio ($P=1+K_{a1}[L]+K_{a1}K_{a2}[L]^2+\dots+K_{a1}K_{a2}K_n[L]^n$), mediante un método numérico (ITC200, n.d.)

El modelo secuencial no hace distinción de cual sitio ha sido saturado; solo describe el número total de sitios que han sido ocupados. El valor de n se elige, seleccionándose el mejor ajuste de acuerdo a los valores y errores obtenidos.

Los conceptos relacionados al proceso de reconocimiento molecular y las metodologías discutidas anteriormente, fueron aplicados en este trabajo de tesis para el estudio de tres sistemas proteicos. En los siguientes capítulos discutimos los resultados de la caracterización de la interacción proteína – ligando de (AtLEA4-5 – iones metálicos, LAOBP – aminoácidos, CreCBD – ácidos nucleicos).

I.4 Bibliografía

- Amero, C. (2014). Introducción a la Resonancia Magnética de Proteínas. Prado, 1-95.
- Bicout, D. J., & Szabo, A. (2008). Entropic barriers, transition states, funnels, and exponential protein folding kinetics: A simple model. *Protein Science*, 9(3), 452–465.
- Boucher, V., Buitink, J., Lin, X., Boudet, J., Hoekstra, F., Hundertmark, M., Leprince, O. (2010). MtPM25 is an atypical hydrophobic late embryogenesis-abundant protein that dissociates cold and desiccation-aggregated proteins. *Plant, Cell & Environment*, 33(3), 418–430.
- Bronowska, A. K. (2011). Thermodynamics of Ligand-Protein Interactions: Implications for Molecular Design. In *Thermodynamics – Interaction Studies – Solids, Liquids and Gases*, 1–45.
- Dickerson, R. E., Gray, H. B., & Haight, G. P. (1979). Coordination chemistry. In *Chemical Principles. Third edition*. 736-785.
- Du, X., Li, Y., Xia, Y. L., Ai, S. M., Liang, J., Sang, P., Liu, S. Q. (2016). Insights into protein–ligand interactions: Mechanisms, models, and methods. *International Journal of Molecular Sciences*, 17(2), 1–34.
- Dunker, K., Silman, I., Uversky, V., & Sussman, J. (2008). Function and structure of inherently disordered proteins. *Current Opinion in Structural Biology*, 18(6), 756–764.
- Freyer, M. W., & Lewis, E. (2008). Isothermal titration calorimetry: experimental design, data analysis, and probing macromolecule/ligand binding and kinetic interactions. *Methods in Cell Biology*, 84(07), 79–113.
- Gilson, M. K., & Zhou, H.-X. (2007). Calculation of Protein-Ligand Binding Affinities. *Annual Review of Biophysics and Biomolecular Structure*, 36(1), 21–42.
- Goyal, K., Tisi, L., Basran, A., Browne, J., Burnell, A., Zurdo, J., & Tunnacliffe, A. (2003). Transition from natively unfolded to folded state induced by desiccation in an anhydrobiotic nematode protein. *The Journal of Biological Chemistry*, 278(15), 12977–12984.
- ITC200, M. (n.d.). MicroCal™ iTC 200 System. *User Manual 29017607AA*, 1–326.
- Keeler, J. (2010). Understanding NMR Spectroscopy. In *Springer*.
- Khrapunov, S. (2018). The Enthalpy-entropy Compensation Phenomenon. Limitations for the Use of Some Basic Thermodynamic Equations. *Current Protein & Peptide Science*, 19(11), 1088–1091.
- Kleckner, I. R., & Foster, M. P. (2011). An introduction to NMR-based approaches for measuring protein dynamics. *Biochimica et Biophysica Acta - Proteins and Proteomics*, 1814(8), 942–968.
- Li, H., Xie, Y., Liu, C., & Liu, S. (2014). Physicochemical bases for protein folding, dynamics, and protein-ligand binding. *Science China Life Sciences*, 57(3), 287–302.
- Liu, S.-Q., Ji, X.-L., Tao, Y., Fu, Y.-X., Tan, D.-Y., & Zhang, K.-Q. (2012). Protein Folding, Binding and Energy Landscape: A Synthesis. *Protein Engineering*, 207–243.
- Martin, S., & Clements, J. (2013). Correlating Structure and Energetics in Protein-Ligand Interactions: Paradigms and Paradoxes. *Annual Review of Biochemistry*, 82(1), 267–293.

- Martin, S. R., & Schilstra, M. J. (2008). Circular Dichroism and Its Application to the Study of Biomolecules. In *Methods in Cell Biology* (Vol. 84, pp. 263–293).
- Mészáros, B., Simon, I., & Dosztányi, Z. (2009). Prediction of Protein Binding Regions in Disordered Proteins. *PLoS Comput Biol*, 5(5). e1000376.
- Nall, B. T., Pierce, M. M., & Raman, C. S. (1999). Isothermal Titration Calorimetry of Protein–Protein Interactions. *Methods*, 19(2), 213–221.
- Pancsa, R., & Tompa, P. (2012). Structural disorder in eukaryotes. *PlosOne*, 7(4), e34687.
- Parson, W. (2008). *Modern Optical Spectroscopy*. Springer Dordrecht Heidelberg London New York. 1–523.
- Persch, E., Dumele, O., & Diederich, F. (2015). Molecular recognition in chemical and biological systems. *Angewandte Chemie - International Edition*, 54(11), 3290–3327.
- Quintanar, L., & Rivillas-Acevedo, L. (2013). Studying Metal Ion–Protein Interactions: Electronic Absorption, Circular Dichroism, and Electron Paramagnetic Resonance. In M. Williams & T. Daviter (Eds.), *Protein-Ligand Interactions: Methods and Applications*. New York: Springer Science+Business Media. 267–296.
- Rule, G., & Hitchens, K. (2006). NMR spectroscopy. In *Fundamentals of Protein NMR Spectroscopy* (pp. 1–529).
- Szalainé Ágoston, B., Kovács, D., Tompa, P., & Perczel, A. (2011). Full backbone assignment and dynamics of the intrinsically disordered dehydrin ERD14. *Biomolecular NMR Assignments*, 5(2), 189–193.
- Tompa, P., & Fuxreiter, M. (2008). Fuzzy complexes: polymorphism and structural disorder in protein–protein interactions. *Trends Biochem Science*, 33, 2–8.
- Tompa, P., Szasz, C., & Buday, L. (2005). Structural disorder throws new light on moonlighting. *Trends Biochem Science*, 30, 484–489.
- Wright, P E, & Dyson, H. J. (1999). Intrinsically unstructured proteins: re-assessing the protein structure–function paradigm. *Journal of Molecular Biology*, 293(2), 321–331.
- Wright, Peter E, & Dyson, H. J. (2009). Linking Folding and Binding. *Current Opinion in Structural Biology*, 19(1), 31–38.

CAPÍTULO II

Interacción AtLEA4-5 - iones metálicos

Capítulo II

II. Interacción AtLEA4-5 – iones metálicos

II.1 Proteínas LEA (Late Embryogenesis Abundant), un grupo de IDPs

Dentro de los eucariontes, las plantas presentan un alto porcentaje de “desorden” estructural en sus proteínas (Pancsa & Tompa, 2012)□. Un grupo particular de las IDPs son las proteínas LEA (*Late Embryogenesis Abundant*). Las plantas sometidas a déficit hídrico ocasionado por cambios de temperatura, salinidad o sequía pueden sufrir pérdida intracelular de agua (Kovacs et al., 2008)□. Uno de los mecanismos de respuesta fisiológica al estrés osmótico es la expresión de las proteínas conocidas como LEA (Battaglia & Covarrubias, 2013)□. Estas proteínas también se acumulan en determinadas condiciones fisiológicas de las plantas, como durante el desarrollo de las semillas, polen y algunos estados de desarrollo de las raíces (Colmenero-Flores et al., 1999)□. Las proteínas LEA fueron identificadas por primera vez en semillas de algodón (Dure & Galau, 1981)□, y desde entonces se han encontrado en varias especies de plantas, bacterias, levaduras, arqueas y nemátodos (Battaglia & Covarrubias, 2013; Battaglia et al., 2008)□. En *Arabidopsis thaliana* (Brassicales: Brassicaceae; Heynh, 1842), una de las plantas modelo en biología molecular de plantas, se han encontrado 51 genes que codifican para proteínas LEA (Hundertmark & Hinch, 2008)□.

La mayoría de las proteínas LEA han sido clasificadas como hidrofílicas, de acuerdo a su índice de hidrofobicidad (<1) y al porcentaje de glicinas en su secuencia (>6%) (Garay-Arroyo et al., 2000)□. En general, las proteínas LEA presentan un alto contenido de aminoácidos pequeños, polares y cargados, mientras que carecen o contienen pocos residuos de triptófanos y cisteínas. Esta composición aminoacídica favorece el estado desplegado y la alta flexibilidad que adoptan en soluciones acuosas (Olvera-Carrillo et al., 2011a)□.

Las proteínas LEA se pueden clasificar en 7 grupos (LEA1-LEA7) de acuerdo con su similitud en secuencia de aminoácidos, siendo posible distinguir motivos característicos y únicos dentro de cada grupo (Battaglia & Covarrubias, 2013; Battaglia et al., 2008)□. En solución, según estudios espectroscópicos, la mayoría de las proteínas LEA se hallan en

forma de *random coil*, por lo que se han clasificado como proteínas intrínsecamente desplegadas.

Específicamente, el grupo 4 de las proteínas LEA de *A. thaliana* (AtLEA) (Tabla 1) está constituido por tres proteínas (Figura II.2) (Olvera-Carrillo et al., 2011b)□, que se conocen como AtLEA4-1, AtLEA4-2, AtLEA4-5 de acuerdo con la ubicación de sus genes en los cromosomas 1, 2 y 5 respectivamente.

Tabla 1: Características de los miembros del grupo 4 de las proteínas AtLEA (Modificado de Olvera-Carrillo et al., 2010)

Proteínas	Gen	# aa	Gravy	kDa	PI
AtLEA4-1	At1g32560	134	-1.042	14.9	8.95
AtLEA4-2	At2g35300	97	-1.156	10.5	9.67
AtLEA4-5	At5g06760	158	-0.815	16.2	9.45

En la tabla se muestra: nombre común (referido a otras nomenclaturas con las que se reconocen a las proteínas AtLEA4), gen (nombre de ADN genómico de cada proteína), #aa (cantidad de residuos en la secuencia de proteína), Gravy (índice de hidropatía según Kyle y Doolittle), kDa (masa molecular de cada proteína), PI (punto isoeléctrico).



Figura II.1: Alineamiento de secuencias de las proteínas LEA del grupo 4 de *Arabidopsis thaliana*: AtLEA4-5, AtLEA4-1, and AtLEA4-2. Los residuos totalmente conservados están enmarcados en rectángulos negros, mientras que los aminoácidos con carga positiva, negativa, las glicinas e histidinas se muestran en los colores azul, rojo, verde y púrpura, respectivamente. Se puede observar que el extremo amino está altamente conservado, mientras que el carboxilo es más variable.

Como se observa en la Figura II.1 la región carboxilo terminal está menos conservada, es de tamaño variable dentro del grupo y se predice que no forma ninguna estructura (Dure, 1993)□. Las tres proteínas muestran un bajo porcentaje de identidad de secuencia (54.64% entre AtLEA4-1 y AtLEA4-2, 29.69% entre AtLEA4-1 y AtLEA4-5, y 26.09% entre AtLEA4-2 y AtLEA4-5). La región amino está conservada y se ha visto que es capaz de

plegarse en presencia de disolventes que inducen plegamiento como trifluoretanol (TFE) (Culik et al., 2014; Rariy & Libanov, 1997) (Cuevas-Velázquez et al., 2016). En forma similar a otras proteínas LEA, las proteínas del grupo 4 abundan en todos los tejidos embrionarios de las semillas deshidratadas (Roberts et al., 1993).

II.1.1 Estudios de la interacción de las proteínas LEA con metales

Entre las funciones de las proteínas LEA en respuesta al estrés causado por frío, sequía o alta salinidad se ha descrito el secuestro de iones metálicos (Hara et al., 2004), lo cual se puede asociar con una posible actividad antioxidante. Se han utilizado varios sistemas redox (Cu-ascorbato, Fe(III)/H₂O₂) para la detección de esta actividad, así como cromatografía de afinidad con ligando inmovilizado (IMAC, de sus siglas en inglés) para determinar la especificidad (Hara et al., 2004; Liu et al., 2011).

La capacidad de las dehidrinas de secuestrar iones metálicos ha sido ampliamente estudiada. La proteína CuCOR19 (cold response de *Citrus unshiu*) mostró su capacidad para secuestrar radicales hidroxilos y peróxidos generados por los sistemas in vitro (Hara et al., 2004). Un análisis de la composición de aminoácidos de CuCOR19 sugirió que glicina, histidina, y lisina, principales residuos que componen muchas dehidrinas, son blanco del radical hidroxilo, lo que sugiere que los tres aminoácidos están involucrados en el secuestro de radicales libres (Hara et al., 2004). Además, mediante IMAC se determinó que CuCOR15 unía Fe(III), Co(II), Ni(II), Cu(II), y Zn(II). Con el uso de 27 péptidos, diseñados a partir de secuencias de dehidrinas de 14 especies de plantas, se encontró que la eficiencia en las actividades de reducción de ROS puede estar determinada principalmente por dos factores: el contenido de histidina y la longitud de los péptidos (Hara et al., 2013).

En estudios de la interacción de las dehidrinas de frijol y soya con Cu(II) mediante IMAC, se demostró que las proteínas LEA2 unían Cu(II) con la participación de residuos de histidinas (Krüger et al., 2002; Tunnacliffe & Wise, 2007). Por otro lado, estudios con AtHIRD11, una dehidrina de Arabidopsis, mostraron que la generación de peróxido de hidrógeno y radicales hidroxilo a partir de Cu(II) quedaba inhibida en presencia de esta proteína (Hara et al., 2013). Mutaciones en los residuos de His de AtHIRD11 causaron la

disminución de su actividad antioxidante, lo que indicó la importancia de las histidinas en la reducción de la generación de ROS. También, se ha descrito que una secuencia con dos His une metales con más fuerza que una con solo una His. Los motivos His-X-X-X-His, His-Glu-X-X-His y N-Asp-Ser-His se han caracterizado como sitios de unión de iones metálicos en muchas metaloproteínas (Gusman et al., 2001)□.

Por otra parte, mediante IMAC se determinó que la proteína ZmLEA3 de maíz (*Zea mays*), perteneciente al grupo 3 une Mn(II), Fe(III), Cu(II) y Zn(II) (Liu et al., 2013)□. Resultados similares se encontraron para las proteínas GmPM1 y GmPM9 de soya (*Glycine max* L.), pertenecientes al grupo 4, las cuales unen Fe(III), Ni(II), Cu(II) y Zn(II) (Liu et al., 2011)□. También se detectó su capacidad de secuestrar ROS mediante el sistema Fe(III)/H₂O₂.

El estudio del efecto que tienen los metales sobre el plegamiento de proteínas LEA se ha abordado mediante DC. En estudios realizados con la dehidrina AtHIRD11 se obtuvo una disminución de la señal a 200 nm (Hara et al., 2013)□. Sin embargo los espectros de absorción electrónica por CD de las proteínas GmPM1 y GmPM9 no sufrieron modificaciones en presencia de FeCl₃ (Liu et al., 2011)□.

El proceso de unión de las proteínas LEA a sus ligandos, en particular a metales, ha sido abordado también mediante ITC, y se han encontrado valores de K_d en el orden micromolar. Existe gran heterogeneidad en las K_d de diferentes metaloproteínas o proteínas/péptidos, asociados a enfermedades neurodegenerativas (Tabla 2).

Tabla 2: Ejemplos de proteínas que unen iones metálicos (Cu(II), Zn(II)).

Proteínas	K_d (Cu(II))	K_d (Zn(II))	Referencia
Catalasa bovina	657 μ M	ne	(Hao et al. 2015)
Transferina humana	ne	25.2 nM	(Terpstra et al. 2014)
Superóxido dismutasa humana	ne	1 nM	(Potter et al. 2007)
Péptidos beta amiloides	13 μ M	ne	(Sacco et al. 2012)
Prion	7 nM	ne	(Gogineni et al. 2015)

ne: no especificado en el artículo que se cita.

II.1.2.1 Estudios de caracterización de la proteína AtLEA4-5

La caracterización estructural de los miembros de la familia AtLEA4 de *Arabidopsis thaliana* se ha abordado mediante varios ensayos *in vivo* e *in vitro* (Olvera-Carrillo et al., 2010)□. Cuando se silenció la expresión de uno de los miembros de la familia AtLEA4, las plantas mostraron un fenotipo más susceptible al estrés hídrico. La mayor abundancia relativa de los ARNm de las tres proteínas AtLEA4 se encontró durante la fase de deshidratación de la semilla, mientras que los niveles de los transcritos se redujeron significativamente luego de su germinación (Olvera-Carrillo et al., 2010)□.

En cuanto a la caracterización funcional *in vitro* de la familia AtLEA4, se ha descrito que la proteína AtLEA4-5 es capaz de evitar la pérdida de actividad y la agregación de las enzimas lactato deshidrogenasa, citrato sintasa y malato deshidrogenasa en condiciones de deshidratación (Reyes et al. 2005)□. El mecanismo propuesto se basa en que en el complejo LEA-enzima, las LEA adquieren cierto grado de “orden” o de plegamiento que favorece el correcto plegamiento de las proteínas diana (Kovacs et al., 2013; Olvera-Carrillo et al., 2011a)□□. Además, se ha sugerido que secuencias consenso dentro de cada familia, están implicados en propiciar mayor especificidad a los procesos de reconocimiento y unión de las LEA con sus proteínas diana (Olvera-Carrillo et al., 2011a)□□.

Estos supuestos se han complementado con otros derivados de estudios sobre la capacidad de las AtLEA4 de formar estructuras secundarias en presencia de disolventes que inducen plegamiento como glicerol y TFE (Cuevas-Velázquez et al., 2016). Se ha sugerido que las proteínas de la familia LEA4 pueden encontrarse como IDP y que dependiendo de la disponibilidad de agua algunas regiones pueden sufrir transiciones a estados plegados (Lupas et al., 1991)□.

A pesar de que las proteínas AtLEA 4 no tienen una alta identidad de secuencia, presentan similitud en las funciones que realizan (Olvera-Carrillo et al., 2011a)□. La mayoría de trabajos relacionados a la interacción de proteínas LEA con metales, se han dirigido a identificar mediante IMAC los posibles ligandos metálicos y a determinar su capacidad de secuestrar o evitar la formación de ROS. Sin embargo, poco se conoce de las

características de la interacción en cuanto a los parámetros termodinámicos o a los motivos de la proteína involucrados.

Durante el proyecto de maestría se identificó mediante IMAC que AtLEA4-5 es capaz de unir Cu(II), Zn(II) y Ni(II) y se caracterizaron los posibles modos de coordinación a cobre mediante UV-vis, CD y EPR (French-Pacheco, 2015). También se determinó utilizando ITC que la proteína contiene varios sitios de unión con afinidades en el orden micromolar. Sin embargo, los termogramas obtenidos mostraron que además de la unión, ocurrían otros procesos en la celda de reacción como la oligomerización mediada por los iones metálicos. Estos resultados se completaron y reanalizaron en este trabajo doctoral, determinando el efecto que tiene la interacción con iones metálicos en la estructura secundaria y estado oligomérico de la proteína. Además, se estudiaron algunos fragmentos peptídicos para intentar identificar los motivos involucrados en la interacción.

II.2 Justificación del proyecto AtLEA4-5 – iones metálicos

Las proteínas LEA, un grupo particular de las IDPs, se han asociado a funciones como evitar la inactivación y agregación de enzimas, estabilización de membranas biológicas y el secuestro de metales, entre otras. La correlación altamente significativa entre la acumulación de LEA y, en general de las hidrofílicas, con el déficit de agua en organismos de todos los reinos taxonómicos en combinación con la identificación de motivos conservados dentro de los grupos de LEA, hacen a estas proteínas un buen modelo para el estudio de las propiedades fisicoquímicas de las IDPs (Olvera-Carrillo et al., 2011a)□.

Específicamente, el estudio de la interacción de la familia de proteínas AtLEA4 con iones metálicos nos aportó elementos importantes para la comprensión del fenómeno de unión pues se determinó la especificidad, posibles modos de coordinación y algunos parámetros termodinámicos. Sin embargo, seguimos sin conocer los sitios específicos de unión, el efecto de esta unión sobre el plegamiento y oligomerización de las proteínas, y la importancia de que las LEA unan metales *in vivo*. Determinar estos aspectos se dificulta pues de acuerdo a nuestros resultados previos existen varios sitios de unión de los iones metálicos y las secuencias de estas proteínas tienen un alto porcentaje de residuos que podrían participar en esta unión.

En este proyecto se estudió el efecto de los iones metálicos sobre la estructura secundaria y cuaternaria de la proteína AtLEA4-5. Se determinaron las constantes de afinidad siguiendo distintas metodologías. Se usaron péptidos sintéticos, que contenían los residuos de His para estudiar las posibles zonas de interacción. Se emplearon distintas técnicas espectroscópicas buscando determinar los residuos involucrados en la interacción y caracterizar molecularmente la unión, lo que pudiera contribuir al entendimiento del funcionamiento de estas proteínas.

II.3 Hipótesis

Fragmentos peptídicos de AtLEA4-5 unidos a Cu(II) reproducen las características espectroscópicas del complejo proteína-Cu(II).

II.4 Objetivos

Objetivo general:

Realizar la caracterización espectroscópica y calorimétrica de la interacción de AtLEA4-5 con Cu(II).

Objetivos particulares:

1. Caracterizar la interacción de AtLEA4-5 con iones metálicos mediante UV-vis, CD, EPR e ITC.
2. Caracterizar la oligomerización de AtLEA4-5 inducida por la interacción con iones metálicos.
3. Síntesis y purificación de fragmentos peptídicos de AtLEA4-5.
4. Determinar las zonas de interacción con Cu(II) de los péptidos mediante RMN.
5. Determinar los modos de coordinación del Cu(II) en la proteína y fragmentos peptídicos de AtLEA4-5 mediante espectroscopía de absorción electrónica en la región UV-Vis y DC.
6. Determinar las constantes de afinidad y parámetros termodinámicos de la interacción de los péptidos con Cu(II) mediante ITC.

II.5 Métodos

II.5.1 Caracterización espectroscópica de la interacción AtLEA4-5 – Cu(II)

II.5.1.1 Absorción UV-Vis, dicroísmo circular y EPR

Los espectros de absorción electrónica de UV-vis y DC se adquirieron en un espectrofotómetro UV-visible Agilent 8453 con arreglo de diodos y en un espectropolarímetro Jasco J-825 respectivamente. Se usaron muestras de AtLEA4-5 a 120 μM en 10 mM NEM (4-ethylmorpholine) pH 7.5. La titulación se realizó usando 0.5 equivalentes de CuCl_4 . Los cambios se monitorearon entre 200 y 800 nm. Cada experimento se realizó por duplicado.

Los espectros de EPR se colectaron en una frecuencia de microondas de 9.5 Ghz (banda X) usando un equipo EMX Plus Bruker System, a 150 K, equipado con un sistema de temperatura variable ER4131VT. Las condiciones de cada corrida fueron: potencia de microondas 10 mW, amplitud modulada 5 G, frecuencia de modulación 100 kHz, constante de tiempo 327 ms y tiempo de conversión 82 ms. Una cantidad progresiva de equivalentes de CuCl_4 se adicionaron a muestras de AtLEA4-5 (120 μM in 10 mM NEM, pH 7.5).

II.5.1.2 Determinación de la constante de unión utilizando curvas de EPR

Para estimar una constante de unión se graficó el promedio de las señales de EPR a $g = 2.12$ y $g = 2.45$ como función de la concentración de Cu(II). Los valores normalizados se ajustaron mediante una aproximación no lineal a una ecuación hiperbólica:

$$y = \frac{p[\text{Cu}]}{(K_b + [\text{Cu}])}$$

donde: y es la absorbancia, $[\text{Cu}]$ es la concentración de Cu(II), p es el máximo de unión específica, K_b es la constante de disociación aparente.

II.5.2 Caracterización del estado oligomérico de AtLEA4-5 en ausencia y presencia de Cu(II).

Se estudió el estado oligomérico de AtLEA4-5 mediante la determinación del radio hidrodinámico (R_h) utilizando un equipo de dispersión dinámica de luz (DLS, de sus siglas

en inglés) (Zetasizer Malvern Nano ZSP), con un ángulo de dispersión de 173°. Se utilizaron muestras de AtLEA4-5 a 120 μM en 10 mM NEM, pH 7.5 en ausencia y presencia de 1, 2, 3 o 4 equivalentes de CuCl_2 . También se utilizaron muestras a 12 y 50 μM . A partir de las curvas de pérdida de intensidad de la señal en el tiempo se obtuvieron los coeficientes de difusión traslacionales, D (Stetefeld et al., 2016). Los R_h se obtuvieron de los D , mediante de la ecuación de Einstein-Stokes:

$$R_h = \frac{T \cdot K_B}{6\pi\eta D}$$

donde: K_B es la constante de Boltzmann, T es la temperatura, y η es la viscosidad del medio.

Se utilizó una viscosidad de 0.8872 cP y una temperatura de 25°C. Cada muestra se analizó por triplicado, tomando en cada caso 5 corridas de 10 mediciones de 10 s de duración. Los datos se analizaron mediante los métodos *cumulant* y de distribución implementados en el software SEDPHAT/SEDFIT.

El estimado del R_h para una IDP se calculó usando la siguiente relación (Stetefeld et al., 2016):

$$R_h = (1.24 \times F_{\text{PRO}} + 0.904) \times (0.00759 \times |Q| + 0.963) \times 2.49 \times N^{0.509}$$

donde: N es el número de residuos, F_{PRO} es la fracción de residuos de prolinas, $|Q|$ es la carga neta soluble.

II.5.3 Diseño, síntesis y purificación de los péptidos de AtLEA4-5.

Para completar la caracterización del complejo AtLEA4-5 – iones metálicos los potenciales sitios de unión se estudiaron mediante distintas metodologías. Las secuencias de los péptidos sintetizados corresponden a fragmentos de la proteína AtLEA4-5 que contienen los siete residuos de His. Para discriminar el papel de las His en la coordinación de los iones metálicos se diseñaron péptidos en los que se sustituyó cada His por una Ala. Se seleccionaron los fragmentos 61-69, 82-110 y 134-146.

MQSMKETASNIAASAKSGMDKTKATLEEKAEKMKTRDPVQKQMATQVKEDKINQAE

MQKRETRQHNAAMKEAAGAGTGLGLGTATHSTTGQVGHGTGTHQMSALPGHGTGQL TDRVVEGTAVTDPGRNTGTGRRTTAHNTHVGGGGATGYGGTGGGYTG		
Fragmento	Secuencia	Nombre
61-69	ETRQHNAAM	LEA61-69 (H65A)
82-110	TATHSTTGQVGHGTGTHQMSALPGHGTG	LEA82-110
82-110	TATHSTTGQVGHGTGTHQMSALPGHGTG	LEA82-110 (H106A)
82-110	TATHSTTGQVGHGTGTHQMSALPGHGTG	LEA82-110 (H98A)
82-110	TATHSTTGQVGHGTGTHQMSALPGHGTG	LEA82-110 (H93A)
82-110	TATHSTTGQVGHGTGTHQMSALPGHGTG	LEA82-110 (H85A)
134-146	RTTAHNTHVGGGG	LEA134-146
	RTTAANTHVGGGG	LEA134-146 (H139A)
	RTTAHNTAVGGGG	LEA134-146 (H141A)

La síntesis lineal en fase sólida de los péptidos se llevó a cabo usando la estrategia N-Fmoc. Como soporte sólido polimérico se usó una resina Rink Amida, la cual tiene una capacidad de sustitución de 0.40 mmol/g. Se utilizaron los disolventes diclorometano (DCM) y N,N'-dimetilformamida (DMF). En las reacciones de activación se utilizaron el activador N,N'-diisopropilcarbodiimida (DIC), oxyma para evitar racemización, DMF y el aminoácido correspondiente. En el paso de activación se forma un éster entre el aminoácido a acoplar y el activador, quedando así el aminoácido listo para reaccionar con el grupo amino libre. En la Figura II.2 se muestra un esquema del proceso de síntesis.

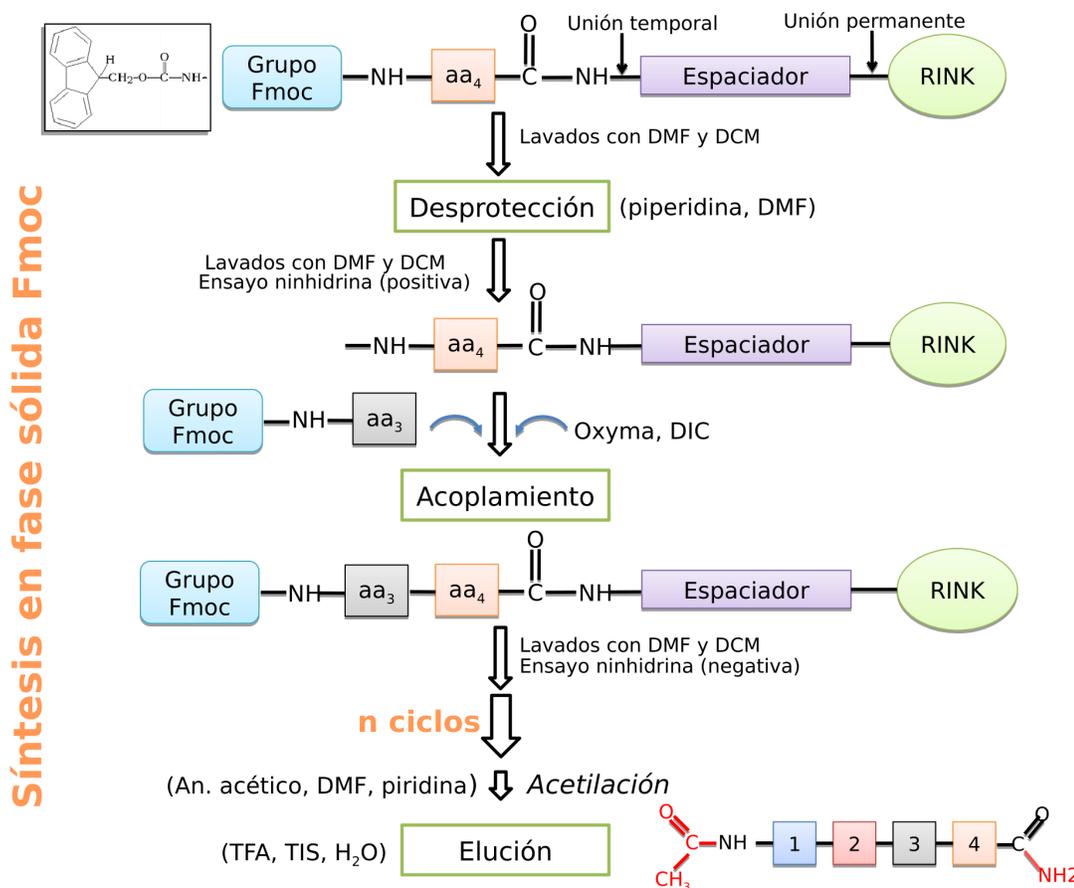


Figura II.2: Protocolo de síntesis de péptidos en fase sólida siguiendo la química Fmoc. La síntesis se hizo del extremo carboxilo al amino. Tras acoplar el primer aminoácido (último de la secuencia de interés) a la resina, este deja disponible su grupo amino, pero protegido con un grupo Fmoc. Para eliminar este grupo se realizó un paso de desprotección con lo cual queda disponible el extremo amino del aminoácido acoplado. Luego de la activación del siguiente aminoácido, se adicionó toda la mezcla de activación al reactor (donde se realiza toda la síntesis) y se incubó con la resina por al menos 4 h a temperatura ambiente en agitación constante. Luego se hicieron varios lavados con DMF y DCM para eliminar el excedente de aminoácido. Nuevamente se realizó la liberación del grupo Fmoc del extremo amino. Básicamente, los ciclos de síntesis consistieron en realizar un paso de desprotección, seguido de otro de acoplamiento, hasta obtener la secuencia deseada. Para comprobar el éxito de cada paso de acoplamiento y de desprotección se realizaron pruebas de ninhidrina, la cual consiste en la reacción de grupos amino libres con el hidrato de ninhidrina. Se forma un aducto que presenta una coloración azul. Entonces, en los pasos de desprotección se espera obtener una prueba positiva (color azul), mientras que en las reacciones de acoplamiento una negativa (color amarillo). Antes del corte final se acetiló el extremo amino. Esta modificación se realizó para evitar interacciones indeseadas del extremo amino con iones metálicos y para mimetizar que en la secuencia de la proteína los extremo amino y carboxilo de los péptidos están formando un enlace peptídico.

El corte final del péptido de la resina se realizó con un tratamiento con 95% ácido trifluoroacético (TFA), 2.5% triisopropilsilano (TIS), 2.5% agua y 2.5 % 1,2-etanoditiol (EDT) (si la secuencia contenía Cys, His, Met) durante 2 horas con agitación constante a temperatura ambiente. La reacción con secuencias con arginina se dejaron por 4 h.

La separación del péptido se realizó por medio de su precipitación en éter etílico frío. Se realizaron sucesivos lavados con éter seguidos de centrifugaciones.

II.5.3.1 Purificación de los péptidos de AtLEA4-5.

Los péptidos se purificaron mediante cromatografía líquida de alta resolución en una columna C18 de fase reversa (RP-HPLC) con todos los solventes con un grado de pureza HPLC. Tras un paso de equilibración de la columna con agua desionizada 0.05% TFA (A), se inyectaron 4 mL de péptidos disueltos en agua. Los péptidos se eluyeron mediante un gradiente lineal 0 a 70 % acetonitrilo 0.05% TFA (B), a una velocidad de flujo de 3.0 mL/min. La columna se lavó con acetonitrilo y se equilibró con la solución A.

La detección de las señales se realizó por absorción a 214 nm mediante un detector UV-vis.

Los péptidos purificados se liofilizaron para su posterior utilización.

II.5.3.2 Espectrometría de masas ESI

La determinación de la masa molecular relativa de los péptidos sintetizados se realizó mediante espectrometría de masas (MS), utilizando el espectrómetro de masas Electrospray Ionization (ESI). Ver anexo III.1

II.5.4 Caracterización de la interacción péptido – iones metálicos mediante absorción UV-Vis y dicroísmo circular

Los péptidos disueltos en tampón NEM (4-etilmorfolina) 10mM pH 7.5, se titularon con 0.5 en 0.5 equivalentes de CuCl_2 (Cu(II)) hasta lograr la saturación. Los cambios se siguieron por absorción electrónica UV-Vis y DC. Los espectros se colectaron en un espectrofotómetro Agilent 8453 y en un espectropolarímetro JASCO J-815.

II.5.5 Determinación de las zonas de interacción de los péptidos mediante RMN

Los espectros de RMN se adquirieron a 298 K usando el equipo de 500 MHz (11.7 T) perteneciente al Laboratorio Nacional de Estructuras de Macromoléculas (LANEM). La

asignación de los desplazamientos químicos de los átomos de péptidos se realizó mediante experimentos de tipo ^1H - ^{14}N TOCSY para permitir la identificación de los protones de cada residuo en ausencia y presencia de iones metálicos.

Los experimentos se analizaron en NMRpipe (Delaglio et al., 1995)□ y CARR (Keller, 2004)□ para obtener la asignación de los espectros.

II.5.6 Determinación de la termodinámica de la interacción de péptidos – iones metálicos mediante ITC.

En cada experimento, se utilizó un volumen de inyección de 2 μL del titulante y 50 μM de proteína. Se empleó un intervalo de 180 s entre inyecciones para permitir una equilibración completa. La titulación blanco consistió en la solución valorante idéntica pero que contenía sólo la solución tampón en la celda de muestra. La solución en la celda muestra se agitó a 750 rpm para asegurar un mezclado rápido. Los experimentos se realizaron en un equipo MicroCalTMiTC₂₀₀ a 25°C. Los datos se analizaron con el empleo del software SEDPHAT y graficados utilizando el programa GUSI (Zhao et al., 2015).

II.6 Resultados: Proyecto interacción LEA4-5 – iones metálicos

La proteína AtLEA4-5 es capaz de unir Cu(II), Zn(II) y Ni(II) de acuerdo a estudios de interacciones usando IMAC. La caracterización espectroscópica de la interacción (UV-vis, DC y EPR) arrojó que amidas desprotonadas y residuos de His participaban en la esfera de coordinación de Cu(II) (French-Pacheco, 2015)□. Además, mediante calorimetría se determinó que las K_d aparentes están en el orden micromolar.

Teniendo en cuenta estos antecedentes se realizó un análisis de los datos espectroscópicos con un enfoque estructural, lo cual permitió la caracterización de la estructura secundaria y del estado oligomérico de AtLEA4-5 en presencia de iones metálicos. Además, se estimó la K_d aparente a partir de las curvas obtenidas durante las titulaciones espectroscópicas, las que solo son sensibles a la formación de complejo metálico, sin que las señales se afecten por algún otro proceso como la oligomerización. Todo lo anterior nos permitió proponer un modelo de interacción y de oligomerización de la proteína.

II.6.1 Caracterización estructural de AtLEA4-5

La utilización de herramientas bioinformáticas y experimentales permitió el análisis estructural de AtLEA4-5. La predicción de desorden basada en la secuencia de AtLEA4-5 y en su espectro de DC, evidenció el carácter intrínsecamente desordenado de la proteína (Figura II.3). Los resultados están en concordancia con publicaciones previas (Cuevas-Velázquez et al., 2016). Utilizando el programa CIDER (Intrinsically Disordered Ensemble Relationships) la proteína AtLEA4-5 fue clasificada en la región R1 (Holehouse et al., 2017)□. Esta región se corresponde a estructuras débilmente cargadas y colapsadas. El análisis de los datos de DC usando el programa CAPITO señaló que AtLEA4-5 puede adquirir una estructura en forma de pre-glóbulo fundido. En consecuencia, AtLEA4-5 puede contener algunos elementos de estructura secundaria pero no tiene una estructura plegada estable. Además, está parcialmente compacta y no totalmente extendida.

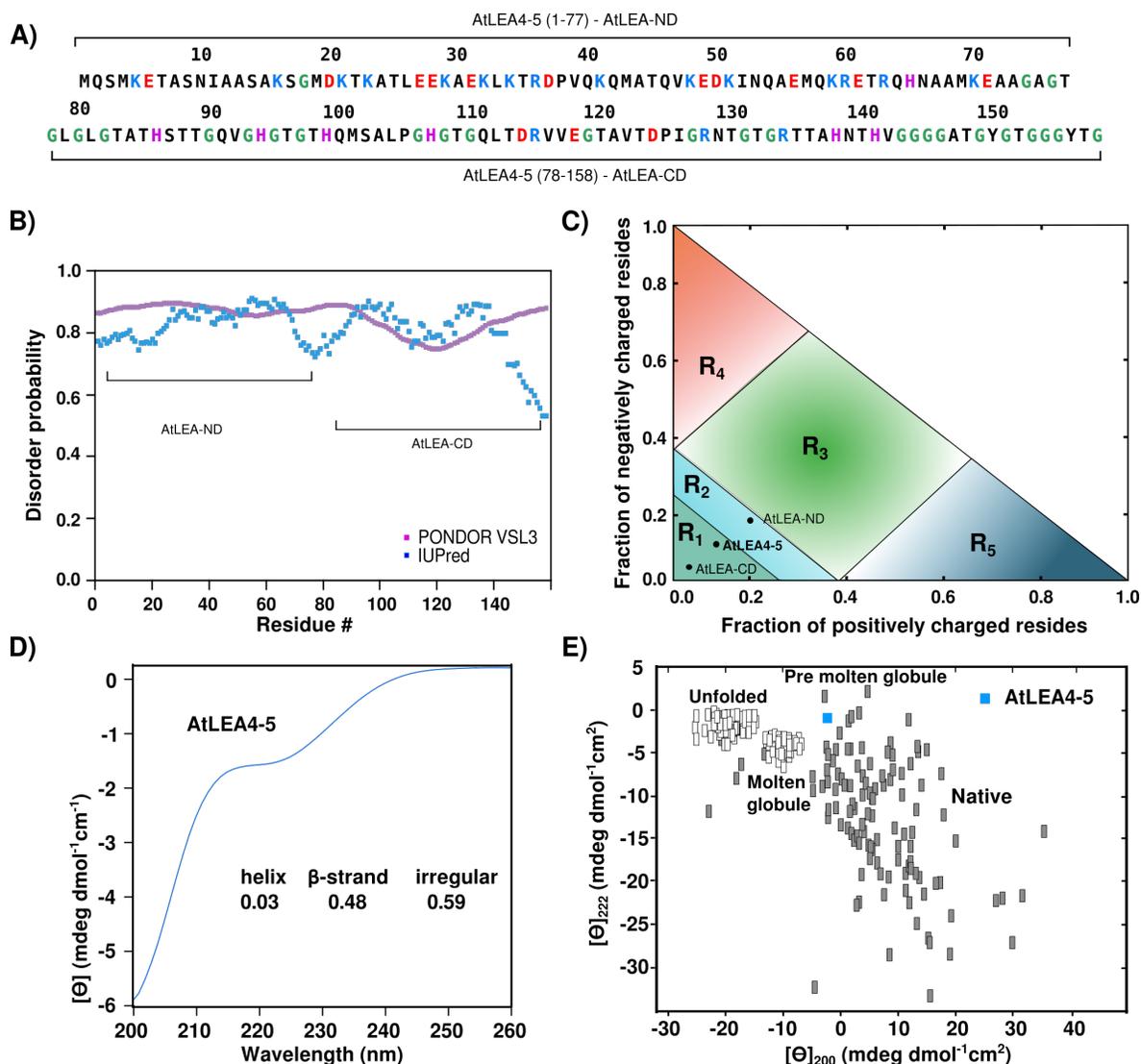


Figura II.3: (A) Secuencia de AtLEA4-5. Los residuos con carga positiva están en azul, los residuos con carga negativa están en rojo, las glicinas están en verde y los residuos de histidina están en púrpura. Se identifican los dominios N-terminal y C-terminal. (B) Probabilidad de desorden predicha de AtLEA4-5 por los métodos IUPred y PONDOR (Dosztányi et al., 2005; Xue et al., 2010). Ambas metodologías muestran valores de probabilidad de desorden por encima del umbral ($p = 0.5$). (C) Clasificación de AtLEA4-5 usando CIDER (siglas en inglés) (Holehouse et al., 2017). La región R1 corresponde a poliamfolitos y polielectrolitos débiles, la región R2 corresponde a secuencias colapsadas o expandidas dependientes del contexto, la región R3 corresponde a poliamfolitos fuertes, mientras que las regiones R4 y R5 corresponden a polielectrolitos fuertes cargados negativa y positivamente respectivamente (D) Espectros de DC y deconvolución de estructuras secundarias de AtLEA4-5. Se observa la señal característica de una estructura random coil (mínimo a $\lambda 200$ nm) (E) Análisis de datos de DC utilizando CAPITO.

La adición de cualquiera de los tres tipos de iones metálicos no modificó los espectros de DC, lo que sugirió que ninguno de los iones metálicos probados indujo cambios en la estructura secundaria de AtLEA4-5 bajo estas condiciones. Tampoco hubo variación en la tendencia de AtLEA4-5 de comportarse como una estructura pre-glóbulo fundido.

II.6.2 Caracterización espectroscópica de la interacción de AtLEA4-5 con Cu(II).

El análisis combinado de los datos de UV-vis, DC y EPR sugirió que AtLEA4-5 une iones Cu(II) mediante dos modos de coordinación (4N y N3O), en los que posiblemente estén involucrados nitrógenos de residuos de histidinas o amidas desprotonadas y oxígenos pertenecientes a grupos carbonilos, Ser, Thr o al agua (French-Pacheco, 2015).

Utilizando las curvas de las titulaciones espectroscópicas se estimó el valor de la K_d aparente del complejo Cu(II) – AtLEA4-5. La ventaja de utilizar esta aproximación es que la determinación sólo se basa en la formación del complejo metálico. Por el contrario las determinaciones calorimétricas tuvieron, en nuestro caso, el inconveniente de que los calores observados tenían la contribución de la unión y de la oligomerización de la proteína. En concordancia con los datos termodinámicos, la K_d estimada a partir de los datos de EPR es del orden micromolar ($\sim 300 \mu\text{M}$) (Figura II.4).

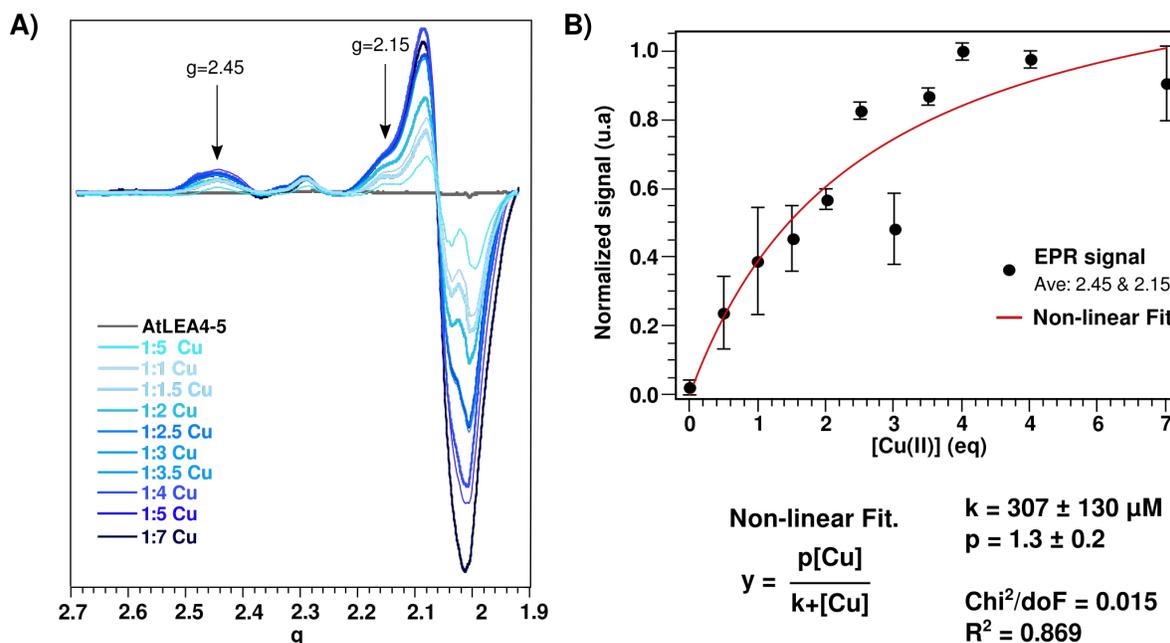


Figura II.4: Estimación de la K_d aparente del complejo AtLEA4-5-Cu(II) a partir de datos EPR. (A) Espectros EPR de AtLEA4-5 en ausencia (línea gris) o presencia de 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 5 y 7 equivalentes de Cu (II) (azul claro a azul oscuro). Los cambios de intensidades en $g = 2.45$ y $g = 2.15$ fueron seleccionados para estimar la constante de unión. Estos valores fueron promediados y normalizados. (B) Cambio de la señal en función del incremento de equivalentes de Cu(II). Los datos se ajustaron a una ecuación hiperbólica (línea roja).

II.6.3 Caracterización del estado oligomérico de AtLEA4-5. Mecanismo de agregación inducida por Cu(II).

Aun cuando la unión del metal no afectó la estructura secundaria de la proteína AtLEA4-5, evaluamos si podía haber una afectación en la organización cuaternaria de esta.

La metodología de DLS es muy útil para estudiar el estado oligomérico de las proteínas, sin embargo en el caso particular de las IDPs resulta difícil obtener un tamaño absoluto de Rh. Como referencia, utilizamos una ecuación empírica propuesta para estimar el radio hidrodinámico de una IDP monomérica (Tomasso et al., 2016) y calculamos el coeficiente de difusión traslacional correspondiente ($R_h = 3 \text{ nm}$, $D = 8.2 \times 10^{-7} \text{ cm}^2/\text{s}$).

La concentración celular de las proteínas LEAs es difícil de determinar debido al hecho de que se expresan diferencialmente durante la embriogénesis y en respuesta al déficit hídrico (Olvera-Carrillo et al., 2011); sin embargo, se ha reportado una concentración de $\approx 250 \mu\text{M}$ (Roberts et al., 1993). Nosotros utilizamos concentraciones en el mismo rango para tener una aproximación más fisiológica. Las mediciones de DLS de AtLEA4-5 a $120 \mu\text{M}$ mostraron una curva cuyos datos son polidispersos. En consecuencia, el ajuste requirió de tres componentes con coeficientes de difusión aparentes de $D = 8.26 \times 10^{-6}$, $D = 1.95 \times 10^{-7}$, y $D = 0.07 \times 10^{-7} \text{ cm}^2 \text{ s}^{-1}$ que corresponden a R_h aparentes de 2.98, 12.6 y 351 nm, respectivamente (Figura II.5A). De esto concluimos que bajo estas condiciones AtLEA4-5 existe en diferentes estados oligoméricos, que incluyen monómeros, probablemente tetrámeros y oligómeros de orden superior. Una concentración de proteína aún más baja todavía da como resultado curvas con datos polidispersos (Figura II.5 B y C). Otras proteínas LEA han mostrado esta variabilidad en su estructura cuaternaria en solución (Rivera-Nájera et al., 2014; Liu et al., 2017).

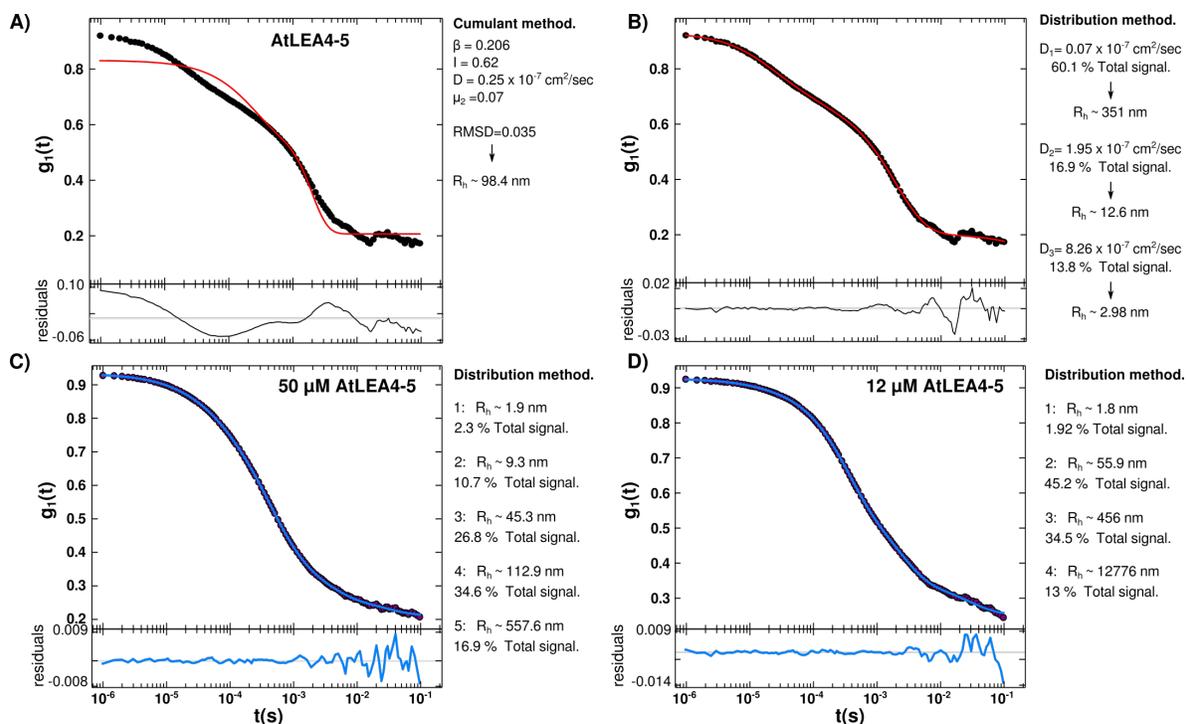


Figura II.5: Análisis de datos de DLS de AtLEA4-5. A partir de la curva de correlación se obtuvo el coeficiente de difusión. Para la determinación de los R_h se utilizaron dos métodos no lineales que permiten ajustar los datos a múltiples exponenciales decauyentes (*cumulant* y *distribución*). El método acumulativo supone una población de partículas única. El primer "acumulativo" corresponde a un coeficiente de difusión promedio, mientras que el segundo corresponde a la varianza. El método de distribución genera una distribución de moléculas con diferentes propiedades hidrodinámicas que pueden describirse mediante parámetros individuales. La señal DLS es muy sensible a la presencia de moléculas grandes debido al hecho de que la luz dispersada es proporcional a la sexta potencia del diámetro (A) Las mediciones de DLS de las soluciones de AtLEA4-5 mostraron una curva con datos polidispersos que no podía ser descrita con una sola población (método acumulativo). El mejor ajuste se obtuvo utilizando tres poblaciones diferentes en la distribución. (B) Mediciones DLS de soluciones AtLEA4-5 a 50 μM y a 12 μM (C). Se muestran los parámetros obtenidos y el residual del ajuste.

El estudio de una IDP mediante DLS tiene el sesgo de que es difícil correlacionar el coeficiente de difusión con el tamaño absoluto de la proteína. Sin embargo, el coeficiente se puede usar de manera confiable para revelar cambios debido a un proceso de oligomerización. Las mediciones de DLS de AtLEA4-5 en presencia de diferentes concentraciones de Cu(II) produjeron una disminución en los coeficientes de difusión traslacional lo cual es consistente con un aumento de tamaño debido a la oligomerización inducida por la unión de metal (Figura II.6). Los coeficientes de difusión de los complejos metálicos se determinaron mediante el método acumulativo, es decir, mediante una población promedio. Se obtuvieron valores de: $D = 0.134 \times 10^{-6} \text{ cm}^2 \text{ s}^{-1}$, $D = 0.1 \times 10^{-6} \text{ cm}^2 \text{ s}^{-1}$, $D = 0.06 \times 10^{-6} \text{ cm}^2 \text{ s}^{-1}$ y $D = 0.03 \times 10^{-6} \text{ cm}^2 \text{ s}^{-1}$ para las estequiometrías 1:1, 1:2, 1:3 y 1:4, respectivamente. Todos estos datos corresponden a oligómeros de gran tamaño.

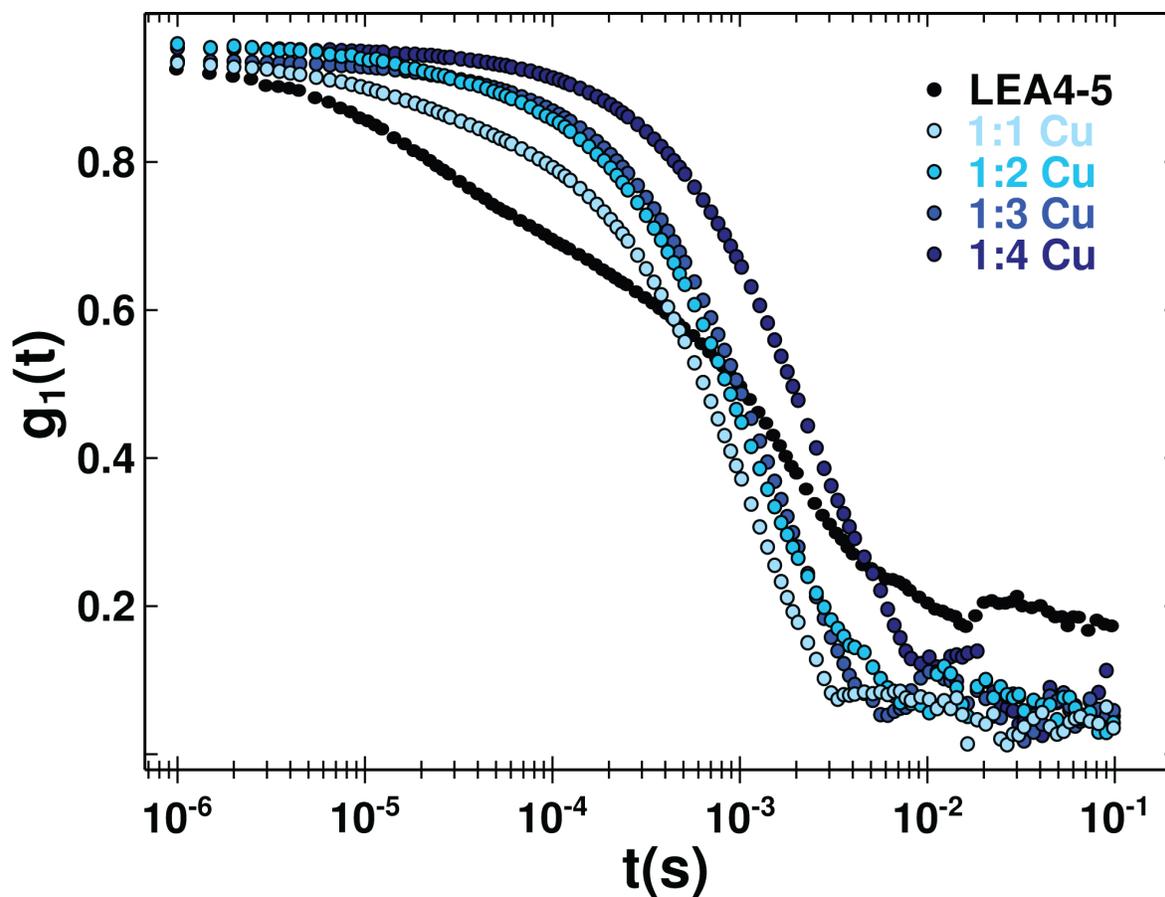


Figura II.6: Función de correlación de AtLEA4-5 en ausencia (línea negra) o presencia de 1, 2, 3 y 4 equivalentes de Cu (II) (azul claro a azul oscuro). Los datos se utilizaron para obtener los coeficientes de difusión traslacional mediante las metodologías acumulación y distribución. Las mediciones de AtLEA4-5 en presencia de diferentes concentraciones de Cu(II) produjeron un desplazamiento hacia la derecha de las curvas, lo que indica un aumento de tamaño debido a la oligomerización inducida por la unión del metal.

Los resultados obtenidos nos permitieron proponer un modelo de unión de AtLEA4-5 a iones metálicos. Sugerimos que la interacción se basa en el reconocimiento de un ensamble de conformaciones (Figura II.7).

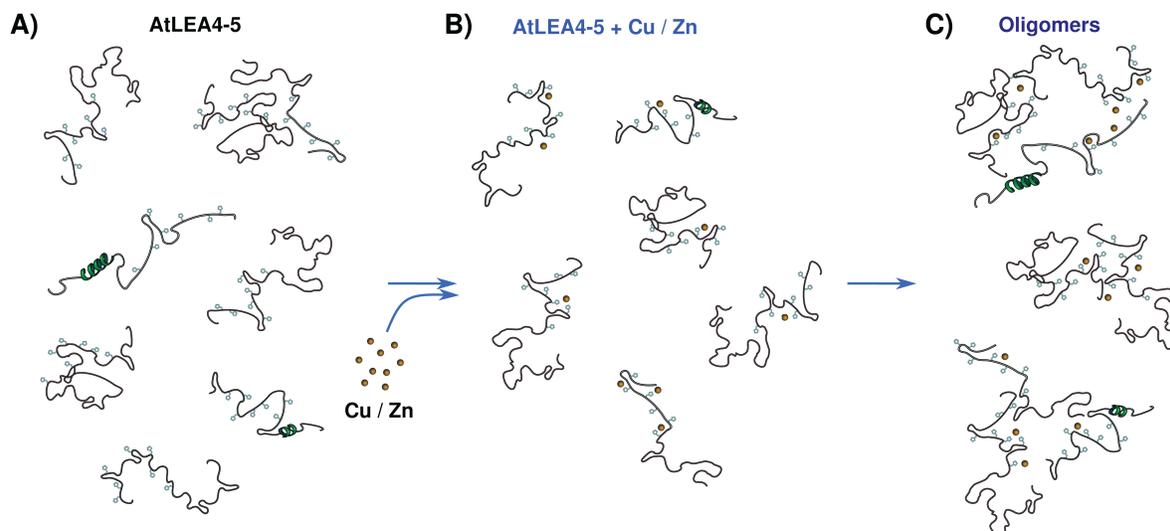


Figura II.7: (A) AtLEA4-5 puede tener un conjunto de conformaciones accesibles para la unión. (B) Los metales interactúan con afinidad moderada, uniéndose a residuos de His en múltiples formas y múltiples sitios de unión, sin formar un único sitio de unión bien definido. (C) Se representa la formación de especies oligoméricas en la medida que aumenta la concentración de iones metálicos.

Los iones metálicos interactúan principalmente con los residuos de His de AtLEA4-5 con moderada afinidad. Reconocen distintas conformaciones de la proteína y están coordinados en múltiples sitios de unión. Estos distintos modos de coordinación no forman un sitio bien definido, más bien pueden existir como un *fuzzy complex*. Un modelo de interacción entre IDPs y pequeños ligandos en el que múltiples conformaciones de ligando reconocen múltiples conformaciones de proteínas ha sido propuesto anteriormente (Jin et al., 2013). Sin embargo, en el caso de AtLEA4-5 pareciera que los iones metálicos reconocen preferencialmente a His. El incremento en la concentración de iones metálicos induce la formación de especies oligoméricas de gran tamaño.

El número exacto y las secuencias responsables de la unión aún no han sido esclarecidos. En este sentido, abordamos el estudio de los posibles sitios de unión mediante la síntesis de distintos péptidos que cubren fragmentos de la proteína que contienen residuos de His.

II.6.4 Síntesis y purificación de los péptidos de AtLEA4-5

Los fragmentos peptídicos de AtLEA4-5 se obtuvieron mediante síntesis química y se analizaron mediante espectrometría de masas. Se obtuvo una excelente correlación entre la masa molecular teórica y la experimental (Tabla 3). Los péptidos sintetizados tienen la característica de tener su extremo amino acetilado y el carboxilo amidado.

Tabla 3: Secuencia y masa molecular teórica y experimental de péptidos de AtLEA4-5

Nombre	MM teórica (Da)	MM experimental (Da)
LEA61-69 (H65A)	1098.208	1098.33
LEA82-110	2726.258	2726.25
LEA82-110 (H106A)	2598.193	2528.18
LEA82-110 (H98A)	2598.193	2528.18
LEA82-110 (H93A)	2598.193	2528.18
LEA82-110 (H85A)	2598.193	2528.18
LEA134-146	1305.376	1305.21
LEA134-146 (H139A)	1239.314	1239.65
LEA134-146 (H141A)	1239.314	1239.72

II.6.5 Determinación de los modos de coordinación del Cu(II) en la proteína y fragmentos peptídicos de AtLEA4-5

Luego de determinar que la proteína AtLEA4-5 es capaz de unir a los iones metálicos Cu(II), Zn(II) y Ni(II) se hicieron análisis espectroscópicos de la interacción de algunos de sus fragmentos con iones metálicos para determinar las zonas específicas de la proteína que unen estos ligandos. La interacción con Cu(II) se analizó por CD, UV, EPR, RMN e ITC. El Zn(II), por ser un ion diamagnético, sólo es posible estudiarlo por RMN e ITC.

II.6.5.1 Caracterización del péptido LEA61-69 – iones metálicos

En la Figura II.8A se muestra la titulación del péptido LEA61-69 con Cu(II) seguida por absorción electrónica: se observa la aparición de señales típicas de LMCT y transiciones d-d, que aumentan de intensidad con el aumento de la concentración de Cu(II). En el espectro de DC se observan al menos dos LMCT, una positiva y una negativa y dos bandas d-d. Las

LMCT concuerdan con las energías de transiciones de amidas desprotonadas e histidinas. La titulación se realizó hasta 3 equivalentes por ambas metodologías, y se observó saturación de la señal a 1 equivalente de Cu(II); por lo tanto la estequiometría del complejo Cu(II)-LEA6-69 es 1:1. De estos resultados podemos decir que este péptido tiene un sitio de unión a Cu(II) en el cual participan la H65 y amidas de la cadena peptídica. El espectro del péptido no reproduce el observado en AtLEA4-5 completa.

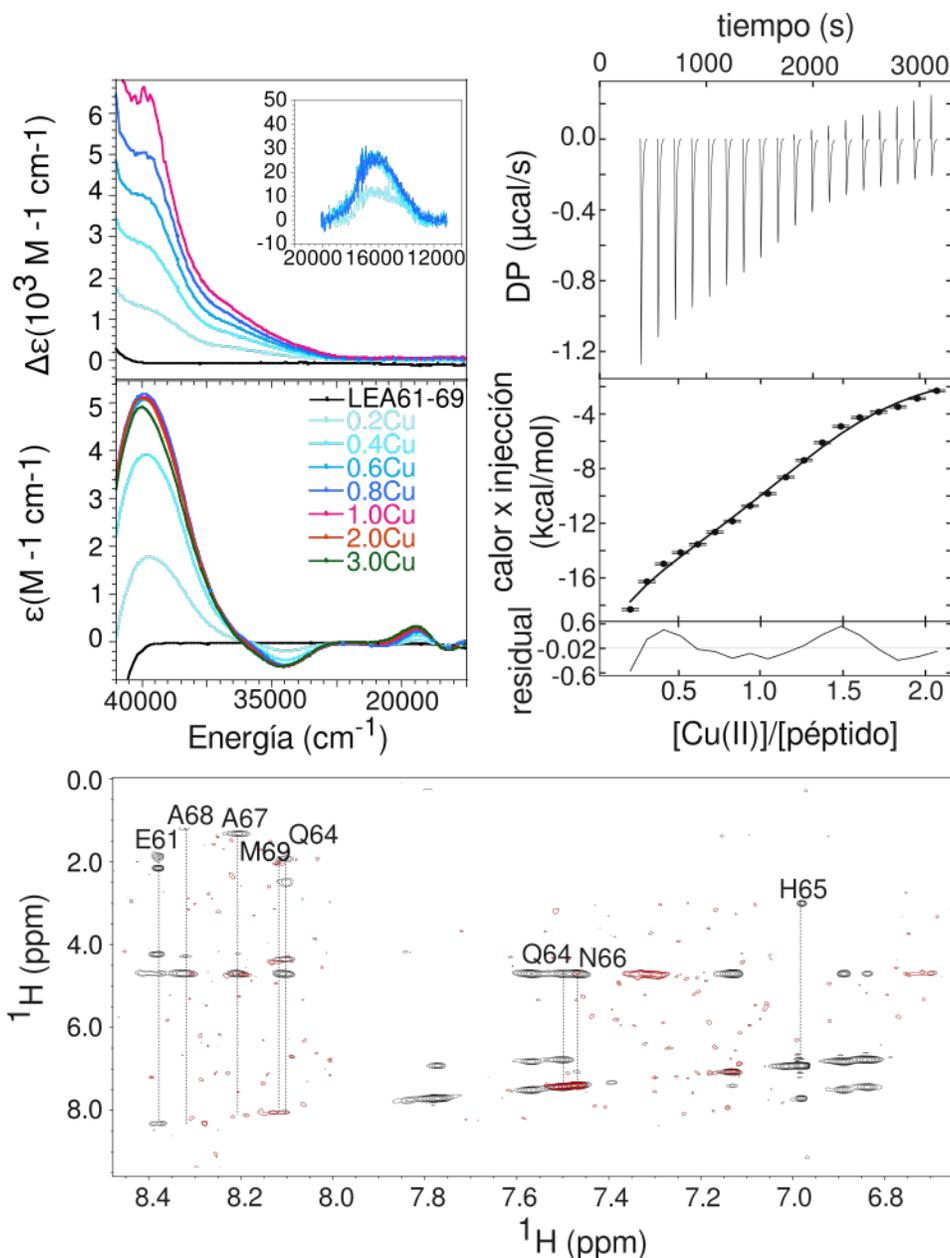


Figura II.8: A) Espectros de absorción electrónica en la región UV-vis del complejo péptido LEA61-69 - Cu(II), obtenidos por UV (panel superior) y DC (panel inferior). B) Ensayo de calorimetría. Se muestra el termograma (panel superior) y la integración de cada pico en función de la relación molar de la proteína y el Cu(II). C) Espectro ^1H - ^1H TOCSY del péptido (negro) y del complejo con Cu(II) en rojo.

Estos resultados están en concordancia con los obtenidos por RMN, donde se observó la pérdida y desplazamiento de algunos residuos, incluyendo la histidina (Figura II.8B). De acuerdo a las señales de RMN, en la coordinación del Cu(II) participan la H65 y los residuos ubicados hacia su extremo amino.

Por otra parte, los resultados de ITC indican que el péptido presenta un sitio de unión de moderada afinidad por Cu(II) (Figura II.8C).

Basado en los datos de DLS del complejo Cu(II) – AtLEA4-5, en los que el metal indujo la formación de oligómeros de mayor tamaño se estudió la agregación de los péptidos por turbidimetría, puesto que es difícil detectar los péptidos por DLS. En estos ensayos se incubaron los péptidos y sus complejos a 37 °C por 12 h y se determinó la absorción a 405 nm en función del tiempo utilizando un espectrofotómetro.

El péptido por si solo no muestra una agregación significativa en el tiempo del ensayo. Sin embargo, en presencia de los iones Zn(II), Ni(II) y Cu(II) aumenta su turbidez en el tiempo lo cual indica que su agregación está mediada por la participación de estos iones (Figura II.9). El orden en cuanto al incremento de la turbidez fue Cu(II) < Ni(II) << Zn(II).

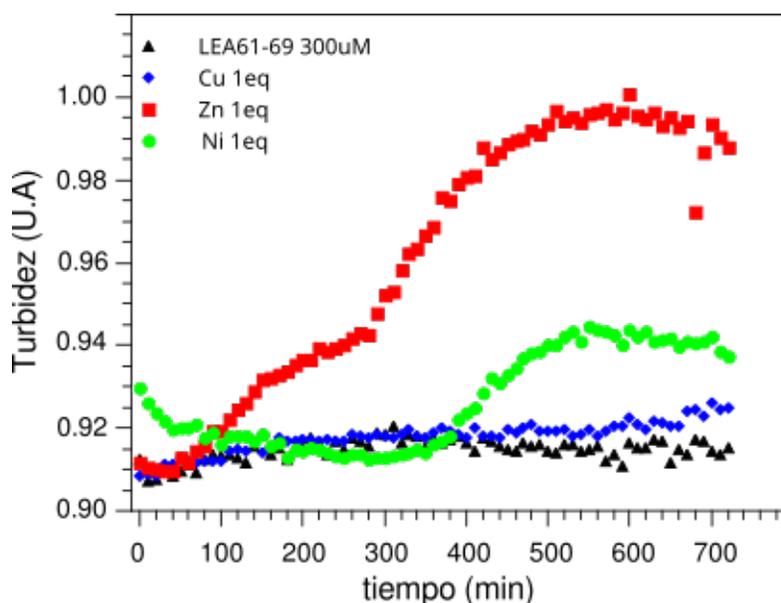


Figura II.9: Cinética de agregación del péptido LEA61-69. Se midió la absorción a 405 nm por 12 h a 37 °C, en presencia de los iones Cu(II), Zn(II) y Ni(II).

II.6.5.2 Caracterización de los péptidos LEA134-146 – iones metálicos

Para el estudio del fragmento 134-146 de AtLEA4-5, se sintetizaron los péptidos LEA134-146, LEA134-146H139A y LEA134-146H141A. En las Figuras II.10A, II.11A y II.12A se presentan las titulaciones de cada uno de los tres péptidos con Cu(II) seguidas por absorción electrónica y DC. Para el complejo Cu(II)-LEA134-146 se observaron tres bandas de transferencia de carga, una positiva y dos negativas, para el complejo Cu(II)-LEA134-146H139A dos LMCT, una positiva y una negativa y para el complejo Cu(II)-LEA134-146H141A tres LMCT, una positiva y dos negativas. Estos resultados sugieren que los tres péptidos coordinan Cu(II) con la participación de los residuos de histidina y amidas desprotonadas.

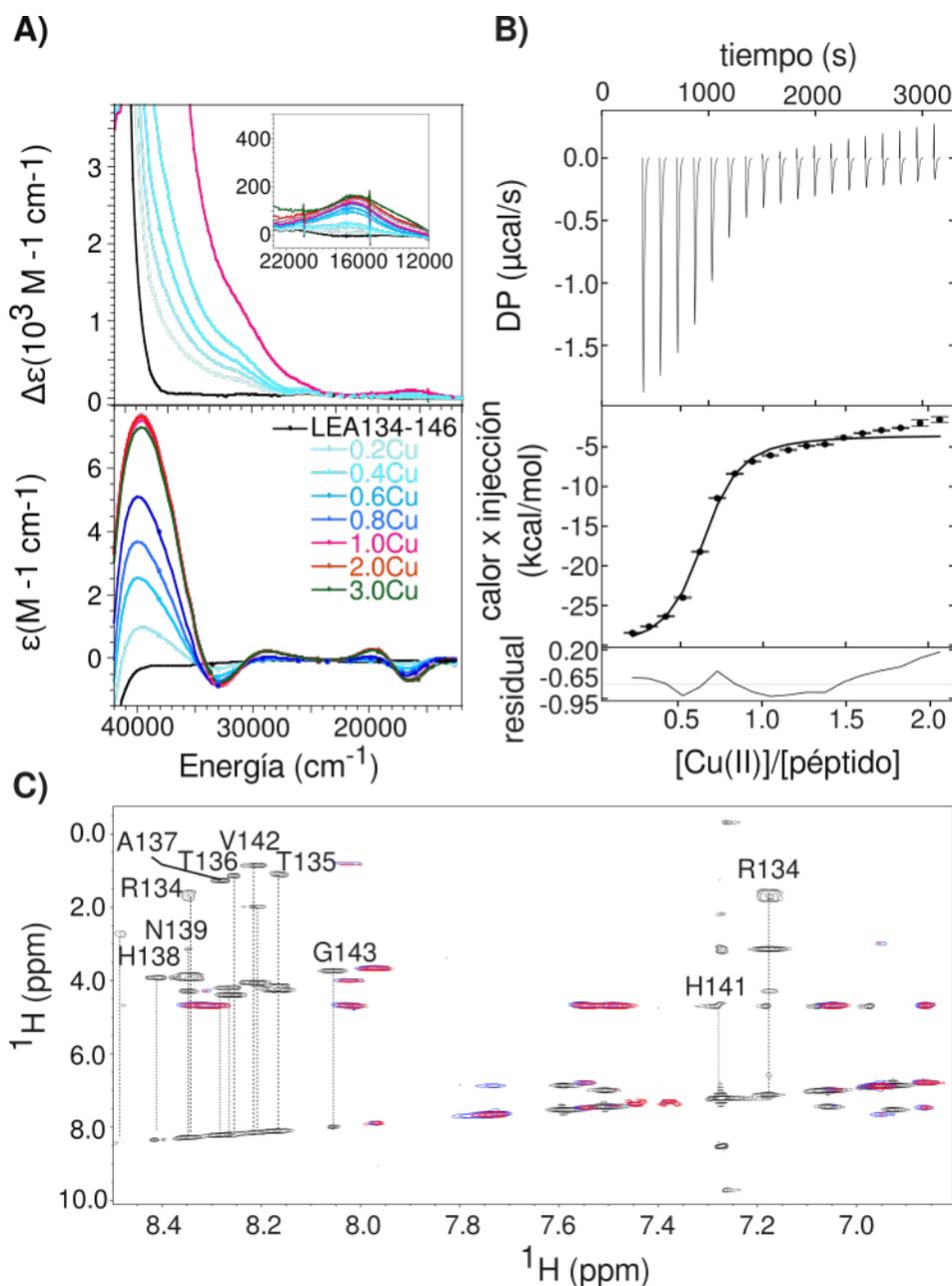


Figura II.10: (A) Espectros de absorción electrónica en la región UV-vis del complejo péptido LEA134-146- Cu(II), obtenidos por UV (panel superior) y DC (panel inferior). B) Ensayo de calorimetría. Se muestra el termograma (panel superior) y la integración de cada pico en función de la relación molar de la proteína y el Cu(II). C) Espectro TOCSY del péptido (negro) y del complejo con Cu(II) en rojo.

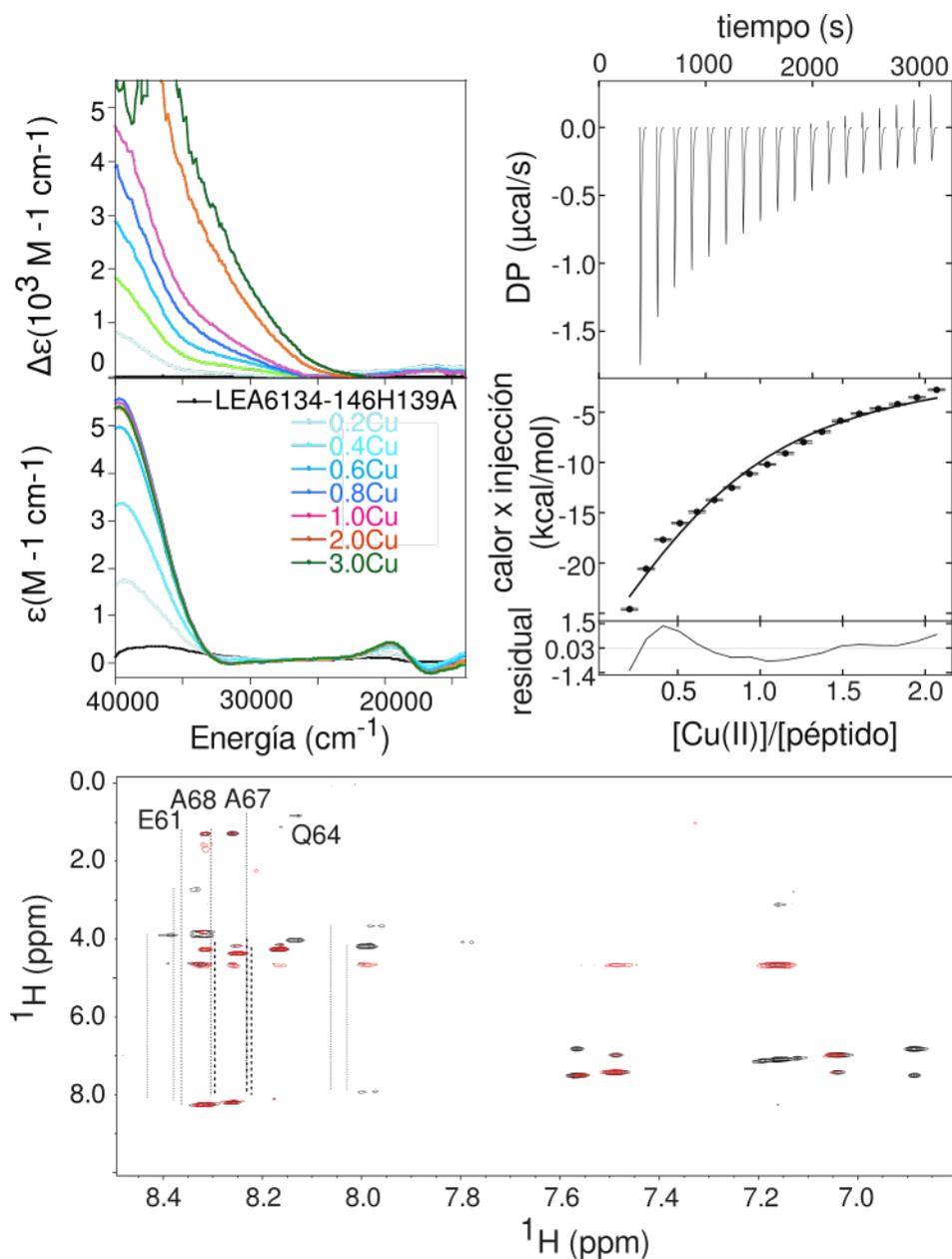


Figura II.11: (A) Espectros de absorción electrónica en la región UV-vis del complejo péptido LEA134-146H139A - Cu(II), obtenidos por UV (panel superior) y DC (panel inferior). B) Ensayo de calorimetría. Se muestra el termograma (panel superior) y la integración de cada pico en función de la relación molar de la proteína y el Cu(II). C) Espectro TOCSY del péptido (negro) y del complejo con Cu(II) en rojo.

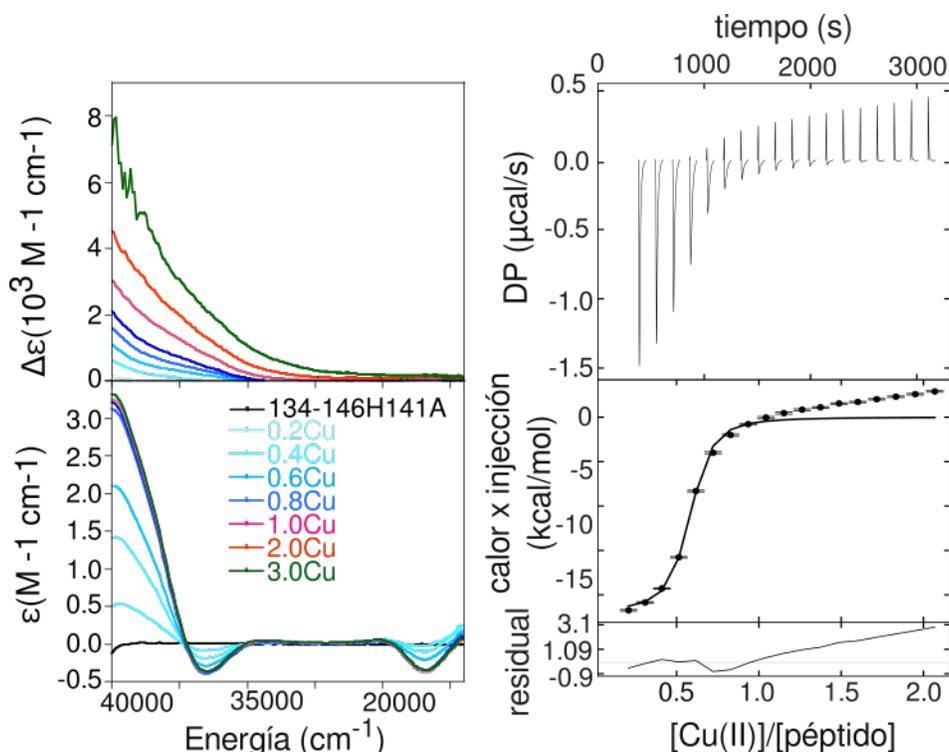


Figura II.12: A) Figura Espectros de absorción electrónica en la región UV-vis del complejo péptido LEA134-146H141A - Cu(II), obtenidos por UV (panel superior) y DC (panel inferior). B) Ensayo de calorimetría. Se muestra el termograma (panel superior) y la integración de cada pico en función de la relación molar de la proteína y el Cu(II).

Al analizar los espectros de CD del complejo Cu(II)-LEA134-146 se observa que a concentraciones mayores de 1 equivalente de Cu(II) las bandas d-d y la LMCT a 32000 cm^{-1} se corren a menor energía, y aparece una nueva señal alrededor de 28000 cm^{-1} , indicando un segundo sitio de coordinación. Al comparar los espectros del complejo Cu(II)-LEA134-146 hasta 1 eq con el espectro del complejo Cu(II)-LEA134-146H141A, se observa claramente la semejanza en los espectros, indicando que el primer sitio de unión a Cu(II) es la His141. De la misma manera, al comparar el espectro del complejo Cu(II)-LEA134-146 con 2eq con el espectro de Cu(II)-LEA134-146H139A se evidencia que el segundo sitio de unión es la His139. En los espectros del péptido LEA134-146H141A se obtienen estas señales, lo que evidencia que este sitio es el de mayor afinidad o el de una estructura más factible para la unión.

Los experimentos de RMN evidencian que el péptido LEA134-146 en presencia de Cu(II) sufre un cambio en la posición de la señal de la mayoría de sus residuos, lo cual indica que el ambiente químico del péptido es susceptible a la presencia de iones cobre en la solución (Figura II.10). Si el metal induce agregación, entonces la formación de un

dímero justificaría el cambio de los desplazamientos químicos de todos los residuos del péptido.

Se adquirieron espectros TOCSY en ausencia y presencia de equivalentes de Zn(II) para determinar cuáles eran las señales más afectadas por la presencia del ión. De acuerdo a los espectros, algunas señales de la cadena lateral sufrieron algún cambio, aunque no se observó una variación significativa en la posición de los picos en los NH.

La asignación del péptido LEA134-146H139A permitió discriminar entre los dos residuos de His del péptido LEA134-146. Lamentablemente, aún no se cuenta con la asignación del péptido LEA134-146H141A. Sin embargo, fue posible determinar, de acuerdo a las señales de el péptido LEA134-146H139A que coordina Cu(II) con la participación de los residuos ubicados hacia su carboxilo (Figura II.11).

En cuanto al estudio de interacción con Zn(II), no se observó variación alguna en los CS de las señales del péptido cuando se adicionó la solución metálica, lo que implicaría que este péptido tiene muy baja afinidad por Zn(II), o no interacciona con éste.

El péptido LEA134-146 en las condiciones utilizadas aumentó su turbidez en el tiempo. Su cinética de agregación se afectó en presencia de los iones Cu(II) y Zn(II). El mayor efecto se alcanzó con el ion Cu(II) (Figura II.13A). La adición de 1 eq de Cu(II) es suficiente para acelerar la agregación, mientras que el zinc requiere de 2 equivalentes para lograr ese efecto (Figura II.13B).

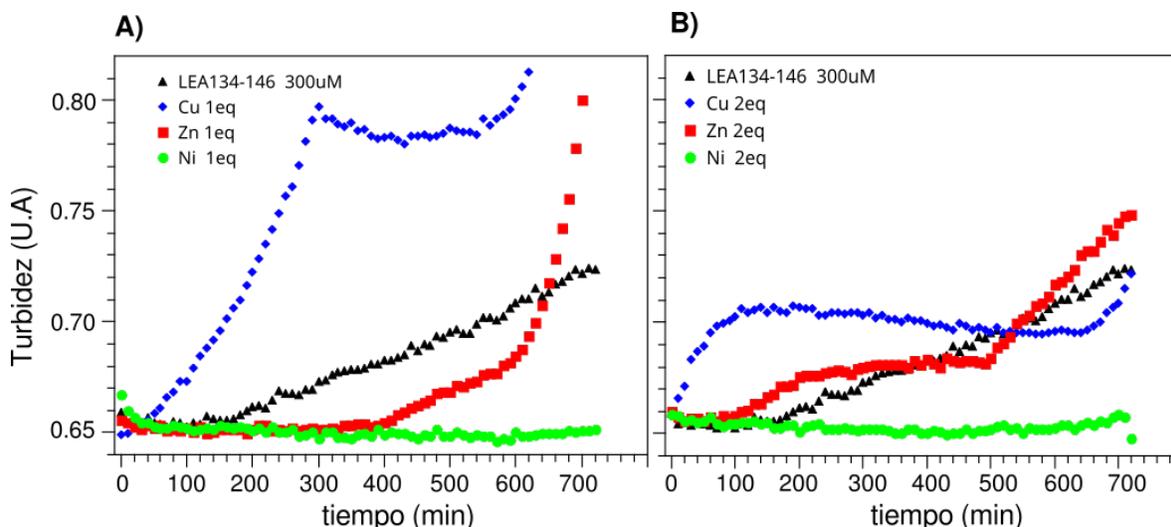


Figura II.13: Cinéticas de agregación del péptido LEA134-146 en presencia de 1 (A) o 2 (B) equivalentes de iones Cu(II), Zn(II) y Ni(II).

De las dos variantes del péptido, LEA134-146H139A agrega ligeramente, mientras el péptido LEA134-146H141A no lo hace (Figura II.14). Los iones Cu(II) no tienen efecto sobre la agregación del péptido LEA134-146H139A, mientras que acelera la agregación del péptido LEA134-146H141A. En cuanto al ion Zn(II), su efecto sobre la cinética de agregación es mayor sobre el péptido LEA134-146H139A.

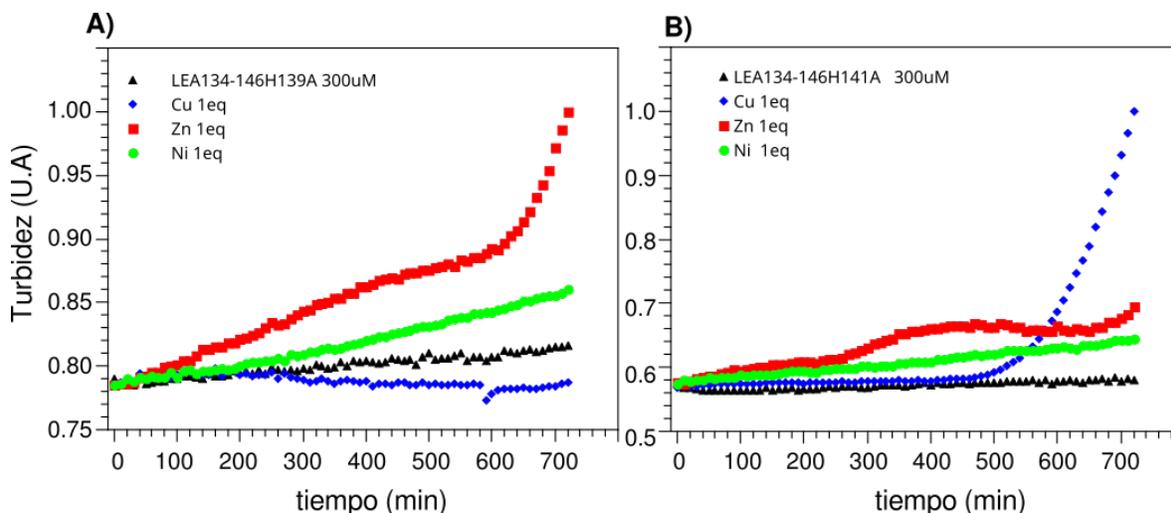


Figura II.14: Cinéticas de agregación de los péptidos LEA134-146H139A (A) y LEA134-146H141A (B) en presencia de 1 equivalente de iones Cu(II), Zn(II) y Ni(II).

Por el contrario, la adición de níquel sin importar si son 1 o 2 equivalentes, causa una inhibición de la agregación del péptido LEA134-146. Cada una de sus mutantes agregan en presencia de Ni(II). Aunque los iones níquel pueden interactuar con las mutantes, no lo

hace con el péptido que contiene ambas His, lo que sugiere que este fragmento pudiera no estar involucrado en la interacción en el contexto de la proteína.

Consideraciones finales

El estudio de la interacción de AtLEA4-5 con iones metálicos estuvo limitado a las técnicas disponibles y a la dificultad de la interpretación de los resultados. Nuestra hipótesis “Fragmentos peptídicos de AtLEA4-5 unidos a Cu (II) reproducen las características espectroscópicas del complejo proteína-Cu(II)” no pudo ser aceptada. Los espectros individuales UV-vis / DC de los péptidos estudiados, o la suma de éstos, no reproducen el de la proteína AtLEA4-5. Aún está pendiente el análisis de otros fragmentos y con ello de otras posibles combinaciones de espectros. Sin embargo, si la formación de los sitios de unión involucra a residuos muy alejados en secuencia o provenientes de diferentes monómeros, no es posible estudiarla con la metodología utilizada. Nuestro proyecto supuso que la coordinación del metal era local y que el estudio de fragmentos permitiría reproducir lo que sucede en la proteína. Esta estrategia ha sido empleada en otras IDP como por ejemplo las proteínas priónicas. No obstante, pudimos evidenciar que los péptidos unen los iones metálicos y que algunos son capaces de agregar por sí solos o lo hacen inducidos por el metal, lo cual está en concoordancia con el comportamiento de la proteína AtLEA4-5.

II.7 Conclusiones

1. La unión de Cu(II) a AtLEA4-5 no modifica su estructura secundaria, sin embargo favorece la formación de grandes oligómeros, en los que diferentes conformaciones de la proteína son reconocidas por los iones metálicos.
2. Los fragmentos peptídicos de AtLEA4-5 conservan su capacidad de unir los iones metálicos Cu(II), Zn(II) y Ni(II). Se unen a los iones metálicos con moderada afinidad (intervalo micromolar), facilitado por el componente entálpico principalmente.
3. El proceso de unión de la proteína AtLEA4-5 a Cu(II), Zn(II) y Ni(II) puede involucrar la coordinación mediada por residuos alejados en la secuencia, lo cual hizo poco eficiente la estrategia de estudiar los modos de coordinación de la proteína mediante péptidos.

II.8 Recomendaciones

- 1) Completar la caracterización de los péptidos.
- 2) Estudiar por RMN la proteína y determinar su paisaje conformacional.
- 3) Realizar estudios de funcionalidad de la proteína en cuanto a la prevención de daño por estrés oxidativo ocasionado por especies reactivas del oxígeno.

II.9 Bibliografía

- Battaglia, M., & Covarrubias, A. a. (2013). Late Embryogenesis Abundant (LEA) proteins in legumes. *Frontiers in Plant Science*, 4, 190.
- Battaglia, M., Olvera-Carrillo, Y., Garcarrubio, A., Campos, F., & Covarrubias, A. (2008). The enigmatic LEA proteins and other hydrophilins. *Plant Physiology*, 148(1), 6–24.
- Boucher, V., Buitink, J., Lin, X., Boudet, J., Hoekstra, F. a, Hundertmark, M., ... Leprince, O. (2010). MtPM25 is an atypical hydrophobic late embryogenesis-abundant protein that dissociates cold and desiccation-aggregated proteins. *Plant, Cell & Environment*, 33(3), 418–430.
- Colmenero-Flores, J. M., Moreno, L., Smith, C. E., & Covarrubias, A. (1999). Pvlea-18 , a Member of a New Late-Embryogenesis-Abundant Protein Family That Accumulates during Water. *Plant Physiology*, 120, 93–103.
- Culik, R. M., Abaskharon, R. M., Pazos, I. M., & Gai, F. (2014). Experimental Validation of the Role of Tri fl uoroethanol as a Nanocrowder. *The Journal of Physical Chemistry*, 118, 11455–11461.
- Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., & Bax, A. (1995). NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *Journal of Biomolecular NMR*, 6(3), 277–293.
- Dosztányi, Z., Csizmok, V., Tompa, P., & Simon, I. (2005). IUPred: Web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics*, 21(16), 3433–3434.
- Dure, L. (1993). *Plant Responses to Cellular Dehydration during Environmental Stress*. (T. J. Close & E. A. Bray, Eds.). Rockville: American Society of Plant Physiologists.
- Dure, L., & Galau, G. A. (1981). Developmental biochemistry of cotton seed embryogenesis and germination : XIII. Regulation of biosynthesis of principal storage proteins. *Plant Physiology*, 68(1), 187–194.
- French-Pacheco, L. (2015). *Caracterización espectroscópica de la interacción con metales de la familia de proteínas AtLEA4 (Late Embryogenesis Abundant) de Arabidopsis thaliana (Brassicales: Brassicaceae)*. Universidad Autónoma del Estado de Morelos.1-68.
- Garay-Arroyo, A., Colmenero-Flores, José M. Garcarrubio, A., & Covarrubias, A. A. (2000). Highly Hydrophilic Proteins in Prokaryotes and Eukaryotes Are Common during Conditions of Water Deficit. *Journal of Biological Chemistry*, 275(8), 5668–5674.
- Gusman, H., Lendenmann, U., Grogan, J., Troxler, R. F., & Oppenheim, F. G. (2001). Is salivary histatin 5 a metallopeptide? *Biochimica et Biophysica Acta - Protein Structure and Molecular Enzymology*, 1545(1–2), 86–95.
- Hara, M., Fujinaga, M., & Kuboi, T. (2004). Radical scavenging activity and oxidative modification of citrus dehydrin. *Plant Physiology and Biochemistry*, 42(7–8), 657–662.
- Hara, M., Kondo, M., & Kato, T. (2013). A KS-type dehydrin and its related domains reduce Cu-promoted radical generation and the histidine residues contribute to the radical-reducing activities. *Journal of Experimental Botany*, 64(6), 1615–1624.

- Holehouse, A. S., Das, R. K., Ahad, J. N., Richardson, M. O. G., & Pappu, R. V. (2017). CIDER: Resources to Analyze Sequence-Ensemble Relationships of Intrinsically Disordered Proteins. *Biophysical Journal*, 112(1), 16–21.
- Hundertmark, M., & Hinch, D. K. (2008). LEA (late embryogenesis abundant) proteins and their encoding genes in *Arabidopsis thaliana*. *BMC Genomics*, 9, 118.
- Keller, R. (2004). The computer aided resonance assignment tutorial. In *Goldau, Switzerland: Cantina Verlag*.
- Kovacs, D., Szabo, B., Pancsa, R., & Tompa, P. (2013). Intrinsically disordered proteins undergo and assist folding transitions in the proteome. *Archives of Biochemistry and Biophysics*, 531(1–2), 80–89.
- Kovacs, Denes, Agoston, B., & Tompa, P. (2008). Disordered plant LEA proteins as molecular chaperones. *Plant Signaling & Behavior*, 3(9), 710–713.
- Krüger, C., Berkowitz, O., Stephan, U. W., & Hell, R. (2002). A metal-binding member of the late embryogenesis abundant protein family transports iron in the phloem of *Ricinus communis* L. *Journal of Biological Chemistry*, 277(28), 25062–25069.
- Liu, G., Xu, H., Zhang, L., & Zheng, Y. (2011). Fe binding properties of two soybean (*Glycine max* L.) LEA4 proteins associated with antioxidant activity. *Plant & Cell Physiology*, 52(6), 994–1002.
- Liu, Y., Wang, L., Xing, X., Sun, L., Pan, J., Kong, X., ... Li, D. (2013). ZmLEA3, a multifunctional group 3 LEA protein from maize (*Zea mays* L.), is involved in biotic and abiotic stresses. *Plant & Cell Physiology*, 54(6), 944–959.
- Lupas, A., Van Dyke, M., & Stock, J. (1991). Predicting coiled coils from protein sequences. *Science*, 24(252), 1162–1164.
- Olvera-Carrillo, Y., Campos, F., Reyes, J. L., Garcíarrubio, A., & Covarrubias, A. a. (2010). Functional analysis of the group 4 late embryogenesis abundant proteins reveals their relevance in the adaptive response during water deficit in *Arabidopsis*. *Plant Physiology*, 154(1), 373–390.
- Xue, B., Dunbrack, R., Williams, R., Duncker, K., & Uversky, V. (2010). PONDR-FIT: A Meta-Predictor of Intrinsically Disordered Amino Acids. *Biochim Biophys Acta*, 1804(4), 996-1010.

CAPÍTULO III

Interacción LAOBP -aminoácidos

Capítulo III

III. Interacción LAOBP – aminoácidos

III.1 Proteína Lysine-Arginine-Ornithine Binding Protein (LAOBP), una proteína transportadora de aminoácidos.

III.1.1 Características estructurales de PBPs

Las proteínas de unión periplásmicas bacterianas (PBPs) han sido utilizadas como modelo para estudiar los principios moleculares de la especificidad proteína-ligando, ya que reconocen una gama de moléculas pequeñas a pesar de que sus complejos comparten una estructura similar (Pistolesi et al., 2011)□.

Las PBPs actúan como quimiorreceptores primarios y como los receptores iniciales para el transporte mediado por el sistema ABC (*ATP-binding cassette*), (Holland et al., 2003)□. Las PBPs están implicadas en la transportación de una amplia variedad de sustratos (Higgins & Ames, 1981)□. Entre las PBPs se han descrito la proteína de unión a maltosa (MBP), la proteína de unión a galactosa/glucosa (GBP), la proteína de unión a ribosa (RBP), y la proteína de unión a dipéptidos (DBP) (Oh et al., 1994)□. El complejo proteína-ligando, además de interactuar con las proteínas transportadoras, interactúa con un receptor quimiotáctico que transmite las señales de movimiento flagelar. Las PBPs tienen entre 25-60 kDa y pueden unir diferentes ligandos con una alta afinidad (0.01–1 μ M).

Los sustratos, luego de atravesar la membrana externa de las bacterias son reconocidos por las PBPs con muy alta afinidad. La unión se caracteriza por un movimiento de cuerpo rígido de los dominios globulares que las conforman (Oh et al., 1993; Sharff et al., 1992)□. La forma unida de las PBPs reconoce e interactúa con un complejo de transporte unido a la membrana interna ABC, lo que desencadena la liberación del sustrato de la PBPs y su translocación al citosol celular. Asociado a este proceso, el complejo ABC hidroliza ATP (Ames & Mimura, 1992)□.

Aunque la familia PBPs tiene una amplia diversidad de tamaños, reconoce una gama de sustratos y carece de una alta identidad de secuencia, las estructuras cristalográficas revelan una arquitectura común para todos sus miembros: dos dominios o lóbulos globulares que

comparten un tamaño y una topología similares, cada uno formado por hebras β rodeadas de hélices α .

En ausencia de sustrato, los dos dominios de una PBP están típicamente lejos uno del otro, separados por una hendidura o ranura profunda accesible al solvente (Figura III.1). Un cambio conformacional permite la transformación de una conformación "abierta" a una "cerrada" (Pistolesi et al., 2011)□. En esta última conformación, observada en las estructuras de los complejos, el sustrato ocupa la hendidura y los dos dominios se acercan y establecen múltiples contactos con el sustrato, el que se encuentra protegido de interactuar con el disolvente. Debido a que la estructura de los dominios individuales permanece en gran medida sin cambios en ambas conformaciones (apo y complejo), y solo se observan diferencias en su posición y orientación relativas, la transición conformacional ha sido descrita como un movimiento de cuerpo rígido donde los segmentos que unen los dos dominios funcionan como una bisagra. Se han encontrado PBPs que existen en una sola conformación en la forma libre, mientras que otras han sido cristalizadas en diferentes conformaciones.

Dado que el cambio conformacional entre ambos estados representa un mecanismo de control esencial para los procesos de transporte, la caracterización de las estructuras de la forma apo y unida a ligando ha sido objeto de muchas investigaciones.

III.1.2 Estudios de caracterización de LAOBP

Una de las PBPs que ha sido estudiada es la proteína de unión a lysina, arginina, y ornitina (LAOBP). Esta es uno de los receptores del sistema de histidina permeasa (Ames & Nikaido, 1978; Kerppola et al., 1991)□. La histidina permeasa está compuesta por las proteínas de unión periplasmáticas LAOBP, HisJBP (proteína de unión a histidina – HisJ), por un complejo unido a la membrana citoplasmática que incluye dos proteínas transmembranales hidrofóbicas: HisQ y HisM, y un transportador ABC hidrofílico adosado a la membrana por el lado citosólico HisP (Ames & Nikaido, 1978; Kustu & Ames, 1973)□.

Las proteínas HisJBP y LAOBP tienen una masa molecular de ~ 26 kDa, comparten un 70% de identidad de secuencia y tienen estructuras terciarias muy similares (Kang et al., 1991)□. Ambas proteínas son capaces de unir y transportar L-histidina, L-arginina, L-lisina y L-ornitina, pero con diferentes afinidades (Higgins & Ames, 1981)□. LAOBP tiene mayor afinidad por la L-arginina (rango nM) y menor por His (μM) (Chu et al., 2013)□, mientras que HisJBP tiene mayor afinidad por histidina (rango nM).

Particularmente, LAOBP de *Salmonella typhimurium* tiene dos dominios globulares, cada uno de los cuales consta de cinco hebras β y cuatro hélices α, conectadas por dos segmentos cortos (Kang et al., 1991)□. El sitio de unión está ubicado en la interfaz entre los dos dominios (Kang et al., 1991; Yang & Mancera, 2008)□. El dominio 1 es discontinuo (residuos 1-88 y 195-238) mientras que el dominio 2 es continuo (residuos 93-185) (Figura III.1).

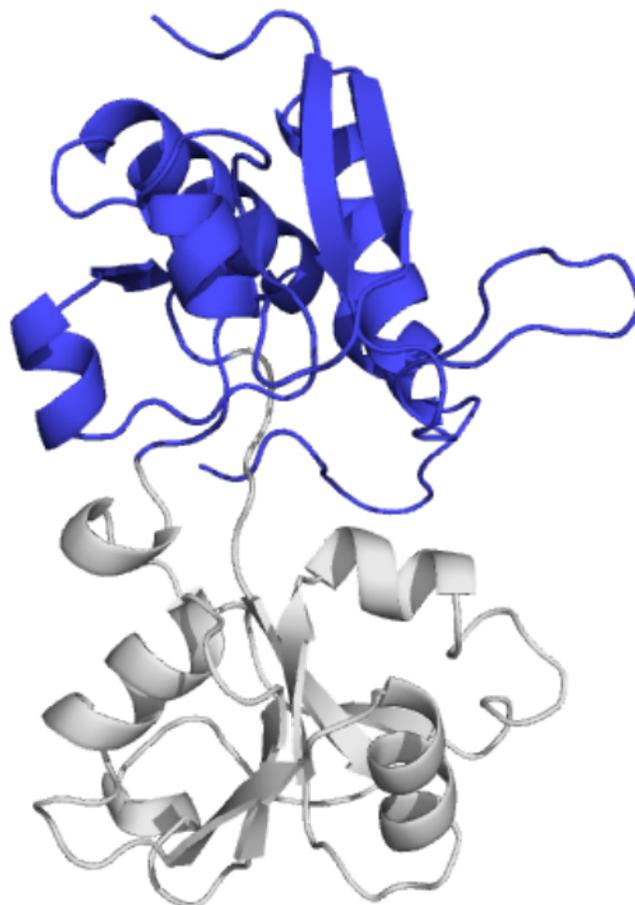


Figura III.1: Estructura cristalográfica de LAOBP (#PDB: 2LAO). En gris se representa el dominio o lóbulo I, y en azul el dominio II.

La estructura cristalina de LAOBP libre muestra una conformación abierta (Kang et al., 1991) □ mientras que todos los complejos cristalizados mostraron una conformación cerrada (lisina (Oh et al., 1993) □, arginina (Oh et al., 1994) □, ornitina (Oh et al., 1994) □ e histidina (Oh et al., 1994) □). En la conformación cerrada, los dos lóbulos giran en relación con la conformación abierta ($\sim 50^\circ$) (Kreimer et al., 2000; Silva et al., 2011) □. No está claro si los ligandos inducen el cambio conformacional directamente (Pulido et al., 2015) □ o simplemente estabilizan la conformación cerrada (Yang & Mancera, 2008) □. Sin embargo, la unión debe ocurrir para que la PBP interactúe con el complejo de membrana (Higgins & Ames, 1981) □. La movilidad intramolecular de la forma unida se ha correlacionado directamente con su capacidad para inducir a la ATPasa y se ha propuesto que los dos dominios HisJBP deben orientarse uno con respecto al otro en una posición única para inducir a la ATPasa (Kreimer et al., 2000) □.

Recientemente, se modificó la especificidad de LAOBP con el diseño de una proteína mutada (LAOBP_L117K) (Banda-Vázquez et al., 2018)□. La mutante tuvo una afinidad por Gln de 1.6 mM, y aunque aún reconocía aminoácidos básicos, lo hacía con mucha menor afinidad. Resultó interesante que la introducción de esta mutación se realizó en un sitio no muy conservado y que estaba alejado del bolsillo primario de interacción con los ligandos fisiológicos.

Aunque las simulaciones de dinámica molecular han proporcionado información sobre los cambios conformacionales presentes en las PBP (Keskin et al., 2000; Silva et al., 2011; Yang & Mancera, 2008)□, aún no se ha descrito experimentalmente la dinámica de la proteína en su forma libre o de sus complejos. Realizar estudios de dinámica, como los documentados en la caracterización de HisJBP por RMN (Chu et al., 2014; Chu et al., 2013; Igarashi et al., 2010)□ proporcionaría evidencia experimental para comprender la diferencia en especificidad entre las PBPs. Tampoco se han publicado estudios experimentales que describan la vía de los cambios conformacionales de la proteína con resolución atómica.

III.2 Justificación del proyecto: interacción LAOBP - aminoácidos

La proteína periplasmática LAOBP de *Sallmonella* transporta aminoácidos hacia la membrana interna donde los libera al interactuar con un receptor transmembranal específico. No se ha elucidado cómo la proteína pasa de su conformación abierta a la cerrada en presencia de los aminoácidos y cómo en su interacción con las proteínas de la membrana citoplasmática libera a los residuos.

La especificidad dependiente del ligando no se puede explicar completamente por la diferencia estructural debido a que el RMSD entre cada estructura unida a ligando es menor que 0.18 Å. Además, los estudios mediante cristalografía de rayos X no han proporcionado información sobre el comportamiento dinámico de la proteína en solución o el cambio conformacional entre los estados abierto y cerrado. Mediante simulaciones de dinámica molecular se ha logrado ampliar la caracterización de los cambios conformacionales que sufre la proteína para transitar de un estado abierto a uno cerrado tras la unión del ligando. Sin embargo, realizar estudios de caracterización mediante RMN proporcionaría evidencia experimental para comprender esta transición. Aquí presentamos estudios de interacción y conformacionales de LAOBP con arginina por RMN.

III.3 Hipótesis

La unión de arginina a LAOBP afecta los desplazamientos químicos de varias regiones de la proteína además de las directamente relacionadas con la unión de ligando.

III.4 Objetivos

Objetivo general:

Estudiar la interacción de LAOBP con arginina mediante RMN.

Objetivos particulares:

1. Asignar los desplazamientos químicos de los átomos de LAOBP.
2. Determinar las zonas de interacción de LAOBP con arginina mediante RMN.
3. Estudiar cambios conformacionales de LAOBP mediante RMN.

III.5 Métodos

III.5.1 Expresión de la proteína recombinante LAOBP

La expresión heteróloga se realizó a partir del plásmido p23A que contiene el gen que codifica para LAOBP. El vector fue donado por el Dr. Alejandro Sosa de la Facultad de Medicina de la UNAM. La cepa bacteriana *Escherichia coli* BL21-Rosseta DE3 se utilizó como sistema de expresión de LAOBP. Se creció una colonia en 20 mL de medio LB (Luria Bertani) (NaCl:peptona:extracto de levadura 1:1:0.5%) toda la noche a una temperatura de 37°C (precultivo). Se adicionó el antibiótico ampicilina (0.1 mg/mL). El precultivo se adicionó a 500 mL de medio LB con antibiótico y se dejó crecer hasta que alcanzó una absorbancia de 0.8-1 a 600 nm. La sobreexpresión de la proteína se indujo con IPTG 0.3 mM durante toda la noche a 25 °C. Para los estudios de RMN se utilizaron 10 mL del precultivo y la inducción se realizó en 1 L de medio mínimo M9 (Na₂HPO₄·7H₂O, KH₂PO₄, NaCl, ¹⁵NH₄Cl, glucosa ¹³C, MgSO₄, CaCl₂, MnCl₂, FeCl₃, ZnSO₄ y vitaminas), para obtener las proteínas isotópicamente marcadas.

III.5.2 Preparación de la muestra y purificación de LAOBP mediante cromatografía de intercambio aniónico

Posterior a la inducción de la sobreexpresión de las proteínas, los cultivos se centrifugaron a 2.000 x g a 4 °C por 30 minutos. El sedimento de los cultivos con LAOBP se resuspendió en una solución de acetato de potasio (KOAc) 10 mM pH 5.1, sacarosa 20% y EDTA 1 mM (tampón de equilibrio) el cual contenía inhibidores de proteasas. La fracción periplásmica se recogió mediante choque osmótico utilizando KOAc 10 mM pH 5.1. Nuevamente, se realizó una centrifugación 2.000 x g a 4 °C por 30 minutos y se recuperó el sobrenadante. El lisado se filtró a través de un filtro GVS (tamaño de poro de 0.22 µm) para eliminar los residuos celulares insolubles. Posteriormente, el lisado se dializó varias veces contra cloruro de guanidinio (GdnHCl) 2 M y bis-tris propano 5 mM pH 8.5, de modo que se eliminasen posibles aminoácidos unidos. La muestra se concentró con una unidad de filtro centrífugo Amicon (10 kDa) y se intercambió en un tampón de unión de bis-tris propano 5 mM pH 8.5.

La única etapa de purificación fue una cromatografía de intercambio aniónico utilizando una columna Source Q. LAOBP se eluyó usando un gradiente lineal (NaCl 0–250 mM). La

columna se conservó a 4 °C en etanol 20%. La presencia de la proteína se siguió a lo largo de la purificación mediante electroforesis en geles de poliacrilamida en condiciones desnaturizantes.

III.5.3 Asignación de los desplazamientos químicos de los átomos de LAOBP y de su complejo con arginina.

Los espectros de RMN se adquirieron a 298 K usando un campo magnético de 16.4 T que se corresponden con la frecuencia de ^1H de 700 Mhz, en un equipo del Laboratorio Nacional de Estructuras de Macromoleculas (LANEM). Para la asignación de la cadena principal de las proteínas se adquirieron espectros del tipo HNCO, HNCACB, CBCA(CO)NH (Grzesiek & Bax 1992), ^1H - ^{15}N HSQC (Palmer et al. 1991), mientras que la asignación de la cadena lateral se obtuvo usando los experimentos H(CCO)NH y C(CO)NH (Montelione et al. 1992).

Los experimentos se analizaron utilizando NMRpipe (Delaglio et al., 1995) □ y CARAM (Keller, 2004) □ para obtener la asignación de los espectros.

III.5.4 Determinación de las zonas de interacción de LAOBP con ligandos mediante RMN.

La determinación de la perturbación o cambios en la posición de los desplazamientos químicos (CSP o ΔCS , de sus siglas en inglés) de los residuos de LAOBP tras la adición de un ligando (Arg o Ala) permite identificar cambios en el ambiente químico de los posibles residuos involucrados en la interacción. La ΔCS de cada residuo se obtuvo realizando una comparación de espectros tipo ^1H - ^{15}N HSQC en ausencia y presencia del ligando. Se comparó la posición de los átomos usando la siguiente ecuación:

$$\Delta\text{CS} = |\Delta\delta\text{HN}| + |\Delta\delta\text{N}|/5$$

Los cambios fueron mapeados en la estructura tridimensional de las proteínas.

III.5.6 Estudio del paisaje conformacional de LAOBP mediante RMN.

Para explorar los estados conformacionales de LAOBP se realizaron experimentos del tipo ^1H - ^{15}N HSQC a diferentes temperaturas (25, 30 y 37 °C). Los datos fueron analizados en NMRPipe y NMRView.

III.6 Resultados

III.6.1 Asignación de los desplazamientos químicos de los átomos de LAOBP

Los desplazamientos químicos de los átomos de la cadena principal de la proteína LAOBP y de su complejo (LAOBP-Arg) se asignaron secuencialmente en un 92 % y 90 % respectivamente (Figura III.2), mientras que solo se logró asignar el 59 % de las cadenas laterales. Básicamente no se asignaron los siguientes aminoácidos: A1, L71, D75, R77, S92, G99, L106, G111, T121, T132, G134, A141, Q143, E205, D232, N234 y los átomos de nitrógeno amida de las 7 prolinas de la secuencia de LAOBP (Figura III.2). Las resonancias asignadas de LAOBP y LAOBP-Arg se depositaron en la base de datos *Biological Magnetic Resonance Data Bank* (BMRB) con los números de acceso 27318 y 27319 respectivamente.

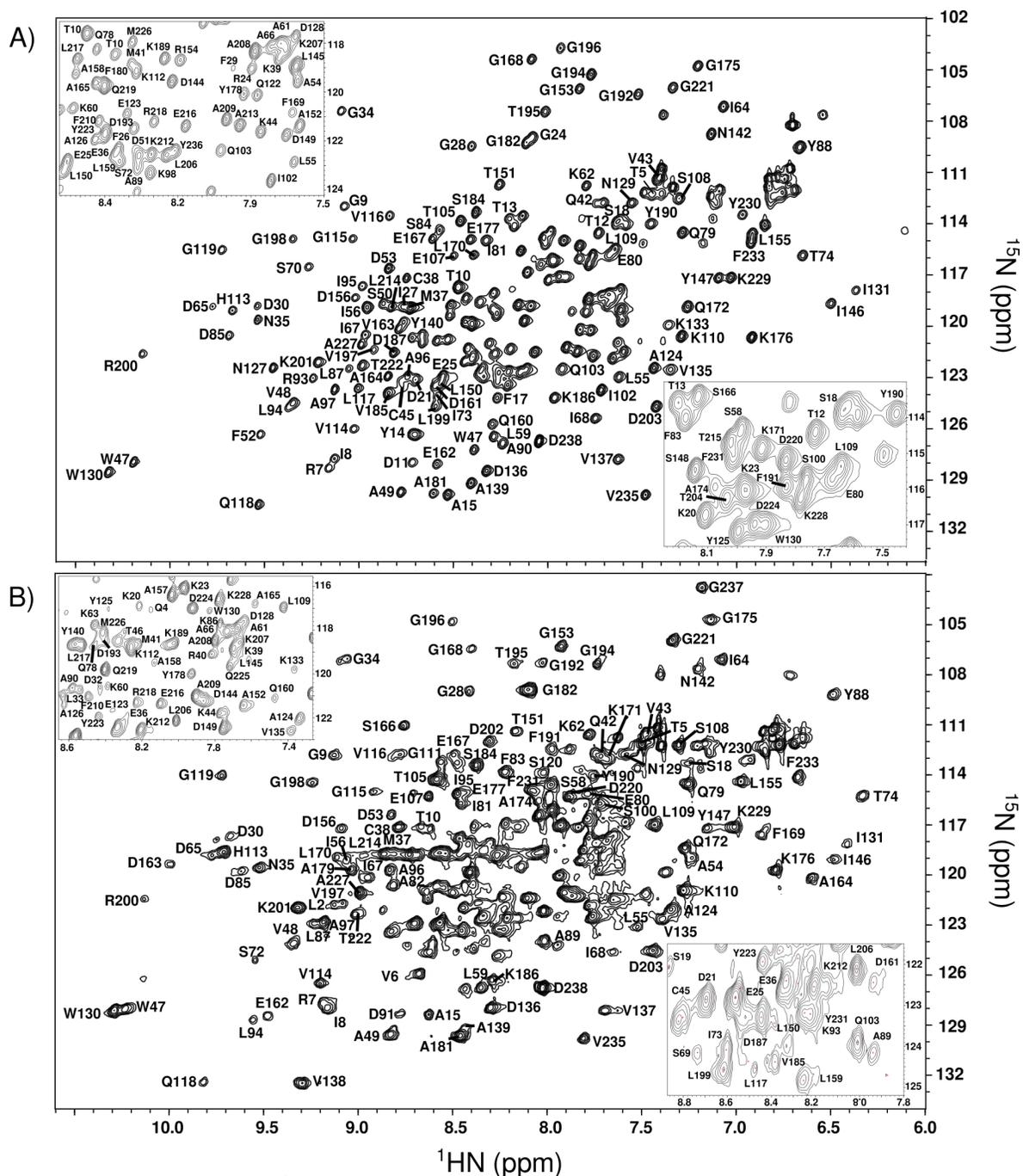


Figura III.2: Espectros ^1H - ^{15}N HSQC de LAOBP (A) y del complejo LAOBP-Arg (B) a pH 7,5, 303 K y 700 MHz. Se indican las asignaciones de los residuos de LAOBP.

III.6.2 Determinación de residuos involucrados en la interacción LAOBP y arginina

La comparación de los cambios en los desplazamientos químicos entre LAOBP y LAOBP-Arg permitió identificar los residuos que probablemente están involucrados en las

interacciones (Figura III.3). Al mapear la diferencia de los desplazamientos químicos (Δ CS) en la estructura terciaria, se pudo observar que la unión del ligando a la proteína produce una perturbación en los CS de residuos localizados en la cavidad entre los dos dominios y en otros aminoácidos que no interactúan directamente con el ligando (Figura III.3).

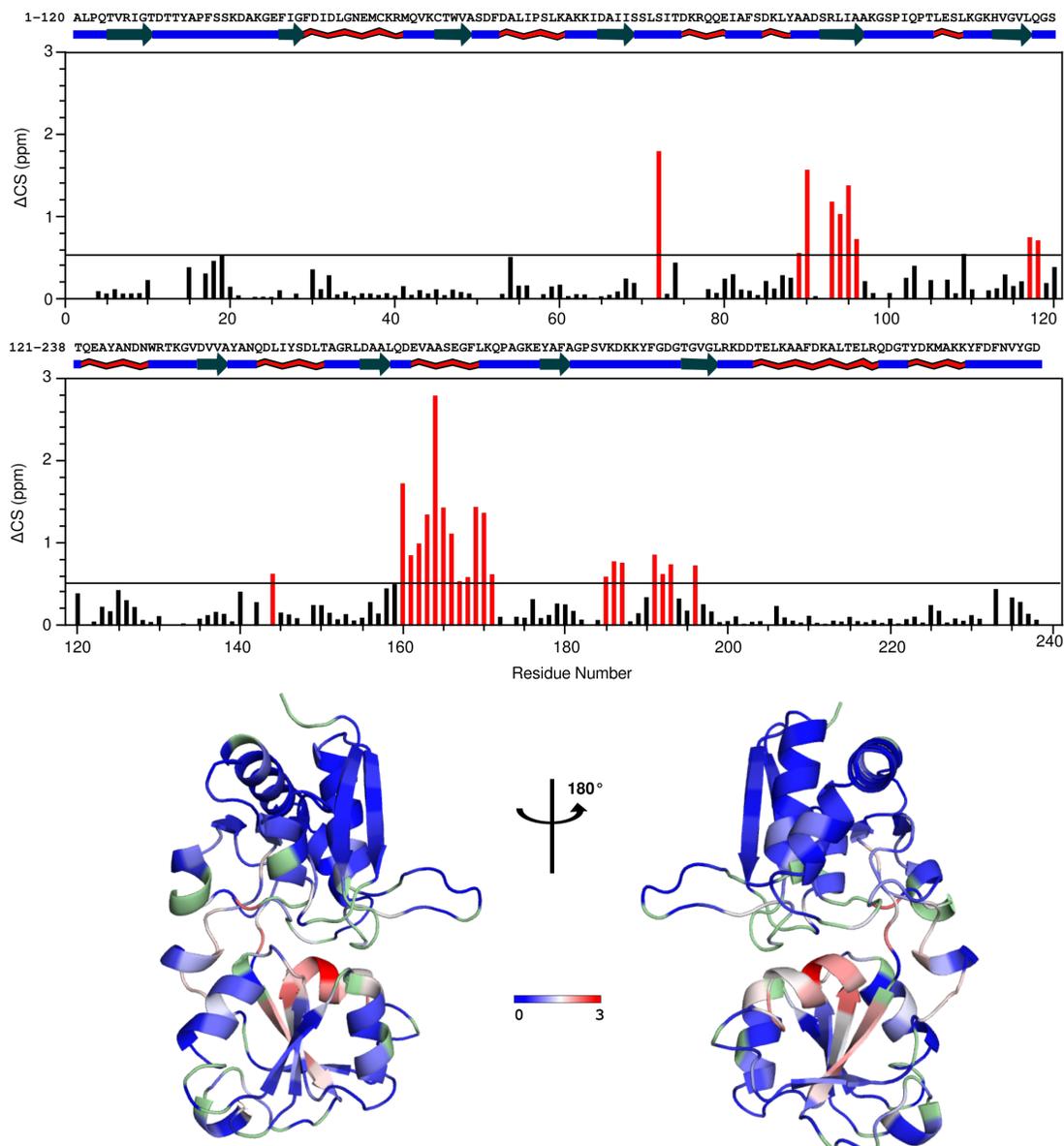


Figura II.3: A) Diferencia de los desplazamientos químicos (Δ CS) entre el espectro de LAOBP y del complejo LAOBP-Arg. La línea negra representa el valor de corte utilizado (media de la población más 0.5 desviación estándar). (B) Estructura tridimensional del complejo LAOBP-Arg. En un gradiente de blanco a rojo (de menor a mayor, respectivamente) se muestran aminoácidos con un Δ CS mayor que el valor de corte. De azul a blanco (de menor a mayor, respectivamente) son aquellos aminoácidos que tuvieron Δ CS menor que el valor de corte. Los residuos sin asignación se muestran en verde. (PDB: 1LAF).

Los residuos S18, S72, L117, D161 son algunos de los CS afectados por la presencia de arginina. Estos son parte del sitio de unión a Arg, de acuerdo con la estructura cristalográfica del complejo LAOBP (S18, F52, S120, Y190 y D193, Kang et al., 1991) y (D11, Y14, D30, F52, S69, S70, S72, R77, L117, T121, D161, Oh et al., 1994).

Otros residuos afectados fueron F52, A90, D91, S92, R93, L94, F191, G192 y D193. La F52 ha sido descrito como clave en la determinación de las afinidades relativas de las proteínas LAOBP e HisJBP por diferentes ligandos (Oh et al 1994). La unión del sustrato a LAOBP se logra mediante un movimiento de cuerpo rígido de sus lóbulos los cuales forman un único ángulo de torsión alrededor de la A90 (Oh et al 1993). También se han descrito algunos cambios en el ángulo de torsión para la Y190 y el D193 (Oh et al 1993), mientras que se ha sugerido que el fragmento 190-194 constituye un segmento de comunicación interdominio, (Keskin et al., 2000). Los residuos A90 y E167, entre otros, corresponden a los residuos que recubren la hendidura (Keskin et al., 2000).

La mayoría de los residuos que no son del sitio de unión y se afectan por la unión del ligando son similares a los afectados en el complejo de HisJBP (Igarashi et al 2010). Algunos de los residuos cuyo CS se vio afectado (160-171) podrían correlacionar con residuos que participan en la interacción con la membrana. Esta hipótesis se basa en que estos aminoácidos están dentro de dos regiones altamente conservadas y expuestas al solvente (residuos 53-94 ; 146-173) (Kang, 1991).

La mayoría de los CSP están localizados en el lóbulo más grande (I), en concordancia con los datos cristalográficos a partir de los cuales se sugirió que el sustrato se une con preferencia a este lóbulo (Oh et al., 1993). Además, estudios de los dominios individuales de HisJBP mostraron que el lóbulo I de HisJBP se une a His mientras que el lóbulo II no lo puede hacer, por lo que se sugirió que el lóbulo I participa en los primeros pasos de la unión (Chu et al., 2013). No obstante, observamos CSP en ambos dominios, lo que ha sido descrito anteriormente. Wolf reportó que algunos residuos en cada dominio experimentan cambios significativos cuando tiene lugar la rotación relativa del lóbulo (Wolf et al., 1996).

Otro factor importante es el tipo de sustrato. La movilidad intramolecular de los complejos está directamente correlacionada con su capacidad para inducir la actividad

ATPasa en el receptor de membrana (Kreimer et al., 2000). Las proteínas LAOBP e HisJBP se unen al mismo sustrato pero con diferente afinidad (Kang, 1991), por lo que se ha sugerido que algunos de los residuos de su hendidura son diferentes, como los residuos de las posiciones 18, 52, 120, 190 y 193. En concordancia con eso, detectamos algunos cambios en los residuos 18, 120 y 193, mientras que el ambiente químico de los residuos correspondientes no se ven afectados en HisJBP en complejo (Chu et al., 2014). Estas diferencias en el centro de unión de las dos proteínas se deben a interacciones polares entre la proteína y la cadena lateral del ligando (Pulido et al., 2015).

III.6.3 Estudio de las conformaciones de LAOBP

Aunque Oh et al. 1993 propusieron que LAO existe en un equilibrio dinámico entre la conformación abierta y cerrada y que el sustrato estabiliza esta última, estudios posteriores sugirieron que el estado apo de LAOBP tiene principalmente una conformación abierta (Silva et al., 2011, Pulido et al., 2015), siendo capaz de sufrir el cambio conocido como “Venus flytrap”. En contraste, los estudios de RMN y MD de HisJBP mostraron que la forma apo es capaz de existir en una conformación cerrada (Chu et al., 2013).

Según los resultados de MD de Yang 2008, los ligandos estabilizan la conformación cerrada en lugar de inducir el cambio conformacional directamente (Yang et al., 2008). Los estudios de MD de LAOBP mostraron dos poblaciones llamadas "apo-open-state" y "open-twisted" con diferentes ángulos de torsión (Silva et al., 2011), pero los autores concluyeron que este estado no era un intermediario obligatorio en la transición hacia el estado abierto.

Para describir con datos experimentales cómo esta transición puede tener lugar y el papel del sustrato en la selección o inducción de la conformación cerrada de LAOBP se estudiaron los espectros de LAOBP en ausencia de sustrato a diferentes temperaturas y con otro sustrato como Ala (Figura III.4).

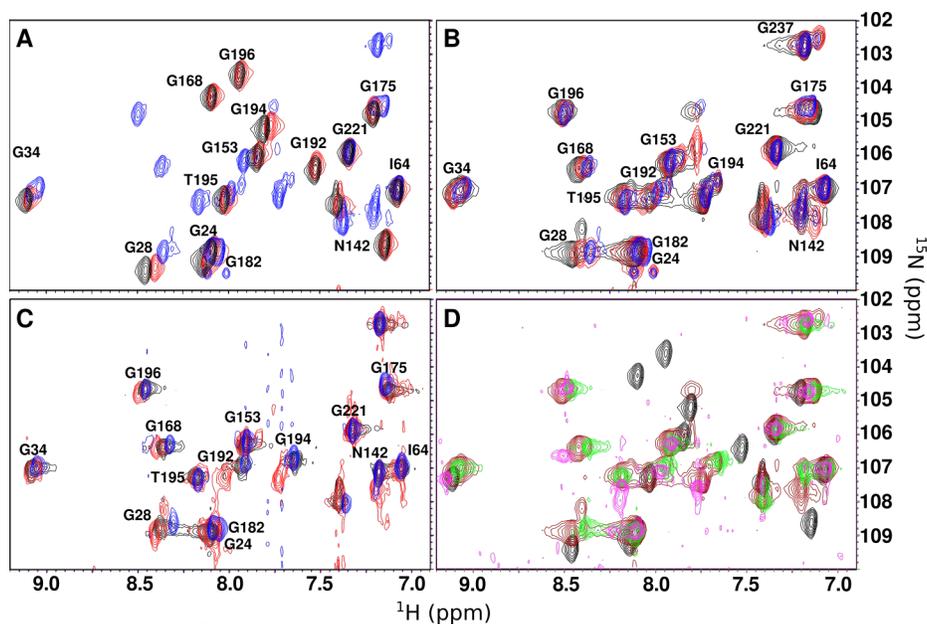


Figura III.4: Espectros ^1H - ^{15}N HSQC a distintas temperaturas en ausencia y presencia de ligando. Los espectros a 25 °C (negro), 30 °C (rojo) y 37 °C (azul) se muestran para LAOBP (A), 2P-LAOBP (B) y el complejo LAOBP-Arg (C). En el panel D se muestra la superposición de los espectros a 30 °C de LAOBP (negro), 2P-LAOBP (rojo) LAOBP-Arg (violeta) y LAOBP-Ala (verde).

En estudios previos de RMN el factor temperatura se ha utilizado para facilitar la asignación de PBPs (Igarashi et al., 2010). Sin embargo, este factor experimental resulta útil para explorar las diferentes conformaciones que las proteínas pueden adoptar. Aquí exploramos la CSP a diferentes temperaturas (25, 30 y 37 °C). Como se muestra en la Figura III.4, el aumento de la temperatura de 25 a 30 °C no indujo un cambio significativo en las señales, pero el incremento a 37 °C ocasionó perturbaciones en algunos residuos particulares. Este resultado indica que la forma apo de LAOBP puede existir en otra conformación, la cual llamamos 2P-LAOBP en referencia al segundo estado poblacional.

La forma y distribución del espectro 2P-LAOBP es similar a la observada en espectros de los complejos de LAOBP con Arg y Ala. Esto sugiere que LAOBP en su estado libre puede acceder a una conformación que se puede detectar con la modificación de la temperatura. Al bajar la temperatura de 37 a 25 °C no se eliminó la segunda población, lo cual sugiere que cuando se alcanza esta segunda conformación es posible que el regreso a la conformación inicial esté energéticamente desfavorecido.

En el estudio estructural de LAOBP-ornitina mediante MD se detectaron dos conformaciones (A y B) de la cadena principal de los residuos D11 y T12, y en el estado cerrado sin ligando, la conformación más estable fue la misma encontrada en la estructura del complejo 1LAF (Yang et al., 2008).

Tal como se había propuesto por Wolf, 1996, el paisaje conformacional de LAOBP se puede describir mediante un modelo en el que LAOBP existe en diferentes conformaciones dependiendo del sustrato, alguna de las cuales puede ser visitada en el estado no unido. Además, Silva et al., 2011, propusieron que el efecto de la unión a cada uno de los diferentes ligandos es desplazar el equilibrio a una región particular del estado cerrado de todas las conformaciones posibles.

Estudiando la dinámica de LAOBP, quizás se explique la diferencia en afinidad y especificidad mostradas por la proteína LAOBP.

Consideraciones finales

El estudio conformacional de LAOBP y su interacción con aminoácidos había estado limitado al uso de dinámica molecular y algunas técnicas experimentales. Aun cuando se cuenta con múltiples estructuras cristalográficas de la proteína y de sus complejos, no se ha logrado una descripción del cambio conformacional desde el estado abierto al cerrado o viceversa. Tampoco se ha logrado entender que aspectos estructurales modulan la amplia especificidad de la proteína. En este proyecto se estudió este sistema mediante NMR. Con los resultados obtenidos se logró la asignación de las formas libre y unida de LAOBP, lo cual facilitará próximos estudios de caracterización. Además, se detectó una segunda población en la forma libre que pareciera ser similar a la conformación cerrada en presencia de ligando. Estos resultados sirven de base para explorar mediante NMR los aspectos aún no esclarecidos de la interacción de LAOBP con sus ligandos. La hipótesis del proyecto “La unión de arginina a LAOBP afecta los desplazamientos químicos de varias regiones de la proteína además de las directamente relacionadas con la unión de ligando” fue aceptada.

III.7 Conclusiones

- 1) La asignación de los desplazamientos químicos de LAOBP y su complejo con arginina, fue depositada en la BMBR, lo que facilitará posteriores estudios de caracterización de esta proteína.
- 2) La forma apo de LAOBP visita distintas conformaciones semejantes al estado cerrado, lo que sugiere que el ligando selecciona una conformación para establecer el complejo binario.
- 3) Los cambios en los desplazamientos químicos de LAOBP tras la unión con arginina aparecen en los residuos del centro activo y en regiones que quizás interactúan con membrana, lo que sugiere que el complejo adquiere tras la unión la conformación óptima para que sea reconocido por el transportador transmembranal.

III.8 Recomendaciones

- 1) Realizar experimentos de dinámica por RMN en distintas escalas de tiempo de la forma libre y unida de LAOBP, para caracterizar la especificidad de esta proteína.
- 2) Caracterizar LAOBP-2P por RMN, como un posible intermediario en la transición de LAOBP libre al complejo.

III.9 Bibliografía

- Ames, G. F.-L., & Mimura, C. S. (1992). *Traffic ATPases: A Superfamily of Transport Proteins Operating from Escherichia coli to Humans*. 65, 1–47.
- Ames, G. F.-L., & Nikaido, K. (1978). Identification of a membrane protein as a histidine transport component in *Salmonella typhimurium*. *Proceedings of the National Academy of Sciences of the United States of America*, 75(11), 5447–5451.
- Banda-Vázquez, J., Shanmugaratnam, S., Rodríguez-Sotres, R., Torres-Larios, A., Höcker, B., & Sosa-Peinado, A. (2018). Redesign of LAOBP to bind novel l-amino acid ligands. *Protein Science*, 27(5), 957–968.
- Chu, B. C.H., Chan, D. I., Dewolf, T., Periole, X., & Vogel, H. J. (2014). Molecular dynamics simulations reveal that apo-HisJ can sample a closed conformation. *Proteins: Structure, Function and Bioinformatics*, 82(3), 386–398.
- Chu, Byron C H, Dewolf, T., & Vogel, H. J. (2013). *Role of the Two Structural Domains from the Periplasmic Escherichia coli Histidine-binding Protein HisJ*, 288(44), 31409–31422.
- Chu, Byron C H, DeWolf, T., & Vogel, H. J. (2013). Role of the two structural domains from the periplasmic *Escherichia coli* histidine-binding protein HisJ. *The Journal of Biological Chemistry*, 288(44), 31409–31422.
- Higgins, C. F., & Ames, G. F. (1981). Two periplasmic transport proteins which interact with a common membrane receptor show extensive homology: complete nucleotide sequences. *Proceedings of the National Academy of Sciences of the United States of America*, 78(10), 6038–6042.
- Holland, I., Cole, S., Kuchler, K., & Higgins, C. (2003). *ABC proteins: from bacteria to man*. London, UK: Academic Press.
- Igarashi, S., Osawa, M., Shin-ichiro, O., & Ichio, S. (2010). Backbone resonance assignments for the ligand binding subunit of the histidine permease complex (HisJ) from *Escherichia coli*, under histidine-bound and unbound states. *Biomolecular NMR Assignments*, 4(1), 17–20.
- Kang, C. H., Shin, W. C., Yamagata, Y., Gokcen, S., Ames, G. F., & Kim, S. H. (1991). Crystal-Structure of the Lysine-Binding, Arginine-Binding, Ornithine-Binding Protein (Lao) from *Salmonella-Typhimurium* at 2.7-Å Resolution. *Journal of Biological Chemistry*, 266(35), 23893–23899.
- Kerppola, R. E., Shyamala, V. K., Klebba, P., & Ames, G. F. L. (1991). The membrane-bound proteins of periplasmic permeases form a complex: Identification of the histidine permease HisQMP complex. *Journal of Biological Chemistry*, 266(15), 9857–9865.

- Keskin, O., Jernigan, R. L., & Bahar, I. (2000). Proteins with similar architecture exhibit similar large-scale dynamic behavior. *Biophysical Journal*, 78(4), 2093–2106.
- Kreimer, D. I., Malak, H., Lakowicz, J. R., Trakhanov, S., Villar, E., & Shnyrov, V. L. (2000). Thermodynamics and dynamics of histidine-binding protein, the water-soluble receptor of histidine permease implications for the transport of and low affinity ligands. *European Journal of Biochemistry*, 267(13), 4242–4252.
- Kustu, S. G., & Ames, G. F.-L. (1973). The *hisP* protein, a known histidine transport component in *Salmonella typhimurium*, is also an arginine transport component. *Journal of Bacteriology*, 116(1), 107–113.
- Oh, B H, Pandit, J., Kang, C. H., Nikaido, K., Gokcen, S., Ames, G. F., & Kim, S. H. (1993). Three-dimensional structures of the periplasmic lysine/arginine/ornithine-binding protein with and without a ligand. *The Journal of Biological Chemistry*, 268(15), 11348–11355.
- Oh, Byung Ha, Ames, G. F. L., & Kim, S. H. (1994). Structural basis for multiple ligand specificity of the periplasmic lysine-, arginine-, ornithine-binding protein. *Journal of Biological Chemistry*, 269(42), 26323–26330. <https://doi.org/DOI 10.1007/s10822-008-9215-6>
- Pistolesi, S., Yjandra, N., & Bermejo, G. (2011). Solution NMR studies of periplasmic binding proteins and their interaction partners. *Biomolecules Concepts*, 2(1), 53–64.
- Pulido, N. O., Silva, D.-A., Tellez, L. a, Pérez-Hernández, G., García-Hernández, E., Sosa-Peinado, A., & Fernández-Velasco, D. A. (2015). On the molecular basis of the high affinity binding of basic amino acids to LAOBP, a periplasmic binding protein from *Salmonella typhimurium*. *Journal of Molecular Recognition : JMR*, 28(2), 108–116.
- Sharff, A. J., Rodseth, L. E., Spurlino, J. C., & Quioco, F. A. (1992). Crystallographic Evidence of a Large Ligand-Induced Hinge-Twist Motion between the Two Domains of the Maltodextrin Binding Protein Involved in Active Transport and Chemotaxis. *Biochemistry*, 31(44), 10657–10663.
- Silva, D. A., Bowman, G. R., Sosa-Peinado, A., & Huang, X. (2011). A role for both conformational selection and induced fit in ligand binding by the lao protein. *PLoS Computational Biology*, 7(5), e1002054
- Silva, D. A., Domínguez-Ramírez, L., Rojo-Domínguez, A., & Sosa-Peinado, A. (2011). Conformational dynamics of L-lysine, L-arginine, L-ornithine binding protein reveals ligand-dependent plasticity. *Proteins: Structure, Function and Bioinformatics*, 79(7), 2097–2108.
- Yang, A. Y. C., & Mancera, R. L. (2008). Molecular dynamics simulations of ligand-induced backbone conformational changes in the binding site of the periplasmic lysine-, arginine-, ornithine-binding protein. *Journal of Computer-Aided Molecular Design*, 22(11), 799–814.

CAPÍTULO IV

Interacción CreCBD – ácidos nucleicos

Capítulo IV

IV. Interacción CreCBD – ácidos nucleicos

IV.1 Dominio amino de la proteína Cre, una recombinasa de ADN

IV.1.1 Recombinasas específicas de sitio: tirosinas recombinasas

Las recombinasas específicas de sitio son enzimas esenciales que catalizan el reordenamiento del ADN en una variedad de procesos biológicos, incluida la replicación y transcripción del genoma bacteriano, la integración y la escisión del ADN viral de los cromosomas del huésped. Precisamente, su alta especificidad las ha hecho blanco de numerosos estudios de caracterización que tienen como finalidad alguna aplicación biotecnológica o biomédica (Gibb et al., 2010)□.

Las recombinasas específicas del sitio se pueden agrupar en dos grandes familias: las tirosina recombinasas (YRs) o las serina recombinasas (SRs). Esta clasificación está basada en la naturaleza del residuo catalítico del mecanismo de reacción (Yang, 2010)□.

Las YRs presentan una tirosina catalítica que forma un enlace covalente fosfoproteína, para lograr el reordenamiento topológico del ADN, y el empalme de los fragmentos cortados del ADN (Meinke et al., 2016)□. Todas las YRs reconocen un sitio de recombinación central con una arquitectura similar, en el que dos moléculas de recombinasa se unen a dos elementos de unión a recombinasa (RBE). Los sitios de recombinación suelen oscilar entre 30 y 200 nucleótidos de longitud y consisten en dos motivos que están dispuestos como repeticiones invertidas que flanquean la región (secuencia espaciadora) donde la recombinación ocurrirá.

La secuencia espaciadora en sí no está estrictamente conservada, lo que permite una cierta flexibilidad en el reconocimiento. La mayoría de las YRs se han encontrado en procariontes y bacteriófagos donde realizan varias funciones, incluyendo la escisión e integración del ADN, cortar el plásmido en varios segmentos, la regulación de la expresión génica, la segregación de cromosomas durante la división celular, y la separación de ADN circulares catenados (Grindley et al., 2006)□.

Los sistemas de YR mejor caracterizados incluyen la integrasa Int del bacteriófago lambda, Cre del bacteriófago PI, Flp de levadura y XerC/D de *E. coli*. Las secuencias de

estas proteínas presentan un bajo porcentaje de identidad entre ellas y en general exhiben una estructura cuaternaria conservada en la conformación holiday junction y tienen una alta similitud en sus dominios amino y carboxilo (Meinke et al., 2016)□.

IV.1.2 Recombinasa Cre del bacteriofago P1

La recombinasa Cre (Cause recombination), está estructurada en dos dominios, el N-terminal (NTD) o Core Binding Domain (CBD) de 130 residuos y el C-terminal (CTD) o Catalytic Binding Domain (CAT) de 211 residuos (Van Duyne, 2014)□ (Figura IV.1). Cre no requiere proteínas accesorias del hospedero para organizar el complejo sináptico que inicia la reacción de recombinación. Esta propiedad ha permitido su utilización en el campo de la ingeniería genética.

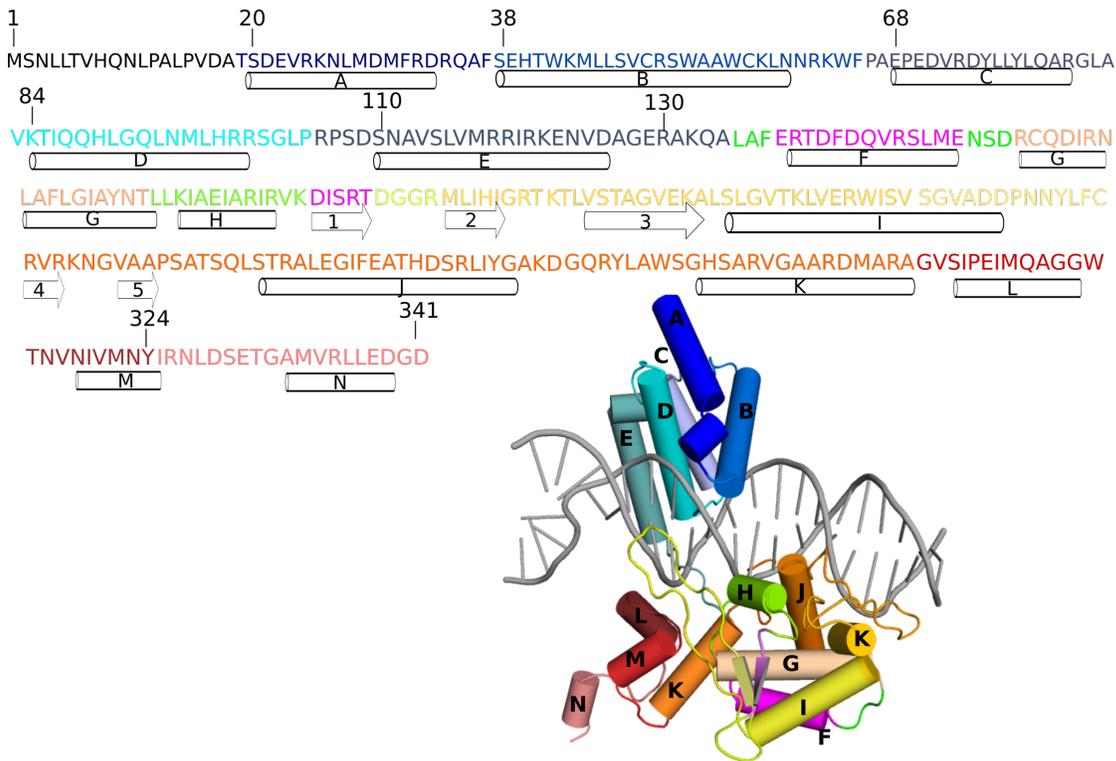


Figura IV.1: Secuencia y estructura terciaria de Cre. Arriba se muestra la secuencia de aminoácidos y su correspondiente estructura secundaria, haciendo corresponder los colores con la estructura tridimensional mostrada en el panel inferior. El esquema representa la estructura de un monómero de Cre unido a medio sitio loxP (gris). Las hélices alfa mostradas como tubos están localizadas en los dominios N-terminal (CBD) y C-terminal (CAT) y están nombradas de la A-N. Las hélices B y D de CBD se extienden sobre una cara del surco mayor, mientras la hélice J de CAT se sienta en el surco mayor desde la cara opuesta.

Los ensayos de la actividad de Cre son sencillos, y se han descrito metodologías para estudios *in vivo* e *in vitro*, los que permiten una rápida evaluación de la actividad de la enzima y sus mutantes (Gibb et al., 2010; Ringrose et al., 1998; Smedley et al., 2011)□. Cre se une con alta afinidad a los sitios loxP, con una Kd aparente menor a 1 nM determinada en ensayos de cambio de movilidad electroforética (EMSA, de sus siglas en inglés) y resonancia de superficie de plasmones (SPR, de sus siglas en inglés).

El sitio de reconocimiento loxP cuenta con 34 bp de los cuales 14 pb corresponden a cada uno de los dos elementos de unión a la recombinasa (RBEs) dispuestos a ambos lados del sitio de corte asimétrico de 6 pb (Figura IV.2). Los RBEs difieren solo en los pares de bases adyacentes al sitio de corte. Cuando Cre se une a un RBE del sitio loxP, los dos dominios de la proteína rodean el dúplex de ADN, formando una "abrazadera en forma de C" que establece contactos con los surcos mayor y menor del ADN (Figura IV.2). Dos elementos de unión claves son la hélice J del dominio CAT que se une al surco mayor cerca de la región central del sitio loxP y las hélices B/D del CBD que se extienden por el surco mayor cerca de la región de corte.

La proteína Cre se une de manera cooperativa al primer RBE con afinidad nanomolar y al segundo con afinidad subnanomolar (Guo et al., 1997; Rüfer et al., 2002)□.

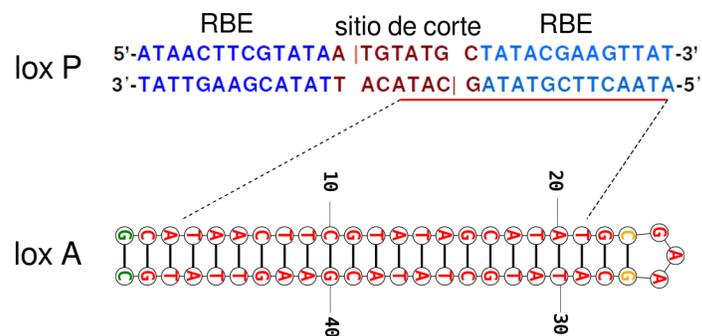


Figura IV.2: Secuencia de loxP. Se señala medio sitio de LoxP, identificado como lox A, el cual forma una horquilla.

Se establecen relativamente pocos contactos directos entre las cadenas laterales de la proteína y las bases del surco mayor. Los contactos más claros involucran a la R259, que interacciona mediante dos puentes de hidrógeno con los pares de bases C · G en las posiciones ± 10, y la N90, que se une por puentes de hidrógeno a los pares de bases A · T en las posiciones ± 5. De los dos contactos, la Arg259 juega un papel crucial en la

interacción pues es importante para la afinidad y la especificidad (Kim et al., 2001)□. Los residuos E262 y S257 también forman puentes de hidrógeno. Por otra parte, en el complejo ternario Cre-Cre-loxP los dos dominios CBD interactúan mediante las hélices A y E y los dos CAT lo hacen a través de hélice N. La hélice N de una subunidad está enterrada en un bolsillo hidrófobic del monómero adyacente.

Se ha demostrado que el dominio CAT es necesario y suficiente para el reconocimiento y la escisión del ADN. El dominio CAT contiene los residuos catalíticos críticos, incluida la tirosina conservada (Y342) que escinde el ADN a través del ataque nucleofílico para formar un enlace fosfotirosina estable (Duyne, 2014)□.

Hay menos información sobre la función del dominio CBD, pero se ha propuesto que es importante para la unión al ADN, en el control de la direccionalidad de la recombinación y también en el mejoramiento de la actividad de la proteína completa.

El dominio amino de la YR del bacteriofago lambda (Int^{CBD}) ha sido previamente estudiado (Kamadurai & Foster, 2007; Rutherford & Van Duyne, 2014)□. Se reportó que a consecuencia de la unión, su estructura sufre un cambio conformacional de desplegado a plegado a través de un mecanismo de ajuste inducido. Mientras que para Cre, aún cuando se ha sugerido que el fragmento del residuo 1 al 20 de CBD es intrínsecamente desordenado (Meinke et al., 2016) porque no se ha observado en ninguna estructura cristalográfica del complejo, no se han realizado otros estudios estructurales de este.

IV.2 Justificación del proyecto: interacción CreCBD - ácidos nucleicos

El mecanismo de recombinación homóloga del bacteriofago T4 es mediado por la integrasa de DNA Cause Recombination (Cre). El sitio de recombinación es una secuencia palindrómica, de 34 pares de base (loxP). El mecanismo de recombinación ha sido ampliamente estudiado, sin embargo el papel de su dominio amino CBD no ha sido esclarecido.

De acuerdo al porcentaje de desorden estructural predicho para CBD y la ausencia de los primeros 20 residuos en la estructura cristalográfica de Cre, CBD podría contener una región intrínsecamente desordenada. Todas las estructuras reportadas para Cre, son de la proteína en complejo con el ADN, pero aún no se dispone de una estructura en solución de la proteína monomérica, lo que permitiría explorar aspectos como cambios conformacionales y su desorden estructural en el extremo amino.

Se ha reportado que el dominio CBD de la integrasa del bacteriofago lambda sufre un cambio conformacional de desplegado a plegado tras su unión al ADN. Es posible entonces que el mecanismo para el reconocimiento del ADN difiera entre las proteínas. A pesar de la baja identidad, existe cierta coevolución en esta región de las proteínas YR, con lo cual sería interesante saber si los dominios amino están involucrados en algún proceso de reconocimiento común para estas proteínas.

Este trabajo se realizó principalmente durante una estancia de investigación en la Universidad de Ohio, OSU, USA en una colaboración con el Prof. Mark Foster.

IV.3 Hipótesis

La unión de loxP al dominio amino de Cre induce cambios conformacionales en ambas moléculas.

IV.4 Objetivos

Objetivo general:

Estudiar la interacción de CBD con lox P mediante RMN.

Objetivos particulares:

1. Asignar los desplazamientos químicos de los átomos de la proteína CBD.
2. Determinar las zonas de interacción CBD con LoxP mediante RMN.
3. Caracterizar la interacción CBD – lox P mediante dicroísmo circular.
4. Determinar mediante ITC la constante de afinidad y parámetros termodinámicos de la interacción CBD – lox P.

IV.5 Métodos

IV.5.1 Expresión de la proteína recombinante CBD

La expresión heteróloga de proteínas se realizó a partir del plásmido p23A que contiene el gen que codifica para CBD. La cepa bacteriana *E. coli* BL21-Rosseta DE3 se utilizó como sistema de expresión de CBD. Se creció una colonia en 20 mL de medio LB (Luria Bertani) (NaCl:peptona:extracto de levadura 1:1:0.5%) toda la noche a una temperatura de 37 °C (precultivo). Se adicionó el antibiótico carbenicilina (0.1 mg/mL). El precultivo se adicionó a 500 mL de medio LB con antibiótico y se dejó crecer hasta que alcanzó una absorbancia de 0.8-1 a 600 nm. La sobreexpresión de la proteína se indujo con IPTG 0.3 mM durante toda la noche a 25 °C. Para los estudios de RMN se utilizaron 10 mL del precultivo y la inducción se realizó en 1 L de medio mínimo M9 (Na₂HPO₄·7H₂O, KH₂PO₄, NaCl, ¹⁵NH₄Cl, glucosa ¹³C, MgSO₄, CaCl₂, MnCl₂, FeCl₃, ZnSO₄ y vitaminas), para obtener las proteínas isotópicamente marcadas.

IV.5.1.1 Preparación de la muestra y purificación de CBD mediante cromatografía de intercambio iónico y exclusión molecular.

Los cultivos con CBD se centrifugaron a 2.000 x g a 4 °C por 30 minutos y se resuspendieron en una solución con Tris 40 mM pH 7.0, NaCl 100 mM, DTT 5 mM. Las células se sonicaron y se eliminaron los sedimentos (4000 x g durante 50 minutos). El lisado se filtró a través de un filtro GVS (tamaño de poro de 0.22 µm) para eliminar los residuos celulares insolubles.

La purificación de CBD se realizó mediante dos cromatografías de intercambio aniónico utilizando las columnas Hitrap SPFF y Hitrap Heparin y una cromatografía de exclusión molecular en la columna S75. La columna de intercambio aniónico Hitrap SPFF con un volumen de 5 mL de matriz (Vc=5 mL) se lavó con 5 Vc del tampón de equilibrio. El extracto crudo se aplicó lentamente a la matriz. La proteína se eluyó de la columna con 3Vc de tampón de elución (Tris 40 mM pH 7.0, NaCl 500 mM, DTT 5 mM), utilizando un gradiente lineal de NaCl (0–250 mM) y colectándose fracciones de 5 mL. Las fracciones correspondientes a la proteína CBD se juntaron y diluyeron antes de su aplicación en la siguiente columna. Las condiciones del segundo paso de purificación fueron las mismas. Ambas columnas se lavaron con 5Vc de tampón de elución y se reequilibraron con 10Vc

del tampón de equilibrio. Por último, tras la concentración de la muestra purificada se utilizó una cromatografía de exclusión molecular utilizando la columna S75. La presencia de la proteína se siguió a lo largo del protocolo mediante electroforesis en geles SDS-PAGE.

Cada muestra purificada se analizó mediante espectrometría de masas MALDI-TOF y DC para comprobar su masa molecular y estructura secundaria.

IV. 5.2 Ensayos de calorimetría de titulación isotérmica

En cada experimento, se utilizó un volumen de inyección de 2 μ L del titulante y 50 μ M de proteína. Se empleó un intervalo de 180 s entre inyecciones para permitir una equilibración completa. La titulación blanco consistió en la solución valorante idéntica pero que contenía sólo la solución tampón en la celda de muestra. La solución en la celda muestra se agitó a 750 rpm para asegurar un mezclado rápido. Los experimentos se realizaron en un equipo MicroCalTMiTC₂₀₀ a 25°C. Los datos se analizaron con el empleo del software SEDPHAT y graficados utilizando el programa GUSI (Zhao et al., 2015)□.

IV.5.3 Estudio de la interacción proteína – ligando mediante RMN en solución.

IV.5.3.1 Asignación de los desplazamientos químicos de los átomos de CBD.

Los espectros de RMN se adquirieron a 298 K usando campos magnéticos con distintas fuerzas (14.1, 16.4 and 18.8 T) que se corresponden con las frecuencias de ¹H 600, 700 y 800 MHz, respectivamente. Se utilizaron los equipos de The Ohio State University (OSU). Para la asignación de la cadena principal de las proteínas se adquirieron espectros del tipo HNCO, HNCACB, CBCA(CO)NH, HNCA (Grzesiek & Bax, 1992)□, ¹H–¹⁵N HSQC (Palmer et al., 1991)□, mientras que la asignación de la cadena lateral se obtuvo usando los experimentos H(CCO)NH y C(CO)NH (Montelione et al., 1992)□.

Los experimentos se analizaron en NMRpipe (Delaglio et al., 1995)□ y CARA (Keller, 2004)□.

IV.5.3.2 Determinación de las zonas de interacción de CBD y lox A mediante RMN

La comparación de espectros tipo ^1H - ^{15}N HSQC de la proteína CBD y su complejo con loxA permitió identificar los residuos afectados. De forma similar se compararon espectros de 1D de protón para identificar cambios en la molécula de ADN.

IV.5.4 Determinación de la composición de estructura secundaria por DC.

Se adquirieron espectros de DC en la región 200-300 nm. Se utilizaron 20 μM de proteína CBD en PBS 20 mM pH 7.5 y una relación equimolar para el estudio del complejo a temperatura ambiente. Los espectros se colectaron en un espectropolarímetro de DC JASCO J-815 (OSU, USA) y se deconvolucionaron utilizando el programa CAPITO (Wiedemann et al., 2013)□.

IV.5.5 Estabilidad térmica de CBD y del complejo CBD-loxP.

La estabilidad térmica de CBD y su complejo se siguió con la determinación de la T_m de ambos. Se utilizaron 20 μM de proteína CBD en PBS 20 mM pH 7.5 y una relación equimolar para el estudio del complejo. Se siguió el cambio de elipticidad molar a 222 nm en función de la temperatura (25 – 90 °C). Los espectros se colectaron en un espectropolarímetro JASCO J-815 (OSU, USA).

IV.6 Resultados

IV.6.1 Caracterización del estado nativo de CBD

Como parte de los estudios de caracterización de CBD, la proteína purificada isotópicamente marcada fue analizada por espectrometría de masas, RMN y DC (Figura IV.3A). La masa molecular obtenida mediante MALDI se corresponde a la esperada para un monómero de proteína (14.9 kDa) con un porcentaje de marcaje ^{15}N logrado del 99.38%.

La composición de estructura secundaria de CBD se determinó mediante DC, el cual relaciona la elipticidad molar con la longitud de onda (Figura IV.3B). Se observaron bandas alrededor de los 195, 208 y 220 nm, señales típicas de la formación de hélices alfa.

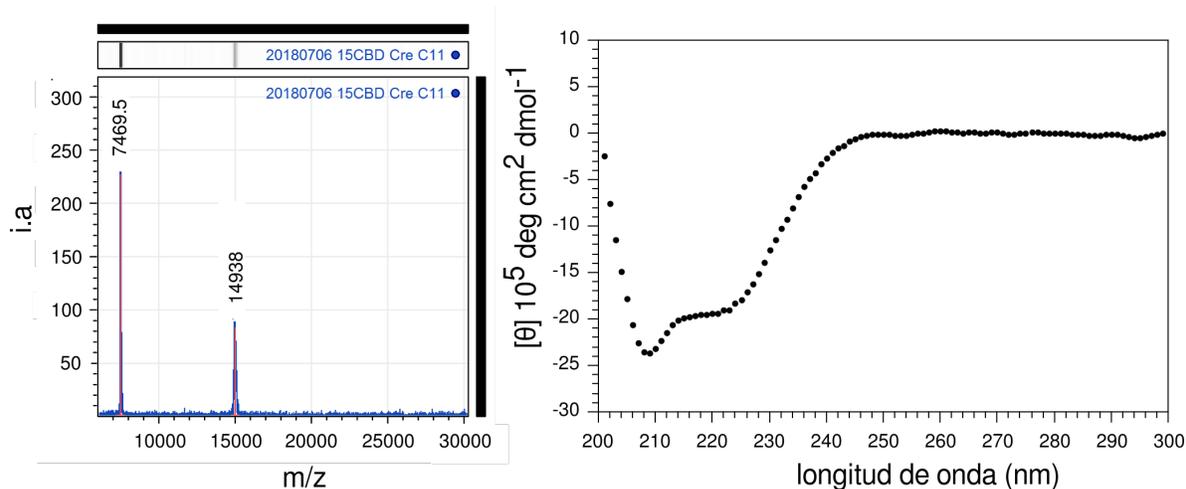


Figura IV.3: Se muestra el espectro de masas MALDI-TOF y el de dicroísmo circular de la proteína marcada.

La deconvolución del espectro de CD, usando el programa CAPITO arrojó que la cantidad de estructura secundaria de CBD está distribuida como: hélice α (0.38), hoja β (0.06) y random coil (0.43). Si se analiza la cantidad de hélice α (~ 70 %) de acuerdo a la estructura cristalográfica, vemos que la cantidad obtenida por DC es significativamente menor. Los datos de DC, sugieren que CBD contiene una IDR de alrededor de 40 residuos.

También estudiamos el estado oligomérico de CBD para lo que realizamos experimentos de difusión tipo DOSY y la estimación del radio hidrodinámico usando el programa HYDRONMR. Este software permite la predicción de los coeficientes de difusión rotacional y traslacional anisotrópicos y parámetros de relajación de RMN de proteínas. El

experimento DOSY permite la determinación del coeficiente de difusión de la molécula y con ello el radio hidrodinámico. Los coeficientes de difusión teórico y experimental fueron de 1.04×10^{-6} y 1.46×10^{-6} cm^2/s los que se corresponderían con Rh de 2.01 y 1.49 nm para moléculas esféricas, respectivamente. Los datos de DOSY indican que aunque CBD son del mismo orden, el Rh no corresponde muy bien con el teórico, tal vez debido a la presencia de una conformación desplegada en sus primeros residuos. Al analizar la secuencia del 1-40 de CBD en el programa CIDER (Classification of Intrinsically Disordered Ensemble Regions) (Holehouse et al., 2017)□, se encontró que el segmento se ubica en la región 2, que corresponde con secuencias que pueden estar extendidas o colapsadas en dependencia del contexto. Este resultado podría justificar el Rh encontrado para CBD, el cual fue menor que el teórico.

IV.6.2 Caracterización termodinámica de la interacción CBD – loxP

El estudio preliminar de la termodinámica de la interacción permitió obtener algunos parámetros de la unión entre CBD – loxP (Figura IV.4). Se obtuvo una K_d de 9.4 μM . Los parámetros termodinámicos están en correspondencia con los reportados para la interacción del dominio N-terminal de la integrasa lambda (Int^{CBD}) (Kamadurai & Foster, 2007)□. En este artículo se discute que a cantidades subestequiométrica de ADN dos unidades de Int^{CBD} reconocen una molécula de ADN, mientras que cuando el ADN está en cantidades mayores la estequiometría es 1:1.

De acuerdo a los datos termodinámicos, la unión proteína - ADN está favorecida por el componente entálpico principalmente, lo que puede estar relacionado a interacciones electrostáticas (Jen-Jacobson, et al., 2000). Se ha sugerido que cuando la interacción proteína - ADN ocasiona una distorsión estructural del ADN leve el proceso está favorecido por la entalpía, mientras que los sistemas que distorsionan fuertemente el ADN tienen un ΔH neto desfavorable como resultado de la tensión molecular, principalmente asociada con el desapilamiento de los pares de bases (Jen-Jacobson et al., 2000)□.

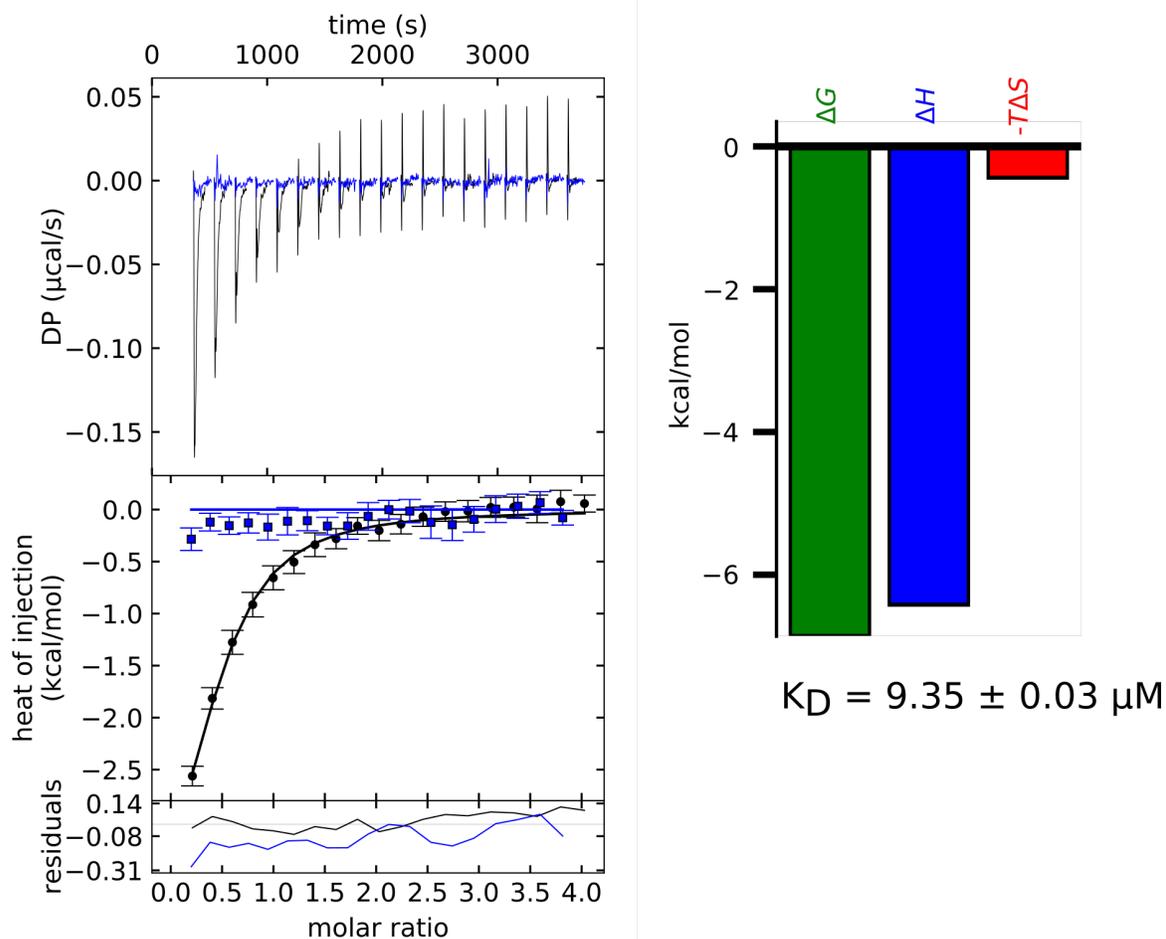


Figura IV.4: Ensayo de calorimetría entre CBD y LoxP a temperatura constante. En el panel superior se muestra el cambio de calor por inyección, y en el panel inferior la integración del área bajo la curva de cada inyección en función de relación molar entre la proteína y loxP. En negro se representan la titulación lox A en la celda con CBD en la jeringa y en azul la del tampón en ADN.

IV.6.3 Asignación de los desplazamientos químicos de los átomos de CBD

Para determinar la viabilidad de estudiar los dominios por separado, se comparó el espectro ^1H - ^{15}N HSQC de CBD con el de la proteína completa (Fig IV.5). Cada señal del dominio amino sobrelapa con alguna de Cre, lo que significa que los dominios por separado actúan independientemente.

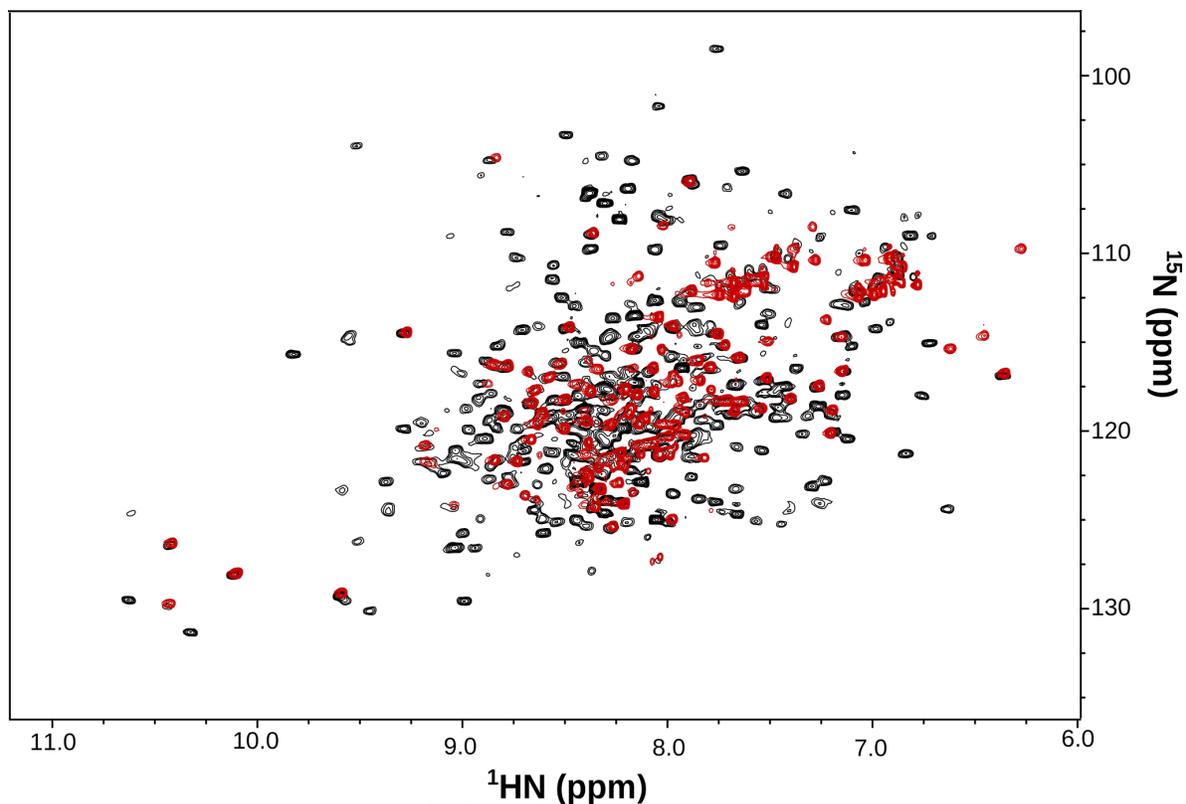


Figura IV.5: Espectros de NMR tipo ^1H - ^{15}N HSQC de Cre (negro) y CBD (rojo).

Con los experimentos HNCA, HNCOC, HNCACB, y TOCSY de nitrógeno y carbono se realizó la asignación de los desplazamientos químicos de los átomos de CBD (Fig IV.6). La cadena principal se asignó en un 68.7% y las cadenas laterales en un 31.7%.

En el espectro de CBD no se observaron todas las señales esperadas de acuerdo a la secuencia primaria, lo que pudiera estar relacionado a la presencia de un fragmento desplegado en la estructura de CBD que adopta múltiples estados conformacionales, lo cual no resultaría en señales. La escasez de señales no puede atribuirse a un estado oligomérico multimérico porque de acuerdo a los datos de DOSY la proteína es monomérica. Tampoco se debe a la ausencia del dominio CAT, porque según los estudios cristalográficos son pocos los contactos entre los dominios y el espectro de los dominios separados explica todos los picos. Entonces, CBD no está totalmente plegado en solución y que contiene una IDR.

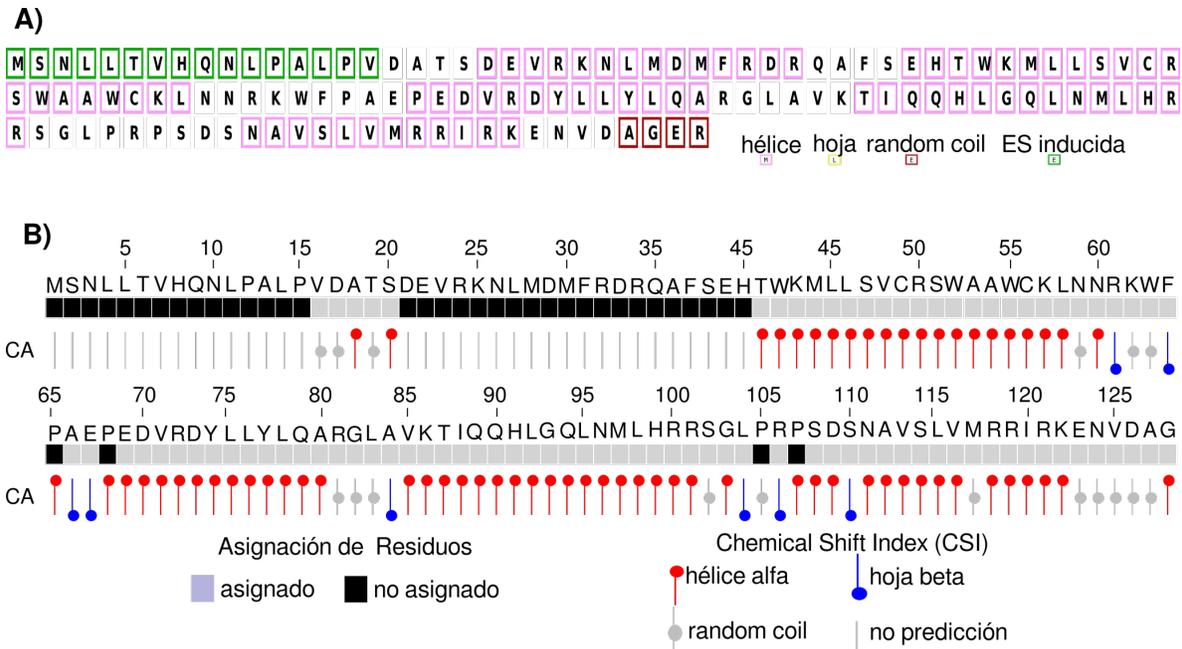


Figura IV.7: Predicción de estructura secundaria usando el programa DISOPRED (A) y análisis del CSI para cada residuo (B). En cuadros negros se identifican los residuos no asignados y en gris los asignados. La propensión a formar algún elemento de estructura secundaria se representa con símbolos rojos (hélices alfa), azul (hebras beta) y gris (random coil).

Como parte de la caracterización del dominio amino de Cre, respecto a si se pliega cuando une ADN, se determinó la composición de estructura secundaria de la proteína en ausencia y presencia de su blanco molecular, loxP, mediante DC (Figura IV.8). El espectro de DC de la proteína muestra las señales características de la formación de hélices alfa (208 y 220 nm), mientras que el de loxP contiene las de ADN a 250 y 280 nm. Bajo las condiciones utilizadas, se observó un incremento en el porcentaje de elipticidad molar de la proteína (~15%) cuando la molécula de ADN estaba presente en el ensayo. Un porcentaje similar de incremento de elipticidad molar fue observado durante el estudio de la interacción entre Int^{CB} y su blanco molecular (Kamadurai & Foster, 2007)□. Se sugirió no considerar como significativo este cambio en la señal de CD debido a dos razones: la primera, que el ADN absorbe en esta región, y la segunda, que cambios en el empaquetamiento interhelicoidal de proteínas puede afectar la señal en esta región del espectro de DC. Concluyéndose que la mayoría de los elementos de estructura secundaria ya estaban presentes en la proteína libre, pero se estabilizaron mediante la unión al ADN (Kamadurai & Foster, 2007)□.

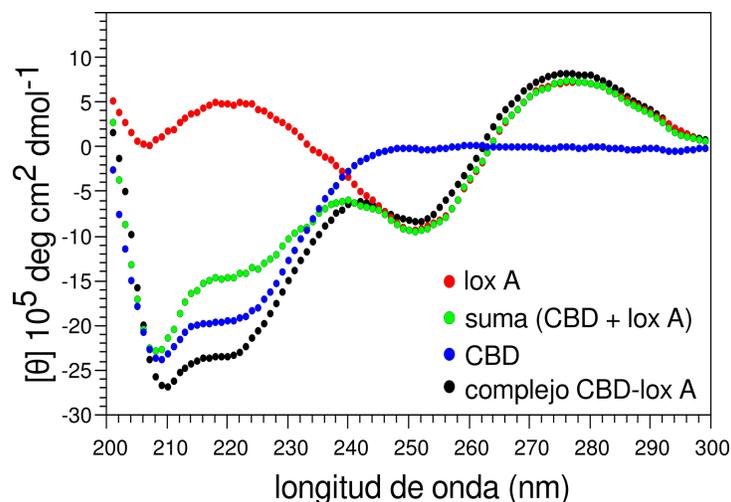


Figura IV.8: Espectros de CD en la región UV. Se muestran los espectros de CBD (azul), medio loxP (lox A) (rojo), del complejo (negro) y de la suma de los espectros de LlxP-CBD (verde).

IV.6.5 Ensayos de estabilidad térmica por CD

Para estudiar la termoestabilidad de CBD seguimos el cambio de la señal a 222 nm en función del aumento de la temperatura (Figura IV.9).

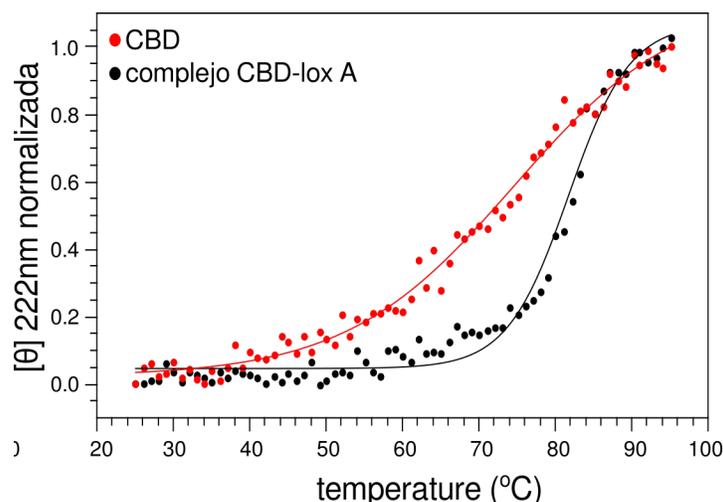


Figura IV.9: Cambio en la señal a 222nm en función del aumento de la temperatura para CBD (rojo) y el complejo lox A - CBD (negro).

La proteína es estable ya que presentó un valor de T_m de 74.5 ± 1 °C, pero se vuelve más estable en presencia del ADN ($T_m=84.5 \pm 0.4$ °C). Entonces, la ganancia en porcentaje de hélice y en estabilidad de la proteína indican la formación de un complejo. Además, la transición se vuelve mucho más cooperativa en el complejo. Las variaciones estructurales y los datos de ITC indican que CDB es capaz de reconocer e interactuar con el ADN.

IV.6.6 Estudios de interacción de CBD con loxP mediante RMN

Para estudiar la interacción proteína-ADN, se colectó un ^1H - ^{15}N HSQC del complejo (Figura IV.10). El espectro del complejo fue de baja calidad. Las señales del complejo son anchas, lo cual podría relacionarse con algún proceso de relajación o deberse a la formación de un complejo CBD – lox A con estequiometría 2:1. No obstante se observaron algunos cambios en los desplazamientos químicos de los átomos de la proteína CBD.

En la Figura IV.10 se puede observar que en el espectro del complejo no aparecen señales nuevas, lo que sugiere que la región desplegada no participa en el reconocimiento del ADN y no hay un evento de plegamiento. Podemos concluir que la proteína CBD muestra CSP en algunos residuos en presencia de lox A.

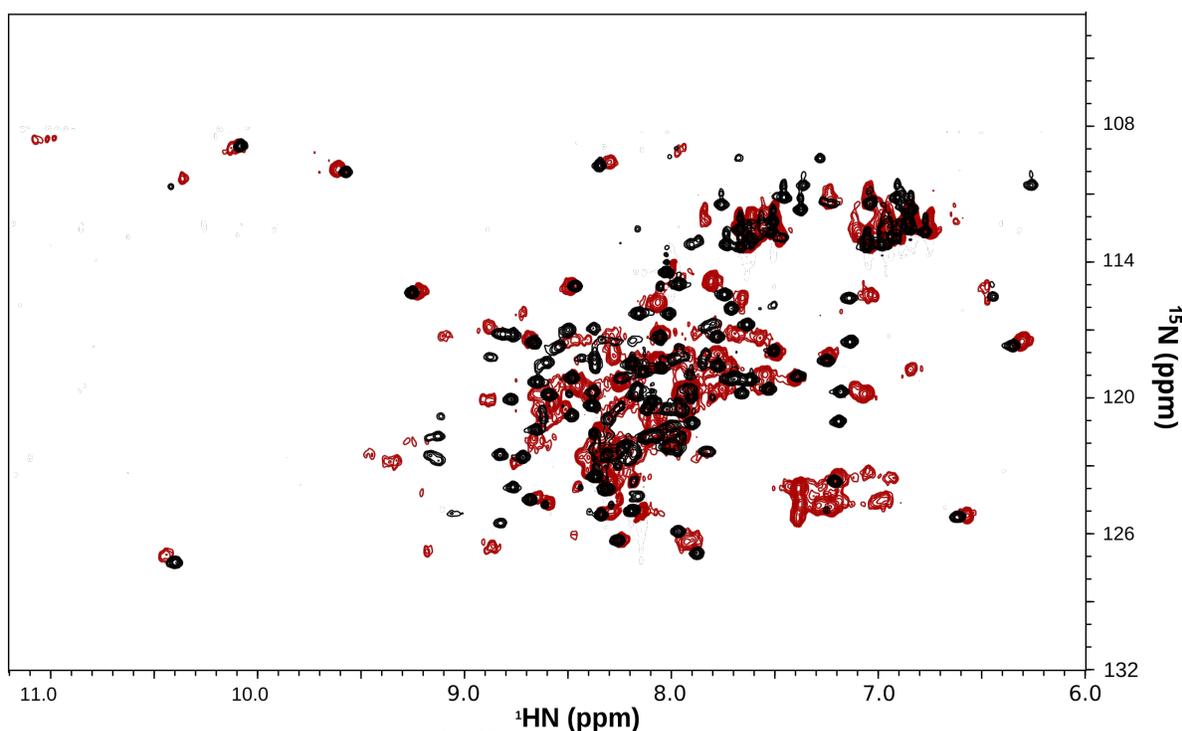


Figura IV.10: Espectros de NMR tipo ^1H - ^{15}N HSQC de CBD (negro) y CBD – LoxP (rojo).

IV.6.7 Cambios en la estructura de loxP asociados a la unión a CBD

Para estudiar el cambio en la estructura de la molécula de ADN, analizamos espectros de 1D de protón (Figura IV.11). En la región del espectro que se muestra (11 a 15 ppm) los protones de la proteína no tienen señal. En esta parte se pueden observar los protones de los grupos imino de los nucleótidos de guanina y timina cuando hay un duplex. De acuerdo con la secuencia de lox A, esperábamos alrededor de 19 señales, de las cuales podemos

diferencial al menos 16, lo que puede estar relacionado a la movilidad de la molécula. En el espectro del complejo CBD – lox A aparecieron nuevas señales, las que resultan interesantes ya que son resultado de la unión. Estas señales podrían corresponder al loop de lox A, el que puede estabilizarse en el complejo.

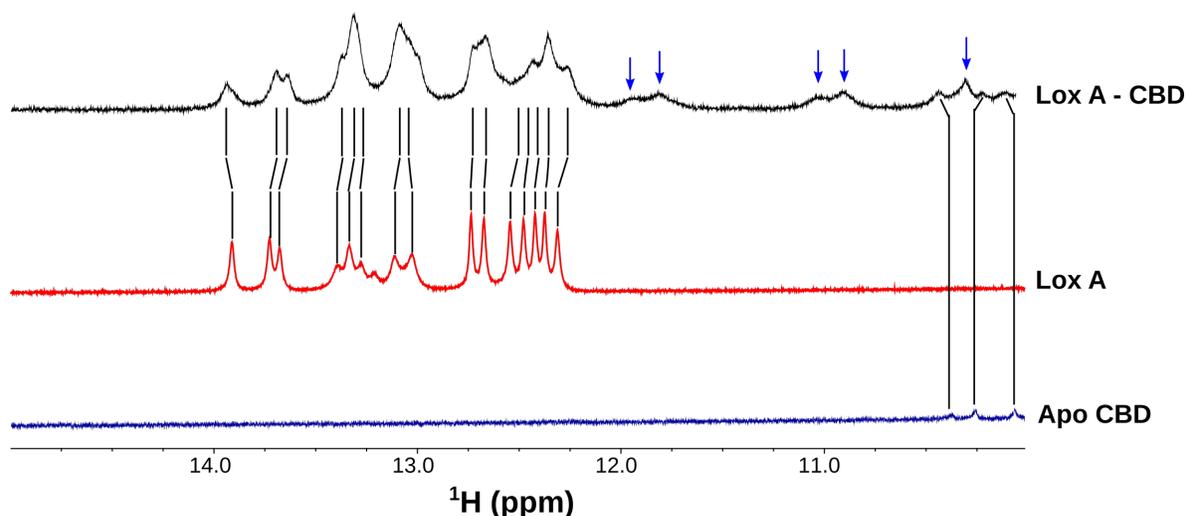


Figura IV.11: Espectros de 1 dimensión de protón. En azul, rojo negro se representan los espectros de CBD, lox A y de su complejo, respectivamente.

Entonces, la unión de CBD indujo cambios en la estructura de loxA. Se ha sugerido que los dominios de unión a ADN aseguran el reconocimiento selectivo de los REB no solo promoviendo la unión del dominio catalítico con una alta afinidad sino también por inducir una curvatura de ADN muy parecida al estado de transición (Kamadurai & Foster, 2007)□.

Consideraciones finales

La recombinasa Cre ha sido ampliamente estudiada por su promisoriosa aplicación biotecnológica. Debido a que su dominio catalítico es responsable de la actividad biológica de la proteína ha sido caracterizado mediante ensayos bioquímicos y varias técnicas biofísicas. Sin embargo el dominio amino, CBD ha sido menos estudiado. En este proyecto se determinó que CBD contiene una IDR, la cual se mantiene desplegada tras la interacción. La unión entre CBD y LoxP tiene una Kd moderada, vuelve al complejo más estable e induce cambios conformacionales tanto en la proteína como en la molécula de ADN, lo cual nos permitió aceptar la hipótesis planteada “La unión de loxP al dominio amino de Cre induce cambios conformacionales en ambas moléculas”, sin embargo no induce un plegamieto de la IDR.

IV.7 Conclusiones

- 1) El fragmento 1-40 de CBD es intrínsecamente desordenado en el monómero y existe en distintas conformaciones.
- 2) La proteína CBD tiene una afinidad moderada por loxP. El complejo macromolecular es más estable que la forma apo y conserva cierta dinámica en residuos particulares. El reconocimiento molecular puede involucrar el mecanismo de selección o inducción conformacional.
- 3) Ambas moléculas muestran cambios en los desplazamientos químicos tras la unión. El ADN puede adquirir una conformación que favorezca la interacción con el dominio catalítico de Cre.
- 4) La unión del ADN a CBD de Cre no induce el plegamiento del extremo amino a diferencia del CDB de la integrasa del phago lambda.

IV.8 Recomendaciones

- 1) Asignar las señales de RMN de CBD en complejo con ADN.
- 2) Realizar estudios de dinámica por RMN en las distintas escalas de tiempo de CBD y del complejo.
- 3) Realizar estudios de interacción entre CBD, CAT y lox P.
- 4) Obtener estructura tridimensional de CBD en solución por RMN.

IV.9 Bibliografía

- Delaglio, F., Grzesiek, S., Vuister, G., Zhu, G., Pfeifer, J. and Bax A. (1995). “NMRPipe: A Multidimensional Spectral Processing System Based on UNIX Pipes.” *Journal of Biomolecular NMR*. 6(3), 277–93.
- Gibb, B., Gupta, K., Ghosh, K., Sharp, R., Chen, J. and Van Duyne, G. (2010). “Requirements for Catalysis in the Cre Recombinase Active Site.” *Nucleic Acids Research*. 38(17), 5817–32.
- Grindley, N. D.F., Whiteson, K. and Rice, P. A. (2006) “Mechanisms of Site-Specific Recombination.” *Annual Review of Biochemistry*. 75 (1),567–605.
- Grzesiek, S, and Bax, A. (1992)m“Improved 3D Triple-Resonance NMR Techniques Applied to a 31 kDa Protein.” *J Mag Reson*. 96, 432–440.
- Guo, F., Gopaul, D. and Van Duyne, G. (1997)“Structure of Cre Recombinase Complexed with DNA in a Site-Specific Recombination Synapse.” *Nature*. 389(6646), 40–46.
- Holehouse, A. S., Das, R., Ahad, J., Richardson, M. and Pappu, R. (2017) “CIDER: Resources to Analyze Sequence-Ensemble Relationships of Intrinsically Disordered Proteins.” *Biophysical Journal*. 112(1), 16–21.
- Jen-Jacobson, L., Engler, L. and Jacobson, L. (2000) “Structural and Thermodynamic Strategies for Site-Specific DNA Binding Proteins.” *Structure*. 8(10),1015–23.
- Kamadurai, H. B., and Foster. M. P. (2007) “DNA Recognition via Mutual-Induced Fit by the Core-Binding Domain of Bacteriophage λ Integrase.” *Biochemistry*. 46(49), 13939–47.
- Keller, R. *The Computer Aided Resonance Assignment Tutorial*. Goldau, Switzerland: Cantina Verlag, 2004.
- Kim, S., Kim, G., Lee, Y. and Park, J. (2001) “Characterization of Cre-loxP Interaction in the Major Groove: Hint for Structural Distortion of Mutant Cre and Possible Strategy for HIV-1 Therapy.” *Journal of Cellular Biochemistry*. 80(3), 321–27.
- Meinke, G., Bohm, A., Hauber, J., Pisabarro, M. and Buchholz, F. (2016) “Cre Recombinase and Other Tyrosine Recombinases.” *Chemical Reviews*. 116(20), 12785–820.
- Montelione, G.T., Lyons, B. A., Emerson, S. D. and Tashiro, M. (1992) “An Efficient Triple Resonance Experiment Using Carbon-13 Isotropic Mixing for Determining Sequence-Specific Resonance Assignments of Isotopically-Enriched Proteins.” *Journal American Chemical Society*. 114, 10974–10975.
- Palmer, AG III., Cavanagh, J., Wright, P. and Rance, M. (1991) “Sensitivity Improvement in Proton-Detected Two-Dimensional Heteronuclear Correlation NMR Spectroscopy.” *J Mag Reson*. 93, 151–170.
- Ringrose, L., Lounnas, V., Ehrlich, L., Buchholz, F., Wade, R. and Stewart, A. (1998) “Comparative Kinetic Analysis of FLP and Cre Recombinases: Mathematical Models for DNA Binding and Recombination.” *Journal of Molecular Biology*. 284, (2) 363–84.
- Rüfer, A., Neuenschwander, P. and Sauer, B. (2002) “Analysis of Cre-loxP Interaction by Surface Plasmon Resonance: Influence of Spermidine on Cooperativity.” *Analytical Biochemistry*. 308(1), 90–99.

- Rutherford, K. and Van Duyne, G. (2014) “The Ins and Outs of Serine Integrase Site-Specific Recombination.” *Curr Opin Struct Biol.* 24, 125–31.
- Smedley, Damian, Ekaterina Salimova, and Nadia Rosenthal. (2011) “Cre Recombinase Resources for Conditional Mouse Mutagenesis.” *Methods.* 53(4), 411–16.
- Van Duyne, G D. “Cre Recombinase. (2014)” *Microbiol Spectrum.* 3(1), 1–19.
- Wiedemann, C., Bellstedt, P. and Görlach, M. (2013) “CAPITO - A Web Server-Based Analysis and Plotting Tool for Circular Dichroism Data.” *Bioinformatics.* 29(14), 1750–57.
- Yang, W. (2010) “Topoisomerases and Site-Specific Recombinases: Similarities in Structure and Mechanism.” *Crit Rev Biochem Mol Biol.* 45(6), 520–34.
- Zhao, H., Piszczek, G. and Schuck, P. (2015) “SEDPHAT – a Platform for Global ITC Analysis and Global Multi-Method Analysis of Molecular Interactions.” *Methods.* 76, 137–48.

Conclusiones

Geerales

de

la

Tesis

Doctoral

IX Anexos

IX.1. Capítulo II

IX.1.1 Determinación de la masa molecular de péptidos sintéticos de AtLEA4-5 Análisis espectrométrico

Las muestras fueron previamente desaladas con Zip Tip C18 (Millipore; Billerica, MA, USA) y aplicadas en un sistema LC-MS (Liquid Chromatography-Mass Spectrometry) compuesto de una bomba Dionex (Thermo-Fisher Co.; San Jose, CA, USA) acoplado a un espectrómetro de masas LTQ-Orbitrap Velos (Thermo-Fisher Co., San Jose, CA, USA) con fuente de ionización tipo nano-electrospray (ESI). Los datos fueron colectados en el Laboratorio Universitario de Proteómica IBT/UNAM.

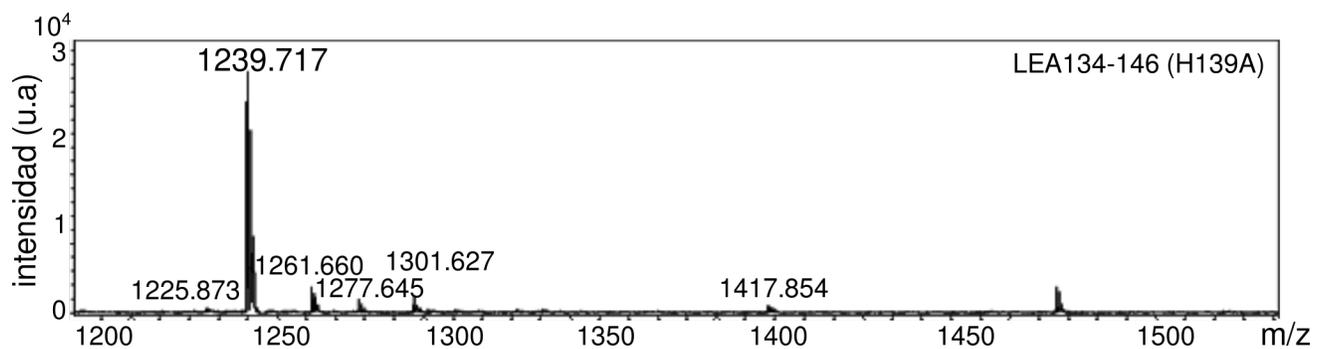
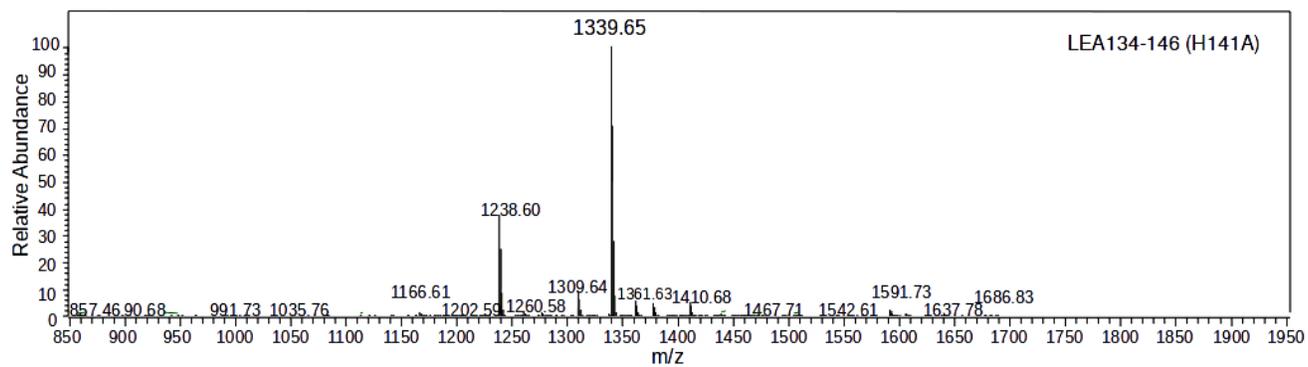
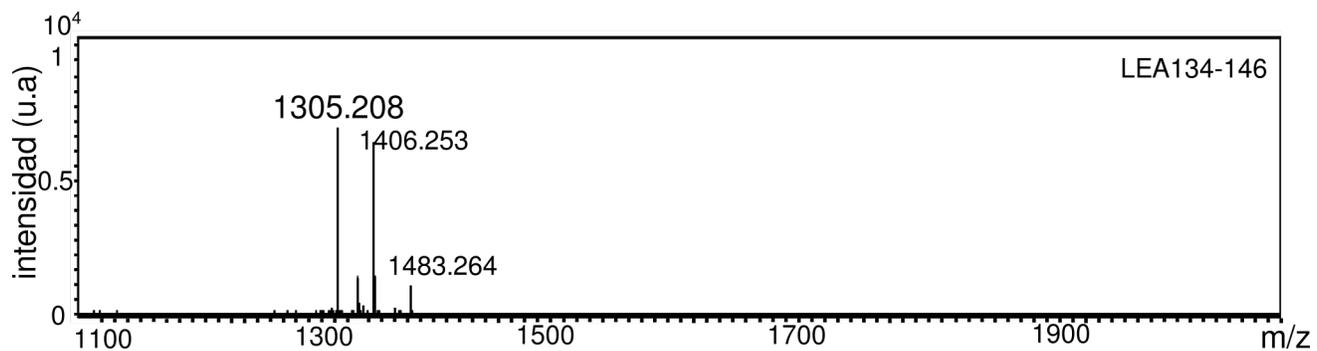
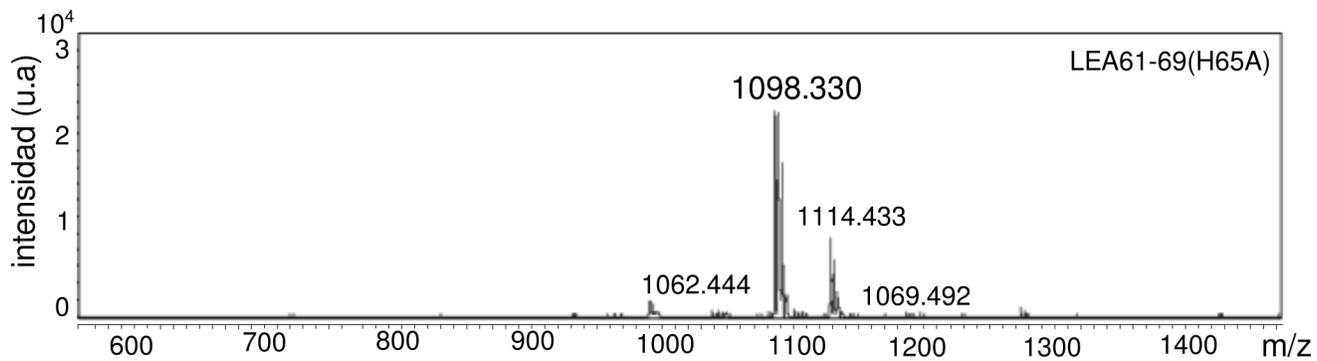
La calibración del espectrómetro se realizó con una solución Calmix (N-butilamina, cafeína, Met-Arg-Phe-Ala (MRFA) y Ultramark 1621. Estos calibrantes se utilizan para calibrar el módulo LTQ Velos con trampa de iones y el módulo Orbitrap con detector de masas FT (Fourier Transform) en modo de ionización positiva ESI. La N-butilamina se incluye para extender la calibración de masas a menores valores de m/z . Este tipo de calibración permite determinaciones de masas moleculares con exactitudes menores a 5 ppm (partes/millón).

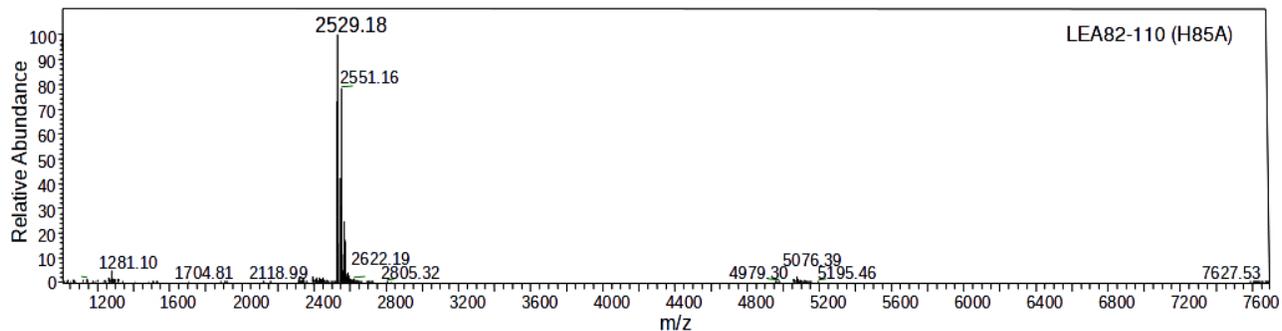
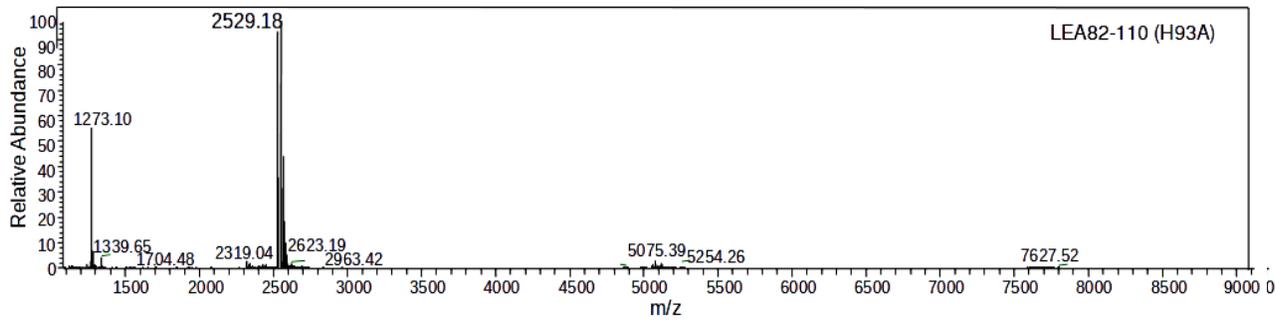
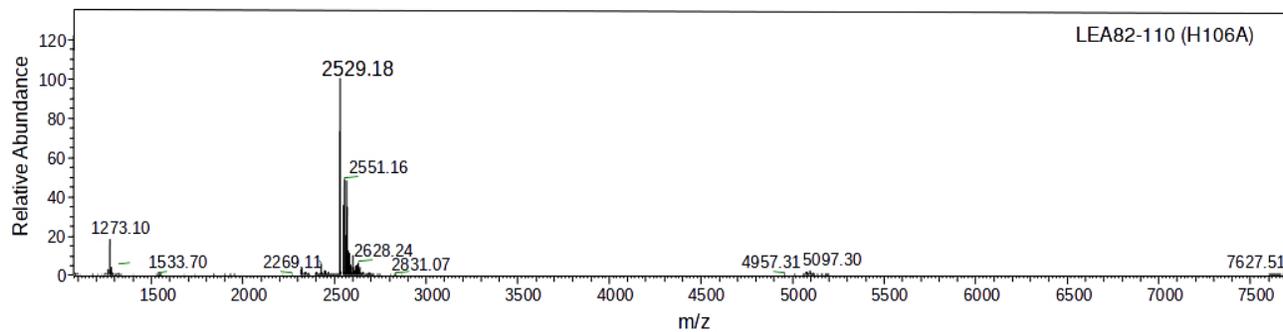
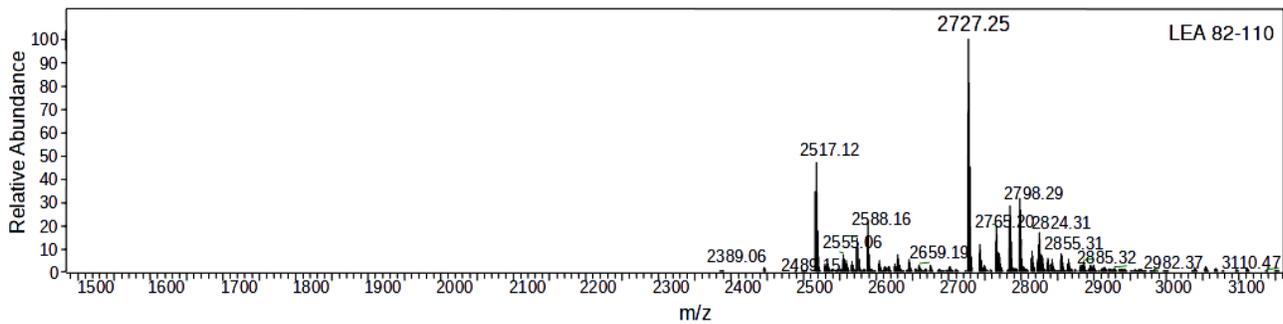
En la cromatografía de líquidos de nanoflujo en línea se utilizó un sistema isocrático de 50-50% de solvente (agua/acetonitrilo con 0.1% de ácido fórmico) con un tiempo de corrida de 30 minutos utilizando una aguja capilar hecha en casa (ID 0.75 μm y 10cm). El flujo del sistema LC fue de 300 nanolitros/minuto.

El escaneo total de iones (*Full Scan*) se realizó en el analizador Orbitrap con un poder de resolución de masas (Resolution Power; $RP=m/\text{FWHM}$) de 60,000.

Las masas moleculares de cada muestra se obtuvieron mediante el procesamiento de datos a través del algoritmo de Deconvolución espectral automática (Xtract RAW file) del software Xcalibur.

En la siguiente tabla se muestran las masas moleculares encontradas para cada muestra. Las masas moleculares 3KDa son reportadas como monoisotópicas y 3KDa son reportadas como masas moleculares AVERAGE debido a la contribución de isotopos pesados.







UNIVERSIDAD AUTÓNOMA DEL
ESTADO DE MORELOS



INSTITUTO DE INVESTIGACIÓN EN CIENCIAS BÁSICAS Y APLICADAS

Coordinación de Programas Educativos

Posgrado en Ciencias

**DR. VÍCTOR BARBA LÓPEZ
COORDINADOR DEL POSGRADO EN CIENCIAS
PRESENTE**

Atendiendo a la solicitud para emitir DICTAMEN sobre la revisión de la TESIS titulada "Caracterización espectroscópica de la interacción proteína – ligando: AtLEA4-5 – iones metálicos, LAOBP – aminoácidos, CreCBD – ácidos nucleicos" que presenta la alumna **Leidys French Pacheco (5620140601)** para obtener el título de **Doctor en Ciencias**.

Nos permitimos informarle que nuestro voto es:

NOMBRE	DICTAMEN	FIRMA
Dr. Enrique Rudiño Piñera IBT-UNAM	Aprobado	
Dra. Carmen Nina Pastor Colón CIDC-UAEM	Aprobado	CARMEN PASTOR 4.11.19
Dra. Rosario Adelaida Muñoz Clares FQ-UNAM	APROBADO	Rosario Muñoz Clares 28/11/2019
Dr. Daniel Alejandro Fernández Velasco FM-UNAM	Aprobado	 3.12.2019
Dra. Claudia Diaz Camino IBT-UNAM	Aprobado	 04.11.2019
Dr. Carlos Daniel Amero Tello CIQ-UAEM	Aprobado	Carlos Amero Tello 4/11/19
Dra. Lina Andrea Rivillas Acevedo CIDC-UAEM	Aprobado	Lina Rivillas 07.11.19