

Editor: Editorial de la Universidad de Granada
Autor: Javier Murciano Calles
D.L.: GR 4529-2011
ISBN: 978-84-694-5727-6

UNIVERSIDAD DE GRANADA
FACULTAD DE CIENCIAS
DEPARTAMENTO DE QUÍMICA FÍSICA

Memoria presentada para aspirar al grado de Doctor en Química

Fdo: Javier Murciano Calles

Licenciado en Química por la Universidad de Granada

Granada, a 6 de mayo de 2011

Vº Bº
DIRECTORES DE LA TESIS

Fdo: **Dr. José C. Martínez Herrerías**
Profesor Titular
Departamento de Química Física
Universidad de Granada

Fdo: **Eva Sánchez Cobos**
Contratada del plan de
incorporación de doctores de la ugr
Departamento de Química Física
Universidad de Granada

Fdo: **Dr. Pedro Luis Mateo Alarcón**
Catedrático de Química Física
Departamento de Química Física
Universidad de Granada

Conócete.
Acéptate.
Supérate
(San Agustín)

a mis padres

AGRADECIMIENTOS

Por fin llega la hora de los agradecimientos, y digo por fin porque tenía decidido escribirlos lo más tarde posible y de un tirón, para que saliesen de dentro de verdad. También porque desde mis comienzos en el departamento, de las primeras cosas que leí fueron agradecimientos de distintas tesis, y mucha gente me ha comentado que son lo más leído... De ahí que los tuviese en la cabeza, sobre todo cuando alguien me ayudaba. Me parecía poco agradecerse directamente y pensaba que tenía que incluirlo en los agradecimientos de la tesis, para que quedara escrito y perdure en el tiempo.

Ante todo, quiero agradecer especialmente a mis tres directores de tesis el haber tenido la estupenda posibilidad de hacer la tesis con ellos. A Pedro, porque ha estado siempre amablemente atento y disponible cuando he tenido que acudir a él, no sólo como director de tesis, sino también como director de departamento y jefe de grupo, con la carga de trabajo que eso supone y lo que hace mi agradecimiento aún mayor.

A Jose le debo la tesis... creo que con esto podría bastar, pero supongo que no debía escribir esto porque él mismo me va a decir que no, que la tesis es mía, etc. Y sí, lleva razón, la tesis es mía, pero contiene todo su impulso y su experiencia investigadora. He de agradecerle muy de veras cómo se ha volcado y cómo me ha ido ayudando en todo para que fuese hacia delante. Algo similar pasa con Eva, que me ha atendido siempre como si únicamente tuviese mi tesis entre manos, y sé lo atareada que ha estado durante todo este tiempo, por lo que le estoy muy agradecido. Especialmente esos ratos en los que me tranquilizaba cuando yo me aceleraba con cualquier cosa...

Y al resto del departamento, también le agradezco el poder trabajar tan a gusto con ellos. Especialmente a Javier, que tanto se ha volcado, y lo sigue haciendo, con cada uno de los que estamos en el laboratorio. Toda su experiencia y saber nos ha dado mucha seguridad cuando teníamos algún problema. A Irene, porque, aunque ella tiene mil historias entre manos, ha demostrado una capacidad tremenda. A Quico y Ana, por ser buenos profesores y compañeros. A Antonio Parody, al que agradezco su interés. A Enrique, por ayudar siempre con los papeleos. A María del Mar, porque su conversación siempre me ha animado mucho! A Salva, por lo mismo: su interés siempre me ha hecho sentir muy bien, gracias.

A Ana Cámara, porque una parte de la tesis se debe a ella. Estoy muy agradecido por colaborar con ella y poder ser ayudado por alguien

con tanta experiencia. También a Gavi, por ayudarme en lo relacionado con la cristalización.

¡Un momento importante: a todos mis compañeros del laboratorio! Especialmente quería mencionar a Bertrand, porque ha estado tantas veces ayudando a resolver los problemas del laboratorio. A Adela, porque con ella empecé a trabajar y ha sido en parte un modelo para mí. A José Manuel, por su carácter siempre amable y dispuesto a ayudar. A Andrés, por el cable que me echó en varias ocasiones. A Lorena y José Luis, por ser compañeros tan estupendos, ahora especialmente a Lorena, por estos momentos finales en que me está atendiendo desde Oxford, donde tiene mucho trabajo. A Manu, por volcarse tanto en las cosas; su generosidad se agradece un montón. A Carles, por ayudar en los asuntos informáticos, tan tediosos muchas veces. A Sara, por ser una gran “jefa” del laboratorio. A Mari Ángeles, aunque haya estado tan perdida... A Ana Mari, por su amable conversación e interés, gracias de veras. A David, por sus “cosas”, que me divierten tanto... y cómo no, a *Francisco*, el gran Fran, al que aquí tengo poco espacio para agradecerle tanta ayuda, interés, colaboración, y amistad.

Gracias también al resto de la gente del departamento, profesores, y becarios del otro grupo: Ángel, Rocío, Asun, Mari Carmen, Inma, Álvaro, Héctor y Sisi, junto a las nuevas incorporaciones de los dos laboratorios, los cuales no he podido conocer mucho aún. Aquí quiero agradecer a Joan, que no sabía dónde situarlo... pero quería ponerlo, por poder haber trabajado y haberlo pasado bien con él.

Quería agradecer también a María José y Juande, de Microscopía, su amabilidad y buen hacer en el trabajo. Y a los de Masas, sobre todo a Juan, por lo bien que se han portado siempre conmigo.

Requieren mención aparte toda la gente de Toronto. Dev, Linda, Wei, Megan, Nick, Raffi, Haiming, Nicolas, Ysabel, y sobre todo Andreas y Nishant, con los que he tenido la oportunidad de trabajar codo con codo, y sobre todo de disfrutar de su amistad. Además quiero agradecer a toda la gente de Ernescliff College, la residencia donde he vivido todas las estancias, que me han hecho sentir muy en casa estando a seis mil kilómetros de ella... No puedo dejar de mencionar a Anthony, Salva, Chuma, Arvin, Angelo, Peter, Joe, Fr. Cris, Fr. Greg, Carlo, Nico, Gabe, Raf, Brix y tantos otros que me dejo, *thanks so much!*

Y ya que hemos empezado con las amistades, quería agradecer a mis amigos de Úbeda, que siempre han estado ahí, aunque nos hayamos podido ver tan poco en esta época de la tesis. A Alberto, Francis, Fran,

Vidal, David, Gonzalo, muchas gracias! Y a la gente de la carrera: Juan y Manolo, a los que estoy agradecido por poder seguir en contacto con ellos, y tantos otros, con los que ya no mantengo tanto trato porque cada uno va siguiendo su camino.

Y terminando ya, a tanta, tanta gente del Colegio Mayor Albayzín, donde no sólo tuve la feliz oportunidad de vivir durante gran parte de la carrera, sino porque después he podido hacer la tesis enteramente allí. Ellos, tanta gente, han hecho que esto de la tesis fuese un camino amable en muchas ocasiones. Quiero agradecer muy especialmente a Santi, por haberme aguantado... tanto. Y a Chema, Yompi, Quique, Álvaro, Flópez, Ignacio (gracias por tus gestiones del photoshop!!!), Rafa y Javier, y más gente aún, por haber colaborado mucho tiempo en buena dirección... Y cómo no, quería hacer un agradecimiento concreto a algunos con los que he aprendido qué es la amistad, de verdad. Pienso que nadie tuvo unos amigos como los míos. A Juanmi, por su inquebrantable lealtad, él ya sabe lo que valoro esa virtud y a él mismo, y a Alberto, porque siempre ha estado ahí, animando, aguantándome y con el que he tenido la gran dicha de hablar tanto, tanto. Además ha sido el compañero de fatigas a la hora de escribir en la sala de estudio los últimos meses, lo cual me ha hecho mucho más grato todo. A Barto, porque siempre ha sabido cómo estaba con mirarme a la cara, y me ha ayudado cuando lo he necesitado. A Berto, porque su buen humor me alegra un montón, además de que ha colaborado profesionalmente con una figura. Al gran Mariano, que me hace descansar tanto siempre. También quería agradecer a Alvarito, Juanfran, Cebox, Santi, Migue y Quique (gracias por tus ánimos en la recta final), porque su amistad la valoro mucho. También Gonzalo, del que no se me olvidarán los ánimos que me dio cuando me fui a defender el DEA en Sevilla. Sé que me dejo mucha gente en el tintero a la que querría poner, a los que quiero disculparme por no mencionarlos expresamente, ellos saben que también los valoro un montón!

Finalmente, a mi familia, a todos mis tíos y primos, los cuales siempre me han preguntado y animado mucho. Se lo agradezco de veras. Especialmente a mi tía Filo, que supone un apoyo continuo que, por otro lado, poca gente sabe valorar, además de que le tengo un cariño inmenso, y quiero que así quede reflejado en esta tesis. A mi abuela, a la que le debo tanto, tanto (igual que a la otra, que ya se nos fue). A mi hermano José María, porque siempre será mi hermano mayor, con lo que eso supone de referencia y aprecio. También a mis tres hermanas, Inmaculada, Mari Carmen y Mari Ángeles, a las que quiero muchísimo, y a

las que el agradecimiento es vital, como a toda mi familia. Y por último, a mis padres. A ellos está dedicada la tesis, porque si a alguien tengo que agradecer todo o, al menos, la mayor parte de lo que tengo o soy en mi vida, es a ellos. Han sido un modelo para mí de lealtad, abnegación, volcarse en los demás, y un etcétera que se puede resumir en que, como tanta gente piensa, pero en mi caso tengo un convencimiento íntimo, personal, que mis padres son los mejores del mundo. Quiero agradecerles de manera particular su apoyo y sus ánimos en toda la tesis, y muy concretamente en la recta final, cuando el cansancio arreciaba. He escuchado varias veces que las guerras las ganan los soldados cansados con el aliento de sus capitanes, y yo, a Dios gracias, no me han faltado muchos capitanes. Gracias.

ÍNDICE

CAPÍTULO 1: INTRODUCCIÓN	1
CAPÍTULO 2: ANTECEDENTES BIBLIOGRÁFICOS	9
2.1 Introducción	11
2.2 Los dominios PDZ	11
2.2.1 Características estructurales y de unión de los dominios PDZ	12
2.3 La proteína PSD95	17
2.3.1 Características de los dominios modulares que componen la proteína PSD95	19
2.3.1.1 Los dominios SH3 y GK de la proteína PSD95	19
2.3.1.2 Los dominios PDZ1 y PDZ2 de la proteína PSD95	20
2.3.1.3 El dominio PDZ3 de la proteína PSD95	20
2.4 Estudios de plegamiento del dominio PDZ3-PSD95 y de otros dominios PDZ	21
2.5 La proteína Erbin	24
2.5.1 Características de la especificidad de unión del dominio PDZ-Erbin	25
CAPÍTULO 3: MATERIALES Y MÉTODOS	29
3.1 Clonación de las dos construcciones del tercer dominio PDZ de la proteína PSD95	31
3.1.1 Clonación de los dominios PDZ3-PSD95 y PDZ3-h-PSD95	31
3.1.2 Ensayos de expresión	32
3.2 Expresión y purificación de las proteínas	33
3.2.1 Expresión y purificación de las dos construcciones del dominio PDZ3 de la proteína PSD95	33
3.2.2 Expresión y purificación de los péptidos unidos a la proteína S-Transferasa de Glutación (GST)	33
3.3 Preparación de tampones y de disoluciones de proteína	34
3.4 Calorimetría Diferencial de Barrido (DSC)	34
3.4.1 Cálculo de la capacidad calorífica molar parcial de la proteína	35
3.4.2 Análisis de las trazas calorimétricas según la termodinámica de equilibrio.	35
3.4.2.1 Modelo de equilibrio de dos estados	36
3.4.2.2 Modelo de equilibrio de tres estados con asociación del intermedio	38

3.4.2.3 Modelo de equilibrio de cuatro estados con asociación del intermedio	42
3.5 Dicroísmo Circular (CD)	45
3.5.1 Dicroísmo circular en la zona del UV-lejano (170-250 nm)	46
3.5.2 Obtención de espectros de dicroísmo circular	48
3.5.3 Análisis de las curvas de desplegamiento térmico según la termodinámica del equilibrio	49
3.6 Dispersión dinámica de la luz (DLS)	50
3.7 Absorbancia de Congo Red (CR)	51
3.8 Fluorescencia de Thioflavina T y ANS	52
3.8.1 Ensayos con ThT	53
3.8.2 Ensayos con ANS	53
3.9 Microscopía de electrónica transmisión (TEM)	54
3.10 Cromatografía analítica de exclusión molecular (SEC)	55
3.11 Calorimetría isotérmica de titulación (ITC)	57
3.12 Cristalización de Proteínas	58
3.13 La metodología de expresión en fagos (<i>Phage Display</i>)	60
3.13.1 El bacteriófago M13	61
3.13.2 Co-expresión de proteínas y péptidos fusionados a una proteína de la cápside	63
3.13.3 Procedimiento experimental	64
3.13.3.1 Mutagénesis de Kunkel	64
3.13.3.2 Proceso de selección	68
3.13.3.3 Ensayo inmunoabsorbente ligado a enzimas (ELISA)	70
3.13.3.4 Secuenciación del ADN y su análisis	72

CAPÍTULO 4: ANÁLISIS ESTRUCTURAL POR DIFRACCIÓN DE RAYOS X DEL TERCER DOMINIO PDZ DE LA PROTEÍNA PSD95

4.1 Resultados	75
4.1.1 Cristalización del dominio PDZ3-PSD95	75
4.1.2 Resolución de la estructura del dominio PDZ3-PSD95 mediante difracción de rayos X	76
4.1.3 Análisis comparativo de las diferentes estructuras del dominio PDZ3-PSD95	77
4.2 Discusión	79
4.2.1 Comparación de las estructuras del dominio PDZ3-PSD95 con las estructuras previamente publicadas	79

4.2.2 Análisis de las consecuencias estructurales de la presencia de una succinimida en el resto aspartato de la posición 332	81
---	----

CAPÍTULO 5: ANÁLISIS TERMODINÁMICO Y CINÉTICO DEL EQUILIBRIO CONFORMACIONAL DEL TERCER DOMINIO PDZ DE LA PROTEÍNA PSD95 EN CONDICIONES DE pH NEUTRO	87
--	-----------

5.1 Resultados	89
5.1.1 Mecanismo de desplegamiento térmico del dominio PDZ3-PSD95 a pH neutro	89
5.1.2 Estudio de la formación reversible de agregados globulares y fibrilares a pH neutro a partir del estado intermedio	93
5.1.3 Aspectos conformacionales y energéticos del intermedio oligomérico obtenidos por el análisis del equilibrio del desplegamiento del dominio PDZ3-PSD95	98
5.1.4 Análisis espectroscópico y cinético de los estados conformacionales iniciales del mecanismo de fibrilación del dominio PDZ3-PSD95	105
5.2 Discusión	109
5.2.1 Aspectos termodinámicos y conformacionales de la formación de fibras durante el desplegamiento del dominio PDZ3-PSD95	110
5.2.2 El esquema de desplegamiento del dominio PDZ3-PSD95 como posible paradigma para el estudio biofísico de interconversión entre estructuras macromoleculares	113

CAPÍTULO 6: DETERMINANTES ELECTROSTÁTICOS DEL EQUILIBRIO CONFORMACIONAL DEL TERCER DOMINIO PDZ DE LA PROTEÍNA PSD95	117
--	------------

6.1 Resultados	119
6.1.1 Estudio del desplegamiento térmico del dominio PDZ3-PSD95 a valores de pH ácido	119
6.1.2 Evaluación de las contribuciones electrostáticas de los residuos ácidos a los parámetros termodinámicos de desplegamiento	121
6.1.3 Mecanismo de formación de fibras del dominio PDZ3-PSD95 a pH ácido	125
6.1.4 Análisis cinético del crecimiento de las fibras a partir del dominio PDZ3-PSD95 a pH ácido	127

6.2	Discusión	131
6.2.1	Análisis termodinámico de la irreversibilidad de los procesos de fibrilación del dominio PDZ3-PSD95 a pH ácido	132
6.2.2	Influencia de las fuerzas electrostáticas como responsables del comportamiento diferencial en el plegamiento del dominio PDZ3-PSD95 a pH neutro y ácido	132
6.2.3	Análisis de las posiciones de los aminoácidos aspartato y glutamato en la estructura del dominio PDZ3-PSD95 que pueden ser responsables de la desestabilización energética del intermedio a pH ácido	133
6.2.4	Racionalización del mecanismo de desplegamiento térmico del dominio PDZ3-PSD95	134

CAPÍTULO 7: ANÁLISIS TERMODINÁMICO Y CINÉTICO DEL EQUILIBRIO CONFORMACIONAL DEL TERCER DOMINIO PDZ DE LA PROTEÍNA PSD95 SIN SU TERCERA HÉLICE α **137**

7.1	Resultados	139
7.1.1	Estudio de la interacción del dominio PDZ3-PSD95 y de su construcción sin la hélice α 3 con el péptido de secuencia KKETAV mediante ITC	140
7.1.2	Mecanismo de desplegamiento térmico del dominio PDZ3-h-PSD95 a pH neutro	142
7.1.3	Mecanismo de fibrilación del dominio PDZ3-h-PSD95	152
7.1.4	Cinética de la formación de fibras en el dominio PDZ3-h-PSD95	155
7.1.5	Influencia del pH en el intermedio de plegamiento del dominio PDZ3-h-PSD95	159
7.2	Discusión	165
7.2.1	Causas estructurales de la estabilización del intermedio de plegamiento del dominio PDZ3-PSD95 cuando se elimina su tercera hélice α	165
7.2.2	Relevancia biológica de la presencia de la tercera hélice α del dominio PDZ3-PSD95	166

CAPÍTULO 8: ANÁLISIS DE LA ESPECIFICIDAD DE UNIÓN DE LOS DOMINIOS PDZ MEDIANTE LA METODOLOGÍA DE EXPRESIÓN EN FAGOS **169**

8.1	Resultados	171
-----	------------	-----

8.1.1 Construcción de la librería del dominio PDZ-Erbin para el estudio de la especificidad relativa a la posición 0 del ligando	174
8.1.2 Construcción de las moléculas diana para el estudio de la especificidad relativa a la posición 0 del ligando	177
8.1.3 Selección de las variantes del dominio PDZ-Erbin más afines para cada uno de los veinte aminoácidos naturales presentes en el extremo C-terminal del ligando	178
8.1.3.1 Selección de las variantes del dominio PDZ-Erbin más específicas mediante una etapa de competición	182
8.1.3.2 Secuenciación de las variantes del dominio PDZ-Erbin seleccionadas	186
8.1.4 Selección de variantes del dominio PDZ-Erbin que unen secuencias peptídicas en un reconocimiento de tipo interno	192
8.2 Discusión	194
CAPÍTULO 9: RESUMEN Y CONCLUSIONES	199
CHAPTER 9: SUMMARY AND CONCLUSIONS	205
BIBLIOGRAFÍA	211
APÉNDICE	229

ABREVIATURAS

AMPA	Ácido α -amino-3-hidroxi-5-metil-4-isoxazolopropiónico
ANS	Ácido 8-Anilino-naftalén-sulfónico (ANS)
ASA	Área superficial accesible al disolvente
B(t)	Campo magnético propio de la radiación electromagnética
CD	<i>Circular Dichroism</i> ; Dicroísmo Circular
C_{p,s}	Capacidad calorífica parcial del disolvente
C_{pN}	Capacidad calorífica del estado N
C_{pD}	Capacidad calorífica del estado D
C_{pI}	Capacidad calorífica del estado I
C_{pIn}	Capacidad calorífica del estado I _n
CR	Congo Red
ITC	<i>Isothermal Titration Calorimetry</i> ; calorimetría isotérmica de titulación
D	Estado desplegado de la proteína
Da	Dalton
DLS	<i>Dynamic Light Scattering</i> ; dispersión dinámica de la luz
DNA	Desoxyribonucleic acid; Ácido desoxirribonucleico
DSC	<i>Differential Scanning Calorimetry</i> ; Calorimetría Diferencial de Barrido
ELISA	<i>Enzyme-Linked ImmunoSorbent Assay</i> ; ensayo inmunosorbente ligado a enzimas
Erbin	ERBB/HERR2 Interacting (protein); (proteína) que interacciona con el receptor ERBB/HERR2
f_N	Población del estado nativo
f_D	Población del estado desnaturalizado
f_I	Población del estado intermedio
f_{In}	Población del estado intermedio asociado
GK	Guanylate Kinase; kinasa de guanilato
GST	Glutathione S-Transferase, proteína S-Transferasa de Glutatión
i(t)	Flujo de fotones
I	Estado intermedio de la proteína
I_n	Estado intermedio asociado de la proteína
IPTG	Isopropil β -D-tiogalactapiranosido
k_B	Constante de Boltzmann
K_d	Constante de equilibrio de disociación del complejo proteína-ligando
K_{N-D}	Constante de equilibrio entre los estados N y D

ABREVIATURAS

K_{N-I}	Constante de equilibrio entre los estados N e I
K_{I-D}	Constante de equilibrio entre los estados I y D
K_{D-I_n}	Constante de equilibrio entre los estados D e I _n
K_{I-I_n}	Constante de equilibrio entre los estados I e I _n
LB	Medio de cultivo bacteriano Laura Bertani
LRR	<i>Leucine-Rich Repeat</i> ; repetición rica en leucina
MAGUK	<i>Membrane-Associated Guanylate Kinase</i> ; kinasa de guanilato asociada a membrana
n	Estequiometría de asociación del estado I _n
nNOS	<i>Neuronal Nitric Oxide Synthase</i> ; sintasa neuronal de óxido nítrico
N	Estado nativo de la proteína
NMDA	N-metil-D-aspartato
P	Concentración total de proteína
PDB	<i>Protein Data Bank</i> ; banco de datos de proteínas
PEG	Polietilenglicol
PBS	<i>Phosphate buffered saline</i> ; tampón salino de fosfato
PCR	<i>Polymerase Chain Reaction</i> ; Reacción en Cadena de la Polimerasa
PDZ	<i>Post synaptic density protein – Discs Large-4 – Zonula Occludens1</i>
PDZ2-PTP-BL	Dominio PDZ2 de la proteína PTP-BL
PDZ-Erbin	Dominio PDZ de la proteína Erbin
PDZ3-PSD95	Tercer dominio PDZ de la proteína PSD95
PDZ3-h-PSD95	Tercer dominio PDZ de la proteína PSD95 sin la hélice α3
pK_a	Menos logaritmo de la constante de acidez
PM	Peso molecular de la proteína
PSD95	<i>Post synaptic density protein</i> (95 kDa), proteína de la densidad post-sináptica
PTP-BL	Protein Tyrosine Phosphatase-Bas Like, proteína fosfatasa de tirosina tipo Bas
PWM	<i>Position Weight Matrix</i> ; matriz de peso de cada posición
q	Función de partición
R_H	Radio hidrodinámico
RMSD	<i>Root Mean Square Deviation</i> ; desviación cuadrática media
RNA	Ribonucleic acid; Ácido ribonucleico
SDS	Lauril sulfato sódico

SDS-PAGE	Electroforesis en gel de poliacrilamida en presencia de lauril sulfato sódico
SEC	<i>Size Exclusion Chromatography</i> ; cromatografía de exclusión molecular
SH3	Src Homology-3; (dominio) homólogo a Src
SH3-GK-PSD95	Dominios SH3 y GK de la proteína PSD95
SNN	Succinimida
SP	<i>Specificity Potential</i> ; potencial de especificidad
T_{N-D}	Temperatura de transición entre el estado N y D
T_{N-I}	Temperatura de transición entre el estado N e I
T_{I-D}	Temperatura de transición entre el estado I y D
T_{D-I_n}	Temperatura de transición entre el estado D e I _n
T_{I-I_n}	Temperatura de transición entre el estado I e I _n
TEM	<i>Transmission Electron Microscopy</i> ; microscopía electrónica de transmisión
ThT	Tioflavina T
TMB	3,3,5,5'-tetrametilbenzidina
UV	Ultravioleta
V_p	Volumen específico parcial de la proteína
V_s	Volumen específico parcial del disolvente
V_R	Volumen de retención
V₀	Volumen de exclusión
ZO-1	Proteína <i>Zonula Occludens 1</i>
ΔAbs (λ)	Diferencia de absorbancia medida en CD
ΔC_{p,ap}	Capacidad calorífica aparente entre la disolución de macromolécula y el disolvente
ΔC_{pN-D}	Cambio de capacidad calorífica entre el estado N y D
ΔC_{pN-I}	Cambio de capacidad calorífica entre el estado N y I
ΔC_{pI-D}	Cambio de capacidad calorífica entre el estado I y D
ΔC_{pD-I_n}	Cambio de capacidad calorífica entre el estado D e I _n
ΔC_{pI-I_n}	Cambio de capacidad calorífica entre el estado I e I _n
ΔG_{N-D}	Cambio de energía de Gibbs entre el estado N y D
ΔG_{N-I}	Cambio de energía de Gibbs entre el estado N y I
ΔG_{I-D}	Cambio de energía de Gibbs entre el estado I y D

ΔG_{D-I_n}	Cambio de energía de Gibbs entre el estado D e I_n
ΔG_{I-I_n}	Cambio de energía de Gibbs entre el estado I e I_n
ΔH_d	Cambio de entalpía de unión
ΔH_{N-D}	Cambio de entalpía entre el estado N y D
ΔH_{N-I}	Cambio de entalpía entre el estado N y I
ΔH_{I-D}	Cambio de entalpía entre el estado I y D
ΔH_{D-I_n}	Cambio de entalpía entre el estado D e I_n
ΔH_{I-I_n}	Cambio de entalpía entre el estado I e I_n
$\Delta H_{N-D}(T_{N-D})$	Cambio de entalpía entre el estado N y D a la T_{N-D}
$\Delta H_{N-I}(T_{N-I})$	Cambio de entalpía entre el estado N y I a la T_{N-I}
$\Delta H_{I-D}(T_{I-D})$	Cambio de entalpía entre el estado I y D a la T_{I-D}
$\Delta H_{D-I_n}(T_{D-I_n})$	Cambio de entalpía entre el estado D e I_n a la T_{D-I_n}
$\Delta H_{I-I_n}(T_{I-I_n})$	Cambio de entalpía entre el estado I e I_n a la T_{I-I_n}
ΔS_{N-D}	Cambio de entropía entre el estado N y D
ΔS_{N-I}	Cambio de entropía entre el estado N y I
ΔS_{I-D}	Cambio de entropía entre el estado I y D
ΔS_{D-I_n}	Cambio de entropía entre el estado D e I_n
ΔS_{I-I_n}	Cambio de entropía entre el estado I e I_n
$\Delta S_{N-D}(T_{N-D})$	Cambio de entropía entre el estado N y D a la T_{N-D}
$\Delta S_{N-I}(T_{N-I})$	Cambio de entropía entre el estado N y I a la T_{N-I}
$\Delta S_{I-D}(T_{I-D})$	Cambio de entropía entre el estado I y D a la T_{I-D}
$\Delta S_{D-I_n}(T_{D-I_n})$	Cambio de entropía entre el estado D e I_n a la T_{D-I_n}
$\Delta S_{I-I_n}(T_{I-I_n})$	Cambio de entropía entre el estado I e I_n a la T_{I-I_n}
θ	Elipticidad
ϵ_D	Coefficiente de extinción molar para la luz polarizada a la derecha
ϵ_I	Coefficiente de extinción molar para la luz polarizada a la derecha
$[\theta]_{MWR}$	Elipticidad molar media por residuo
$[\theta]_N$	Elipticidad molar media por residuo del estado N
$[\theta]_D$	Elipticidad molar media por residuo del estado D
η	Viscosidad

CAPÍTULO 1

INTRODUCCIÓN

1 Introducción

En los sistemas biológicos, los procesos como el crecimiento, la obtención de energía, la división celular o la señalización son consecuencia de un conjunto de interacciones entre numerosas biomoléculas presentes en la célula que están interconectadas unas con otras formando redes entre sí. De esta forma, entender los mecanismos biológicos de estos procesos, que se realizan mediante lo que en la literatura se han denominado “módulos funcionales” (*functional modules*, en inglés) (Hartwell et al., 1999), así como las enfermedades derivadas de estos, requieren una aproximación global, que, yendo más allá del punto de vista reduccionista que estudia interacciones individuales, enfoca estos procesos como análisis globales de la estructura, función y dinámica de estas redes en su conjunto (Cusick et al., 2005).

Abordar el estudio de estas redes de interacción o interactomas no ha de entenderse únicamente como un proyecto de análisis post-genómico, puesto que aunque las secuencias completas de los genomas proporcionan la posibilidad de predecir decenas de miles de proteínas diferentes (unas 25000 para el proteoma humano (Human Genome Sequencing, 2004)), estas secuencias no ofrecen por sí mismas un conocimiento de los mecanismos básicos de los sistemas celulares. Se requiere también información a escala global del interactoma tanto a nivel estructural como funcional y dinámico. Los primeros avances en esta línea de investigación han consistido en desarrollar mapas de interacciones en los que las proteínas, moléculas de RNA, DNA, metabolitos, etc, se representan mediante nodos que están conectados por líneas correspondientes a las interacciones moleculares (interacciones proteína-proteína, DNA-proteína, reacciones enzimáticas, etc). A modo de ejemplo, se puede observar el interactoma de levadura *Saccharomyces cerevisiae* en la Figura 1.1.

Estas redes biológicas, al igual que otro tipo de redes relacionadas con la comunicación como Internet y la telefonía móvil (Goh et al., 2002), son aquellas que se denominan libres de escala (*scale free networks*, en inglés) (Barabasi & Albert, 1999; Jeong et al., 2000). La característica principal de estas redes es la presencia de grandes ejes o centros (*hub*, en inglés). Este tipo de ejes son nodos que tienen cientos, miles o incluso millones de uniones. Existen más tipos de redes de conexión, como las denominadas redes regulares (*regular networks*, en inglés), en las que las interacciones forman una malla con una conectividad que no varía entre los nodos, o también las llamadas redes aleatorias (*random networks*, en inglés), en las que la mayoría de los nodos tienen aproximadamente la misma conectividad, puesto que al estar conectados de forma puramente aleatoria estadísticamente las conexiones se reparten de forma equitativa.

La existencia de *hubs* en redes que presentan muchas conexiones ofrece dos ventajas principales con respecto a los otros tipos de redes. Por un lado, tienen mucha resistencia respecto a fallos de naturaleza aleatoria. Así, cuando en una red de tipo regular falla una conexión aleatoriamente, el resto de las conexiones se mantienen formando subredes que no pueden comunicarse más entre sí. En cambio, la presencia de nodos altamente conectados provoca que se mantenga la integridad de la red, puesto que es

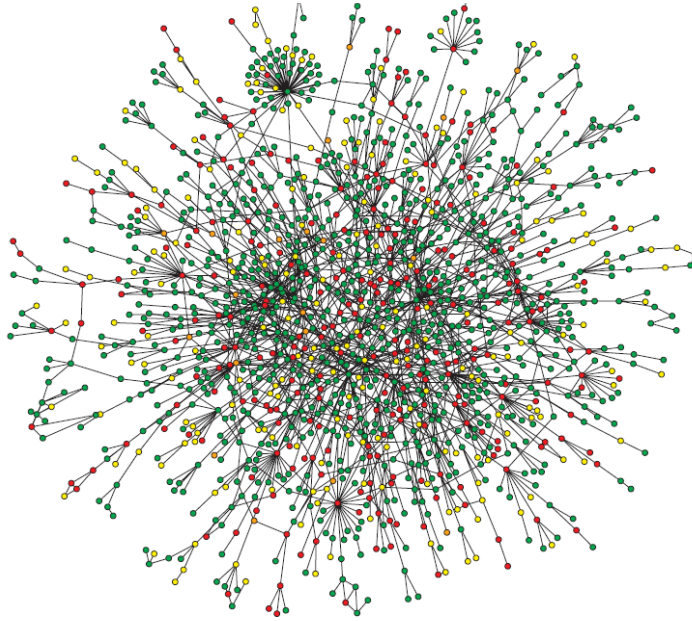


Figura 1.1. *Interactoma o mapa de interacción proteína-proteína de la levadura Saccharomyces cerevisiae (figura tomada de (Dunker et al., 2005)).*

estadísticamente menos probable que se afecten estos nodos, por estar en menor número. Ahora bien, mientras que esta propiedad es una ventaja evidente frente a fallos de tipo aleatorio, estas redes son extremadamente sensibles cuando el fallo no es de esta naturaleza, sino que esté provocado con el fin concreto de eliminar el nodo altamente conectado. Esto causaría el colapso de la conectividad, con la parada consecuente de la comunicación y de los procesos en cuestión. Por otro lado, otra característica que presentan este tipo de redes es que hay distancias más cortas entre dos puntos cualesquiera, en términos de número de nodos que intervienen entre esos dos puntos. Esto confiere a la red la capacidad de que para que una señal viaje desde cualquier punto hasta otro, lo haga a través de un número pequeño y mínimo de conexiones (Watts & Strogatz, 1998).

Abordando ya específicamente las redes de interacción proteína-proteína, y de acuerdo con lo que se ha descrito sobre este tipo de redes, las proteínas *hub* ejercen un papel muy central en los múltiples procesos en los que están implicadas, de tal forma que sin su actuación la supervivencia de los organismos se hace extremadamente difícil (Jeong et al., 2001; Hoffmann & Valencia, 2003). Estas proteínas son elementos centrales en torno a los cuales se organiza todo el proteoma celular, conectando los procesos biológicos entre sí física y temporalmente, o incluso constituyéndose ellas mismas como moldes o soportes con los que coordinar los procesos celulares más complejos (Hartwell et al., 1999). En esta función se han distinguido dos tipos básicos de proteínas *hub*: por un lado, aquellas que interactúan con la mayoría de sus moléculas diana de forma simultánea y por otro las que unen sus distintas moléculas diana en momentos y/o localizaciones diferentes. Los

estudios realizados hasta la fecha han mostrado que aquellas que interactúan simultáneamente con sus diferentes moléculas diana conectan entre sí los distintos procesos biológicos o módulos funcionales, mientras que las proteínas *hub* que las unen en distintas localizaciones o momentos ejercen su función dentro de estos módulos (Han et al., 2004).

Desde un punto de vista funcional, lo más notorio y característico de las proteínas *hub* es que no suelen presentar actividades enzimáticas propias y están organizadas estructuralmente en diferentes dominios o módulos, conformacionalmente independientes, e interconectados por secuencias de aminoácidos relativamente cortas. Dentro del inmenso panorama de la proteómica, se ha hecho cada vez más evidente la existencia de estos dominios que están presentes en proteínas que no tienen relación entre sí (Pawson & Nash, 2003). Así, una representación más informativa de un interactoma sería aquella en la que cada nodo del gráfico esté fragmentado en varios nodos, representando cada uno de estos un dominio de la proteína (Santonico et al., 2005). Estos dominios modulares están presentes en la proteína de tal forma que dejan accesible la superficie donde se produce la interacción o unión a su molécula diana. Suelen reconocer regiones específicas de otras proteínas, aunque también hay algunos que unen los grupos cargados de los fosfolípidos que constituyen las membranas celulares. Cuando interactúan con otras proteínas, los dominios modulares tienen la capacidad de reconocer un elemento distintivo propio que caracteriza a ese tipo concreto de dominios como pueden ser motivos fosforilados, ricos en prolina o secuencias C-terminales. Además, puede haber otros residuos alrededor de ese motivo característico, o también alejados de éste, que proporcionen contactos adicionales que contribuyan a la especificidad de unión del dominio en cuestión.

Dado que al fin y al cabo son las proteínas concretas con los dominios que las componen y sus secuencias de conexión las que ejercen su función dentro de la célula, aparte de desarrollar estrategias de alto rendimiento (*high throughput*) que permitan generar una descripción completa y cuantitativa de los diferentes interactomas, se hace necesario también obtener información a nivel molecular de las propiedades conformacionales y características de unión de las proteínas que componen el interactoma, especialmente de estas proteínas *hub*. De hecho, a pesar de los notables avances obtenidos en el análisis exhaustivo de los interactomas mediante diferentes métodos experimentales y computacionales de alto rendimiento, no siempre se han obtenido resultados concluyentes o suficientemente contrastados (Sprinzak et al., 2003). Todo esto nos conduce a que la solución sería desarrollar marcos en los que integrar la información genómica y proteómica a gran escala con los resultados de aproximaciones de la biología molecular más tradicional, que permitan una comprensión profunda y contrastada de los determinantes de la afinidad y especificidad de los distintos interactomas (Kelly & Stumpf, 2008; Schmid & McMahon, 2007).

La elaboración de esta Tesis Doctoral se ha desarrollado dentro de este contexto. Para ello, se han elegido unos dominios modulares que forman parte de muchas proteínas *hub*, los dominios PDZ. Estos dominios se caracterizan por el reconocimiento de

secuencias de aminoácidos en el extremo C-terminal de sus proteínas diana, lo cual es especialmente conveniente para formar parte de proteínas *hub*, ya que la flexibilidad de estas regiones, que generalmente se encuentran desestructuradas, permite obtener elevados grados de especificidad de unión con secuencias de pocos restos aminoacídicos, dejando además el resto de la proteína diana libre para la realización de su función biológica. Sin embargo, no es ésta la única forma en que los dominios PDZ pueden contribuir a la organización de actividades enzimáticas. En los pocos estudios realizados hasta la fecha se han descrito una serie de aspectos conformacionales en estos dominios que, adecuadamente interpretados en el contexto biológico, los podrían definir como elementos adaptadores y localizadores de funciones celulares (Nourry et al., 2003). Estos estudios describen cómo ciertas propiedades conformacionales pueden ser aprovechadas por la naturaleza para desarrollar sus propios acontecimientos, de ahí su interés, ya que serían los mecanismos por los cuales confieren a las proteínas que los contienen su enorme plasticidad funcional y conformacional. Esta gran plasticidad puede ser especialmente necesaria en procesos particularmente complejos, tales como el crecimiento celular y su división, en los que los dominios PDZ parecen estar muy directamente involucrados.

El interés que suscita el entendimiento de las causas de esta plasticidad funcional y conformacional es lo que nos ha llevado a realizar en esta Tesis Doctoral el primer estudio termodinámico del equilibrio conformacional de un dominio PDZ, en concreto del tercer dominio de la proteína PSD95 (PDZ3-PSD95), una proteína tipo *hub* que es un elemento central en la organización de las sinapsis neuronales. Esta proteína, como se explicará en el Capítulo 2, contiene cinco dominios modulares, de los cuales los tres primeros son dominios PDZ y a continuación tiene un dominio SH3 y un dominio guanilato quinasa (Guanylate Kinase, GK). Es bastante sorprendente que sólo cinco dominios funcionales puedan ser responsables de la organización de una variedad muy amplia de proteínas sinápticas presentes en la membrana interna de la neurona. Hay que tener en cuenta que la plasticidad estructural de las proteínas *hub* ha de deberse a la propia plasticidad de sus dominios y secuencias de conexión entre ellos. Además, trabajos previos demuestran la interrelación existente entre la estructura y la estabilidad de una proteína y sus características funcionales (Luque & Freire, 2000; Luque et al., 2002; Jemth & Gianni, 2007), lo que nos ha conducido a plantearnos el estudio de las propiedades conformacionales del dominio PDZ3-PSD95. Así, se podrían esclarecer algunas de las bases termodinámicas que justifiquen la múltiple funcionalidad de la proteína *hub* PSD95.

Por otro lado, se ha continuado en esta Tesis Doctoral el estudio de los determinantes moleculares de la especificidad de unión de los dominios PDZ. Hasta la fecha, se ha estudiado la especificidad de los dominios PDZ mediante el uso de librerías peptídicas para obtener información sobre algunos dominios concretos con la esperanza incluso de que los análisis de la especificidad de estos dominios permitan establecer algunas reglas generales para predecir la función a partir de la estructura (Tonikian et al., 2008). Más aún, se pensaba que con suficiente información sobre varios dominios PDZ podría ser posible predecir la especificidad en base a la secuencia primaria y así, mediante

ingeniería de proteínas, diseñar nuevas especificidades mediante diseño computacional. De esta manera, en un orden de escala mayor, el conocimiento derivado de estos estudios podría ayudar a elucidar los principios fundamentales que gobiernan las interacciones proteína-proteína en el contexto de los dominios PDZ. Estos estudios, que se resumen en el Capítulo 2, han mostrado una gran diversidad funcional para los dominios PDZ y ha permitido el desarrollo de algunos algoritmos capaces de predecir especificidades para sitios de unión, siempre que sean similares a los de aquellos dominios PDZ estudiados. Sin embargo, este conjunto de datos no proporciona un conocimiento suficiente para poder predecir la relación entre la secuencia de dominios PDZ y su especificidad, ni tampoco ninguna regla general de interacciones proteína-proteína.

De esta forma, se ha comprobado que incluso el estudio del comportamiento de un conjunto grande de dominios PDZ naturales puede no ser suficiente como para completar el conocimiento de las relaciones entre la estructura de cualquier dominio PDZ y su función. La limitación fundamental de este tipo de estudios sería que incluso entre grandes familias de dominios con cientos de miembros, la diversidad natural representa únicamente un pequeño muestreo del conjunto posible de secuencias que es compatible con la estructura de un dominio. Según esto, en nuestros estudios hemos optado por una aproximación más general, que consiste en generar una librería de dominios PDZ que incluyan toda la variabilidad posible, sea o no natural, para conseguir dilucidar los aspectos más generales que interrelacionan la estructura, función y evolución de los dominios PDZ y así poder expandir y completar la base de datos pre-existente. Esto permitirá determinar y localizar nuevas interacciones naturales de interés, así como hacer nuevos descubrimientos sobre los sistemas de señalización en los que estos dominios son clave, amén del potencial que presenta esta información con fines terapéuticos.

Dentro de estos marcos generales, los objetivos específicos perseguidos en este trabajo son los siguientes:

- I. Obtener la estructura del dominio PDZ3-PSD95 para poder evaluar a nivel molecular sus propiedades conformacionales de este dominio.
- II. Analizar en profundidad el equilibrio conformacional del dominio PDZ3-PSD95.
- III. Establecer la influencia del pH como factor regulador del equilibrio conformacional del dominio PDZ3-PSD95.
- IV. Evaluar el papel que tiene la secuencia de conexión existente entre el dominio PDZ3-PSD95 y el módulo SH3-GK-PSD95 en sus propiedades conformacionales.
- V. Determinar a nivel molecular los residuos responsables de la especificidad de unión de los dominios PDZ mediante la metodología de expresión en fagos (*phage display*) usando como modelo el dominio PDZ de la proteína Erbin.

CAPÍTULO 2

ANTECEDENTES BIBLIOGRÁFICOS

2.1 Introducción

Esta Tesis Doctoral se centra en el estudio de unos dominios modulares que forman parte de muchas proteínas *hub*, los dominios PDZ, llegando a haber ejemplos de proteínas constituidas exclusivamente por este tipo de dominios modulares (Feng & Zhang, 2009). En el presente Capítulo se describen las características generales de estos dominios, centrando la atención en las dos proteínas cuyos dominios PDZ han sido objeto de estudio, la proteína PSD95 y la proteína Erbin y comentando los estudios más relevantes de dichos dominios realizados hasta la elaboración de esta Tesis Doctoral.

2.2 Los dominios PDZ

Los dominios PDZ deben su nombre a que fueron descubiertos en tres proteínas: *Post Synaptic Density Protein*, con peso molecular de 95 kDa (PSD95) (Cho et al., 1992), la proteína *Discs large* (Woods & Bryant, 1991), y la proteína *Zonula Occludens-1* (Itoh et al., 1993). Dentro del contexto general del papel de estos dominios como módulos presentes en proteínas *hub* que organizan el interactoma llevando complejos macromoleculares a compartimentos sub-celulares que permiten las diferentes cascadas de señalización, vale la pena destacar tres de las funciones biológicas más comunes en las que estos dominios participan. Por un lado, son adaptadores de los receptores de las enzimas tipo quinasa de tirosina (Pawson, 1995). Estas enzimas ejercen, mediante la fosforilación de un residuo tirosina, la función de la transducción de señales, que consiste en un conjunto de procesos que ocurren de forma concatenada mediante los cuales una célula convierte un estímulo exterior en una respuesta específica. Estos procesos se desencadenan mediante la activación hormonal de los receptores a los cuales están asociadas estas proteínas (Pawson, 1994).

Por otro lado, los dominios PDZ están presentes en las proteínas que se encargan de generar y mantener la polaridad epitelial, fenómeno que permite la formación de la mayoría de los tejidos y que genera barreras con cierta permeabilidad para permitir el transporte de sustancias entre dos compartimentos del cuerpo (Yeaman et al., 1999). Muchas proteínas que contienen dominios PDZ tienen un papel central en la formación y mantenimiento de estas estructuras (Nourry et al., 2003).

Finalmente, dentro del marco de la función de señalización celular de estos dominios, destaca su papel en la unión sináptica entre células del sistema nervioso (Craven & Bredt, 1998). La sinapsis es un ejemplo paradigmático de una unión celular entre compartimentos sub-celulares polarizados, consistentes en una zona pre-sináptica, que contiene las vesículas de los neurotransmisores y que está unida a una región post-sináptica, donde los receptores de los neurotransmisores activan numerosos mecanismos de señalización (Kennedy, 2000). Dentro del elevado número de proteínas que intervienen en estos procesos, multitud de éstas contienen dominios PDZ. El papel de estos dominios en esta red de interacciones, en la que una proteína *hub* puede interactuar con hasta más de 100 proteínas diferentes, permite establecer varios niveles de activación de los

distintos receptores y permitir así la modulación correcta de las señales, con la consecuencia que esto tiene en el aspecto cognitivo.

2.2.1 Características estructurales y de unión de los dominios PDZ

En el contexto de la variada actuación de estos dominios en los distintos procesos celulares, llama la atención su característica general de poseer una organización espacial común consistente en un barril β formado por seis o siete cadenas (enumeradas como $\beta 1$, $\beta 2$,...) y rodeado de dos hélices α , una bastante más corta que la otra (Figura 2.1). A pesar de la organización común, existe una gran variabilidad en secuencia entre los diferentes elementos de estructura secundaria, en la longitud y composición de los lazos y giros que los conectan, así como de sus extremos N- y C-terminales.

Otra característica muy común de los dominios PDZ es su capacidad de reconocer la secuencia del extremo C-terminal de sus proteínas diana. Para ello, estos dominios poseen un motivo altamente conservado en su secuencia, GLGF (Gly-Leu-Gly-Phe), que envuelve al grupo carboxilato del extremo C-terminal del ligando peptídico, cuya carga negativa es neutralizada a través de los grupos amida de la cadena principal de los últimos tres restos del motivo GLGF. Además, en la interacción, las secuencias C-terminales de sus proteínas diana se disponen a lo largo del bolsillo de unión, que es una hendidura formada por la cadena $\beta 2$ y la hélice $\alpha 2$, formando una cadena β adicional antiparalela a $\beta 2$ (Figura 2.1) en la que se presenta el mismo patrón de enlaces de hidrógeno que tiene este elemento de estructura secundaria (Harris et al., 2001). En esta interacción, la zona de unión del dominio PDZ puede interaccionar de forma específica con hasta siete residuos del extremo C-terminal de su proteína diana, permitiendo así el reconocimiento de un ligando entre varios que se diferencien en secuencia. De esta forma, la especificidad de estos dominios es la que determina su papel biológico determinado puesto que dentro de las posibles moléculas diana con las que puede interaccionar, lo hace con una en concreto.

La caracterización de algunas interacciones de ligandos con dominios PDZ permitió una clasificación inicial de estos dominios, atendiendo a su especificidad de unión, en tres clases diferentes en función de la naturaleza de los aminoácidos que reconocen estos dominios en las cuatro últimas posiciones del ligando (Tabla 2.1). Para designar las posiciones de los ligandos se adoptó como criterio asignar el número 0 al extremo C-terminal del ligando y a continuación enumerar negativamente hacia el extremo N-terminal las otras posiciones (Doyle et al., 1996). Según este criterio, por tanto, en la primera clase de especificidad en la posición -2 reconocen los aminoácidos serina o treonina, y en la posición 0 uno de naturaleza hidrofóbica. En la segunda clase de especificidad, en las posiciones -2 y 0 se reconocen aminoácidos de naturaleza hidrofóbica, mientras que en la tercera en la posición -2 se reconocen varios aminoácidos que poseen una cadena lateral con carga habiendo de nuevo en la posición 0 un aminoácido de naturaleza hidrofóbica.

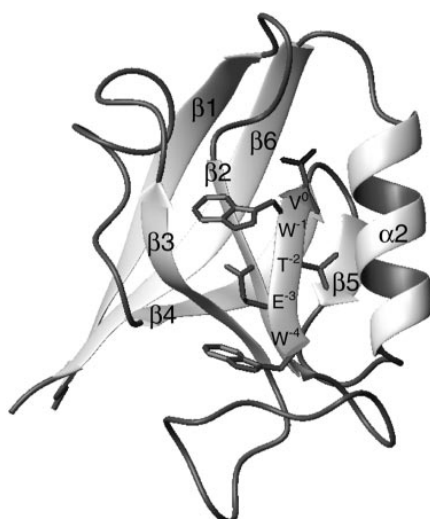


Figura 2.1. Estructura del dominio PDZ-Erbin unido al péptido TGWETWV con los elementos de estructura secundaria señalados. El péptido se representa como una lámina β situada entre la cadena $\beta 2$ y la hélice $\alpha 2$ del dominio (figura tomada de (Appleton et al., 2006)).

Tabla 2.1. Clasificación inicial de los dominios PDZ basada en la especificidad de unión.

	Secuencia consenso C-terminal	Ejemplo de dominio PDZ
Clase I	-X-(S/T)-X- Φ	PDZ-Erbin, PDZ3-PSD95
Clase II	-X- Φ -X- Φ	PDZ3-GRIPP1
Clase III	-X-(D/E/R/K)-X- Φ	PDZ-nNOS

X representa a cualquier aminoácido, Φ sólo hidrofóbico y ψ sólo aromático

Con el fin de comprender más profundamente los aspectos que regulan la especificidad de unión de estos dominios, en el grupo de investigación del Dr. Sachdev Sidhu, con el que nuestro equipo de investigación está colaborando actualmente, se utilizó la técnica de expresión en fagos para intentar establecer un mapa de especificidad más detallado de estos dominios (Tonikian et al., 2008). Mediante el cribado de más de 150 dominios PDZ presentes en la naturaleza frente a una librería de billones de péptidos aleatoriamente diferenciados en secuencia se seleccionaron las secuencias más afines. Del análisis de estas secuencias se concluyó que había que ampliar las clases de especificidad desde las tres iniciales hasta un total de 16 posibles clases de especificidad para el reconocimiento entre estos dominios y sus secuencias diana, cuyas diferencias se deben a su capacidad para reconocer diferentes aminoácidos principalmente en las 4

últimas posiciones del ligando peptídico. De acuerdo con el criterio de numeración mencionado, y según el mapa de especificidad citado, las posiciones del ligando que determinan la especificidad de unión son de esta forma la 0, -1, -2 y -3. Estos estudios los describiremos más a detalle en este mismo Capítulo.

Por otro lado, aunque menos estudiados, los dominios PDZ también pueden presentar interacciones de otra naturaleza, entre las que se encuentran la interacción con otros dominios PDZ (Brenman et al., 1996; Hillier et al., 1999), la unión a secuencias peptídicas que no corresponden al extremo C-terminal (Penkert et al., 2004; Zhang et al., 2009) e incluso a lípidos (Zimmermann et al., 2002; Wu et al., 2007; Feng et al., 2008; Gallardo et al., 2010; Krahn et al., 2010). En el caso de unión a lípidos, los cuales son de tipo fosfatidilinositol, se ha comprobado que la región del giro entre la hélice $\alpha 2$ y la lámina $\beta 6$ junto con el giro entre las láminas $\beta 1$ y $\beta 2$ es la que interacciona con los grupos polares de los estos lípidos (Figura 2.2). Además, en dichos estudios se ha observado que los dominios PDZ presentan poca especificidad frente a estos lípidos, tal y como ocurre con los dominios PH (Yu et al., 2004).

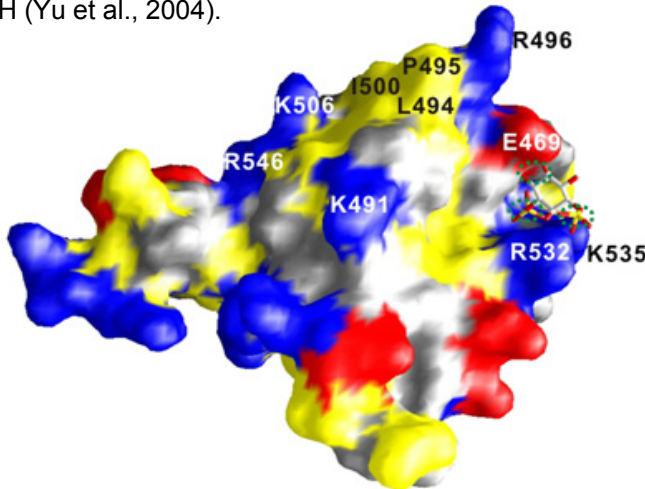


Figura 2.2. Estructura del dominio PDZ2 de la proteína Par-3 unido al grupo polar del lípido PI3P. La superficie del dominio PDZ está coloreada en función de la naturaleza de los residuos que lo componen: en azul están los que presentan carga positiva en su cadena lateral, en rojo los que tienen una carga negativa en su cadena lateral, en amarillo están los que presentan una cadena lateral de naturaleza hidrofóbica, mientras que el resto están coloreados en blanco. La cabeza polar del lípido PI3P está representada con todos sus átomos entre los residuos E469 (presente en el giro entre las cadenas $\beta 1$ y $\beta 2$) y R532 (presente en el giro entre la hélice $\alpha 2$ y la cadena $\beta 6$). Figura tomada de (Wu et al., 2007).

En cuanto a las interacciones que se podrían calificar de reconocimiento “interno” en contraposición al reconocimiento del extremo, se encuentra la unión entre los dos dominios PDZ de sintrofina y la sintasa neuronal de óxido nítrico (*nNOS*, siglas en inglés) (Brenman et al., 1996; Hillier et al., 1999; Oschkinat, 1999), en la que el dominio PDZ de sintrofina es capaz de interaccionar con la lámina $\beta 6$ del dominio PDZ de *nNOS* (Figura 2.3). La lámina $\beta 6$ del dominio PDZ-*nNOS* es un elemento “extra” de estructura secundaria específica de este dominio que mimetiza un ligando peptídico típico de dominios PDZ, con las interacciones propias de las distintas posiciones del ligando. Sin embargo, en vez de

ser un grupo carboxilato terminal el que interacciona con el bolsillo de unión, la conformación del giro β que hay entre las cadenas $\beta 6$ y $\beta 7$ del dominio PDZ-nNOS permite evitar las restricciones estéricas que impone el motivo GLGF y adaptarse a los requerimientos del bolsillo de unión (Oschkinat, 1999).

Por otro lado, hay dos ejemplos conocidos de reconocimiento interno en los que se produce la extensión de la interacción a un residuo más allá de la posición 0, pudiéndose denominar posición +1. En el caso de la unión del dominio PDZ de la proteína Par-6 con Pals-1 (Penkert et al., 2004), se observa que se produce una deformación del bolsillo hidrofóbico con respecto a la interacción típica con el extremo C-terminal del ligando (Figura 2.4). En la interacción, el grupo carboxilato de un aspartato interacciona con el giro entre las cadenas $\beta 1$ y $\beta 2$, que es la región que contiene el motivo GLGF y que coordina el grupo carboxilato del extremo C-terminal en el tipo de interacción más común de estos dominios, y luego la interacción se extiende a varios residuos más. La deformación del bolsillo hidrofóbico implica una plasticidad tal de estos dominios que permite ampliar su capacidad de reconocimiento de ligandos y de propiedades de unión. Esta plasticidad se ve también en el caso de la interacción de los dominios PDZ Dvl (Zhang et al., 2009), que pueden presentar ambos tipos de reconocimiento, tanto del extremo C-terminal con la interacción típica del grupo carboxilato, así como el reconocimiento de otro residuo en la posición +1, siempre que de nuevo exista en el péptido un residuo aspartato cuyo grupo carboxilato interactuare de la misma forma que lo hace normalmente el del extremo C-terminal.



Figura 2.3. Interacción entre los dominios PDZ de sintrofina (color verde) y nNOS (color morado). Para facilitar la visualización de la interacción, se ha coloreado en naranja los elementos estructurales del dominio PDZ nNOS que forman contactos con el dominio PDZ de sintrofina. Los residuos del bolsillo hidrofóbico del dominio PDZ de sintrofina están en color rojo.

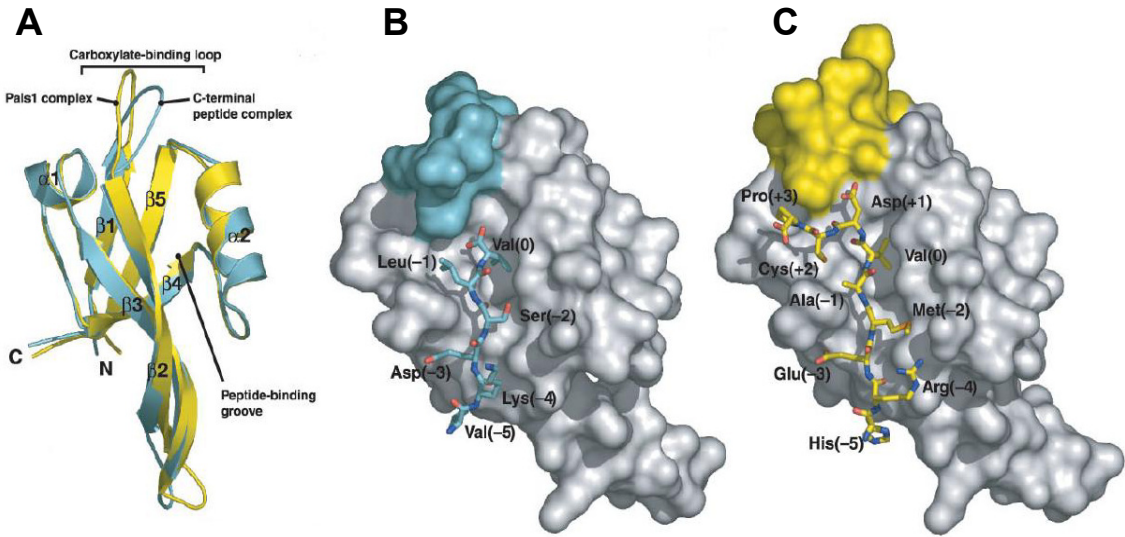


Figura 2.4. Representación de los dos tipos de interacción que presenta el dominio PDZ de Par-6. (A) Superposición de las estructuras del dominio cuando une un péptido a través de su extremo C-terminal (color azul) y cuando une a Pals-1 (color amarillo). (B) Representación del dominio PDZ de Par-6 unido al péptido VKDSL. (C) Representación del dominio PDZ de Par-6 unido al extremo C-terminal de Pals-1. La Figura está tomada de Penkert et al., 2004.

A pesar de que estas propiedades de unión de los dominios PDZ dan algo de versatilidad para su funcionalidad, hacen que parezca imposible imaginar que una proteína *hub* multimodular con unos pocos dominios proteicos tipo PDZ, junto con algunos otros de otra naturaleza, que unirán con una cierta especificidad determinadas secuencias peptídicas, pueda ser la responsable de una multitud tan elevada de procesos diferentes y, por supuesto, no justifica su elevada promiscuidad funcional. Por tanto, deben existir otros mecanismos para el reconocimiento molecular, aparte de los tradicionalmente descritos, que expliquen cómo estas proteínas altamente funcionales consiguen esa promiscuidad sin perder sus especificidades por los respectivos ligandos. Se han descrito experimentalmente algunos ejemplos, como el caso de la proteína HMGA, presente en el núcleo celular y que es un factor de transcripción que regula la expresión de un total de 45 genes (Grosschedl et al., 1994). Se demostró mediante dicroísmo circular y resonancia magnética nuclear que no posee estructura (*random coil*). Éste y más ejemplos están recogidos en una revisión bibliográfica (Dunker et al., 2005) y quizá la conclusión más llamativa de todos los ejemplos descritos en dicha revisión es que estas proteínas suelen mostrar un elevado grado de desorden estructural, presentando regiones con una elevada variabilidad conformacional que podrían ser las responsables, en última instancia, de los cambios estructurales que sufren y que son necesarios para modular su actividad. Estas regiones se encontrarían en lazos de conexión entre estructuras secundarias dentro de los propios dominios e incluso serían las propias de conexión entre los dominios. De ahí que un estudio conformacional detallado de estas proteínas y de sus dominios permita un avance en la comprensión de sus propiedades de unión y, al fin y al cabo, de su capacidad de ejercer la funcionalidad tan elevada que poseen.

2.3 La proteína PSD95

Un ejemplo de sistemas multimodulares *hub* que contienen dominios PDZ es la familia de proteínas MAGUK (*Membrane-Associated Guanylate Kinase*). Esta familia de proteínas ha sido identificada como componentes de estructuras celulares muy importantes como las barreras entre células epiteliales o las densidades sinápticas en las neuronas (Kim, 1995). Una de las proteínas MAGUK es la proteína PSD95 (también denominada SAP90 o DLG4), de la cual se eligió uno de sus dominios PDZ como objeto de estudio en esta Tesis Doctoral. La proteína PSD95 es una de las proteínas a partir de las cuales fueron originariamente identificados los dominios PDZ, y realiza su compleja función a nivel cerebral en la densidad post-sináptica durante el proceso de reconocimiento neuronal (sinapsis).

La densidad post-sináptica es el nombre que se le dio a finales de los años 60 a una zona oscura que se distingue con el microscopio electrónico en la zona superior de las espinas dendríticas (Palay, 1958). En estas espinas que forman las dendritas se organiza la sinapsis, formada por los axones de al menos cientos e incluso miles de neuronas. El oscurecimiento observado, y que se puede distinguir en la Figura 2.5, se debe a la existencia de un complejo supramacromolecular, fundamentalmente formado por proteínas, que aparece como un ensanchamiento de la membrana post-sináptica, extendiéndose unos 30 nm dentro del citosol (Kennedy, 2000).

Este complejo electro-denso contiene una concentración muy alta de receptores del neurotransmisor glutamato, unos de tipo NMDA (N-metil-D-aspartato) y otros tipo AMPA (ácido α -amino-3-hidroxi-5-metil-4-isoxazolopropiónico) (Kennedy, 2000). Cuando se produce un estímulo, la molécula de glutamato se libera de la terminal pre-sináptica y actúa en estos receptores que permiten el flujo de iones específicos (Na^+ , K^+ , Mg^{2+} y Ca^{2+}) que cambian la polaridad de la zona post-sináptica. De esta forma, las señales neuronales se reciben, se descodifican y se propagan mediante la densidad post-sináptica, que se puede ver por esto como un orgánulo asociado a la membrana cuyo papel es la transducción y procesamiento de la señalización post-sináptica (Scannevin & Huganir, 2000; Sheng & Sala, 2001; Okabe, 2007).

Dentro de este complejo, además de los receptores del neurotransmisor glutamato, hay moléculas que permiten la unión entre células, complejos que permiten la síntesis de proteínas y proteínas que ejercen funciones diversas: de señalización (entre ellas, quinasas y fosfatasa), organizadoras del citoesqueleto, reguladoras del tráfico de membrana y otras proteínas tipo *hub*, que, con los dominios modulares que las forman, permiten mediante sus interacciones con otras proteínas o moléculas diana el mecanismo de señalización (Cheng et al., 2006). De entre estas últimas, la más abundante es la proteína PSD95, de la que hay de unas 300 a 400 copias en la densidad post-sináptica (Feng, 2009), aunque puede variar de un tipo de neuronas a otras en función del estado fisiológico y de desarrollo en que se encuentren. Esto se corresponde con los cambios descritos en la composición de proteínas que se dan en las sinapsis y que constituyen la base molecular de la plasticidad de esta conexión neuronal (Malenka & Bear, 2004).

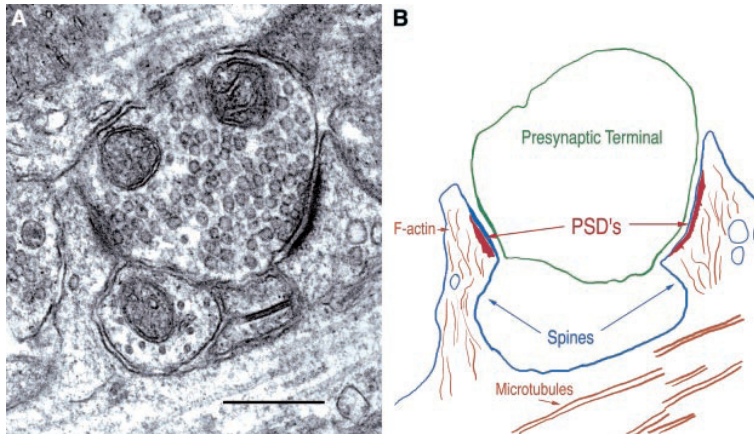


Figura 2.5. (A) Fotografía tomada mediante un microscopio electrónico que muestra una terminal pre-sináptica que forma sinapsis con dos espinas dendríticas. La barra de escala corresponde a 400 nm. (B) Dibujo esquemático de la micrografía (A) en la que se señalan las estructuras sinápticas más relevantes entre las que se encuentra la densidad post-sináptica (PSD's) (figura tomada de (Kennedy, 2000)).

La proteína PSD95 se encarga en concreto, como proteína *hub* que es, de interaccionar con las proteínas de la membrana post-sináptica como los receptores del neurotransmisor glutamato, canales iónicos y moléculas que unen las células, así como con proteínas citoplasmáticas. Durante esas interacciones todas estas proteínas se localizan y agrupan en dicha membrana (Kim & Sheng, 2004). Así, la proteína PSD95 agrupa los receptores tipo NMDA y una familia de canales iónicos de K^+ llamados "Shaker" en la superficie de estas células (Kim et al., 1995). Esta proteína también interacciona con la neuroligina (Irie et al., 1997), una proteína de la membrana post-sináptica cuyo papel es la constitución de un puente con la región pre-sináptica de la siguiente neurona (Dean et al., 2003; Li & Sheng, 2003).

Por otro lado, la proteína PSD95 organiza la señalización post-sináptica, que es posiblemente su función más importante. Así, puede interaccionar con la proteína nNOS, la sintasa neuronal de óxido nítrico (Brenman et al., 1996; Sheng, 2001), una enzima que produce óxido nítrico, que a su vez es un transmisor que está implicado en la regulación de la neurotransmisión y la excitotoxicidad (McGee & Brecht, 2003). También interacciona con la proteína adaptadora de la enzima quinasa A (Colledge et al., 2000), que permite situar estas quinasas, cuyo papel en la señalización celular ya se ha comentado, cerca de sus sustratos específicos en la sinapsis. Finalmente, la proteína PSD95 regula los procesos de transmisión en la sinapsis mediante su interacción con la proteína estargazina, que es esencial para la expresión en la superficie y la acumulación sináptica de los receptores del neurotransmisor de glutamato denominado AMPA (Chen et al., 2000; Schnell et al., 2002). Todos estos ejemplos, los más representativos, ilustran adecuadamente la función *hub* de esta proteína en las neuronas y las conexiones entre ellas, las sinapsis. Va más allá del propósito de esta Memoria exponer con exhaustividad las numerosas funciones específicas conocidas hasta ahora que ejerce esta proteína, las cuales se pueden encontrar en algunas revisiones bibliográficas recientes (Kim & Sheng, 2004; Feng & Zhang, 2009).

2.3.1 Características de los dominios modulares que componen la proteína PSD95

La proteína PSD95, como todas las pertenecientes a la familia MAGUK, contiene tres dominios PDZ, un dominio SH3 y uno GK (Kim, 1995); sin embargo, a pesar de la presencia de este último dominio, no se ha detectado actividad quinasa en esta familia de proteínas. Al igual que con los dominios PDZ y SH3 que la forman, el dominio GK funciona como otro módulo de interacción proteína-proteína. Como ya se ha mencionado, el hecho de que esta proteína sea responsable de una multitud tan elevada de procesos diferentes puede deberse a una plasticidad estructural originada por la plasticidad intrínseca de cada uno de estos dominios que la forman, así como de las secuencias de conexión entre ellos.

2.3.1.1 Los dominios SH3 y GK de la proteína PSD95

En cuanto a los dominios SH3 y GK, estos forman una unidad integrada en la que el dominio SH3 está dividido en dos regiones separadas secuencialmente por una región que hace de bisagra y el dominio GK. Para que se organice correctamente requiere que ambas regiones se unan de forma intra- o intermolecular (McGee et al., 2001). Esta característica puede permitir la auto-asociación de esta proteína requerida para algunas de sus funciones biológicas (Hsueh & Sheng, 1999; Christopherson et al., 2003), en que deben estar los tres dominios PDZ accesibles para la interacción con sus moléculas diana (Figura 2.6). La formación de oligómeros se ha encontrado también en este tipo de proteínas que contienen dominios PDZ (Im et al., 2003), aunque esta oligomerización ocurre mediante interacciones entre los propios dominios PDZ.

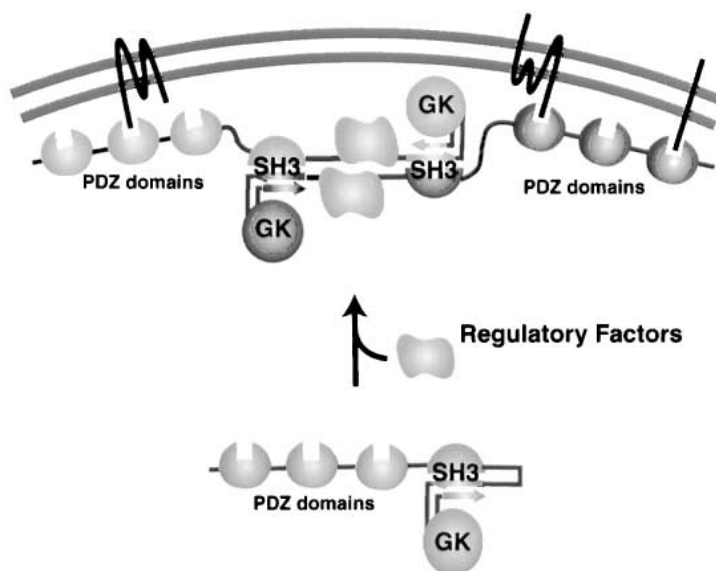


Figura 2.6. Modelo esquemático del papel que tiene el módulo SH3-GK en la proteína PSD95 (figura tomada de (McGee et al., 2001)).

2.3.1.2 Los dominios PDZ1 y PDZ2 de la proteína PSD95

En cuanto al primer y segundo dominio PDZ de la proteína PSD95 se ha observado que ambos dominios aislados tienen estructuras muy similares (Tochio et al., 2000; Long et al., 2003) y propiedades de unión parecidas (Dakoji et al., 2003; Sainlos et al., 2011). La unión entre ambos dominios es una secuencia de aminoácidos corta (5 residuos), rígida y muy conservada entre las distintas especies, por lo que parece que su función consiste más en la restricción del movimiento inter-dominio que un mero conector pasivo. De hecho, la estructura de este tándem muestra que estos dos dominios PDZ forman contactos laterales entre sí y dejan sus sitios de unión orientados de tal forma que se puedan unir sus proteínas diana, por lo que se ha llegado a considerar a este tándem simplemente como un duplicado de un dominio que, por una parte promueve la dimerización de receptores o canales iónicos para que ejerzan su función (Panel A de la Figura 2.7) o bien permite la unión a algunas de sus moléculas diana que, a su vez, presentan dos regiones de unión similares a estos dominios en zonas contiguas (Panel B de la Figura 2.7) (Long et al., 2003; Kalia & Salter, 2003; Sainlos et al., 2011).

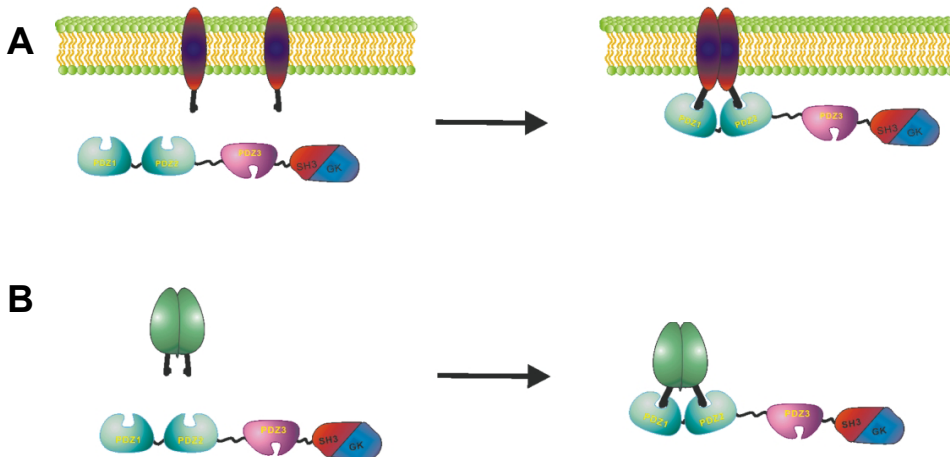


Figura 2.7. Diagrama que muestra los modelos de las funciones de los dominios PDZ1 y PDZ2 de la proteína PSD95 cuando actúan en tándem. En el panel A se representa la dimerización de un receptor de membrana inducida por la actuación de ambos dominios, mientras que en el panel B se muestra la interacción de los dos dominios con una única molécula diana que posee dos regiones de unión muy parecidas (figura tomada de (Long et al., 2003)).

2.3.1.3 El dominio PDZ3 de la proteína PSD95

El tercer dominio PDZ de la proteína PSD95, en adelante citado como PDZ3-PSD95, ejerce, junto con los otros dos dominios PDZ presentes en esta proteína, la función de interactuar con sus distintas moléculas diana. Ha sido objeto de numerosos estudios de sus propiedades de unión, algunos de los cuales han tratado de establecer cómo estas propiedades dependen de su plasticidad estructural y las propiedades dinámicas de su plegamiento.

En concreto, en un estudio computacional en el que se mutaron varios restos del dominio se ha mostrado la existencia de una red intramolecular de residuos que interaccionan entre sí, conservada entre las distintas especies, permitiendo la comunicación entre distintas partes de la molécula (Lockless & Ranganathan, 1999). Sin embargo, hay poca evidencia de que exista una transición conformacional al producirse la unión de la molécula diana al dominio PDZ3-PSD95, por lo que la unión parece ocurrir mediante el mecanismo clásico de llave-cerradura (Gianni et al., 2006; Jemth & Gianni, 2007). Por otro lado, se ha detectado la influencia de la tercera hélice α del dominio mediante un mecanismo alostérico, ya que se ha visto que al eliminar este elemento estructural disminuye la afinidad por el ligando (Petit et al., 2009). Esto parece deberse únicamente a factores entrópicos pues no conlleva cambios estructurales, lo que indica un tipo de alostería dinámica en el que están involucrados esos residuos del dominio. Hay que señalar que esta hélice no es común al resto de dominios PDZ pues suelen presentar sólo dos hélices α en su estructura, tal y como ya se ha mencionado. Además, inmediatamente a continuación comienza la secuencia del dominio SH3, lo que la define como una secuencia de conexión interdominios, radicalmente distinta a la secuencia conectora de los dominios PDZ1 y PDZ2, aunque, a la vez que ésta, con mucha relevancia para la proteína PSD95 entendida en su conjunto.

2.4 Estudios de plegamiento del dominio PDZ3-PSD95 y de otros dominios PDZ

Los estudios de plegamiento que se realizaron con el dominio PDZ3-PSD95 hasta el inicio de esta Tesis Doctoral, han revelado que este dominio presenta un mecanismo de desplegamiento algo más complejo que el de dos estados, donde se podrían poblar estados intermedios. Así, cuando se realizaron experimentos de intercambio de hidrógeno del estado nativo, se observó un estado intermedio en el que la región desplegada es el extremo C-terminal del dominio (Feng et al., 2005). Sin embargo, la construcción empleada en este trabajo contiene algunos aminoácidos que no pertenecen a la proteína natural, sino que provienen del producto de clonación en el vector de expresión. En dicho trabajo, además, al realizar experimentos cinéticos de plegamiento con la técnica de flujo detenido, se observó que sigue aparentemente un modelo de dos estados, lo que hacía concluir que dicho intermedio de plegamiento está “escondido”, es decir, que no se detecta cinéticamente y podría estar por tanto antes o después del estado de transición. Sin embargo, Chi y colaboradores (Chi et al., 2007) ampliaron dichos estudios cinéticos de plegamiento con el dominio PDZ3-PSD95 utilizando más condiciones, lo que le llevó a poder detectar un estado intermedio en condiciones estabilizantes (con Na_2SO_4) o desestabilizantes (en condiciones ácidas de pH).

Al mismo tiempo, en dicho trabajo también se detectó un intermedio cuando se siguió la cinética de plegamiento con la técnica de flujo detenido en los dominios PDZ1 y PDZ2 de la proteína PSD95 y en el dominio PDZ de la proteína nNOS, por lo que es posible que sea una característica común a este tipo de dominios modulares. De hecho,

antes de este estudio ya se había descrito la presencia de un intermedio cinético en el plegamiento del dominio PDZ2 de la proteína PTP-BL (Gianni et al., 2005). Además, las posiciones de los dos estados de transición a lo largo de la coordenada de reacción también son aproximadamente los mismos en todos estos dominios PDZ, por lo que se podría hablar de un mecanismo de plegamiento común para ellos (Chi et al., 2007). Tanto es así, que al construir un permutante circular del dominio PDZ2-PTP-BL, en el que la lámina $\beta 1$ se eliminó del extremo N-terminal de la proteína y se situó en el extremo C-terminal de la lámina $\beta 6$ (Figura 2.8), también se detectó un intermedio cinético en su plegamiento (Ivarsson et al., 2008).

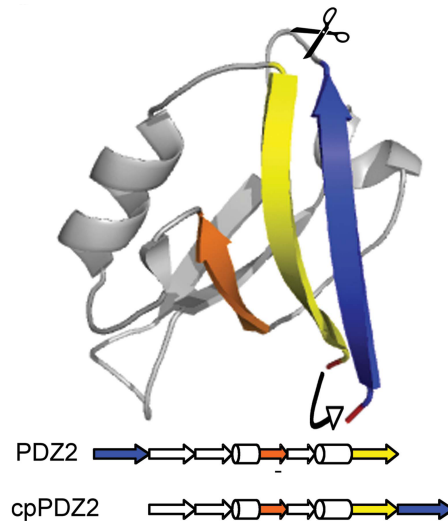


Figura 2.8. Estructura del permutante circular del dominio PDZ2-PTP-BL, en el que la lámina $\beta 1$, en color azul, es unida al extremo C-terminal de la proteína, a través de la lámina $\beta 6$, en color amarillo (figura tomada de (Ivarsson et al., 2008)).

Se ha estudiado si dicho intermedio ha de ser un paso obligatorio (intermedio *on-pathway*) en el mecanismo de plegamiento. Por un lado, Ivarsson y colaboradores comprobaron que en el caso del dominio PDZ2-PTP-BL el intermedio es efectivamente un paso obligatorio (Ivarsson et al., 2007). Sin embargo, en el caso del dominio PDZ2 de la proteína SAP97, se observó que el comportamiento experimental varía en función de la temperatura. Mientras que a 25°C los datos experimentales obtenidos a partir de la técnica de flujo detenido se ajustan a un modelo en el que el intermedio no es un paso obligatorio, a 37°C el modelo que simula bien los datos experimentales implica que dicho intermedio está incluido obligatoriamente en el mecanismo de plegamiento (Haq et al., 2010). De ahí que no se pueda deducir un comportamiento general en los dominios PDZ en cuanto a la obligatoriedad de la presencia del intermedio en su mecanismo de desplegamiento.

Por otro lado, se han realizado hasta cuatro estudios diferentes en los que se obtuvieron los valores ϕ de los estados de transición de diferentes dominios PDZ. Estos valores ϕ , que se obtienen para mutantes determinados de una proteína, son el resultado de dividir el cambio de la energía de activación entre el cambio en estabilidad del mutante

con respecto al estado nativo (Fersht, 1999). Dichos valores dan, de esta forma, información energética para el residuo específico que se muta y pueden ser usados para construir modelos para la organización energético-estructural de los estados de transición (Fersht et al., 2002), usándolos, por ejemplo, como restricciones estructurales en simulaciones de dinámica molecular (Vendruscolo et al., 2001; Paci et al., 2002). El primer estudio, correspondiente al dominio PDZ2-PTP-BL, reveló que el primer estado de transición, que corresponde al paso desde el estado desplegado al estado intermedio, es muy heterogéneo y no posee apenas estructura secundaria, sólo se mantienen algunas interacciones entre las láminas $\beta 1$, $\beta 4$ y $\beta 6$ del estado nativo (Gianni et al., 2007). En cambio, el segundo estado de transición es mucho más compacto y se parece más al estado nativo, presentando las mismas interacciones entre las láminas $\beta 1$, $\beta 3$, $\beta 4$ y $\beta 6$, junto con las interacciones entre la cadena $\beta 1$ y la hélice $\alpha 1$ (Figura 2.9).

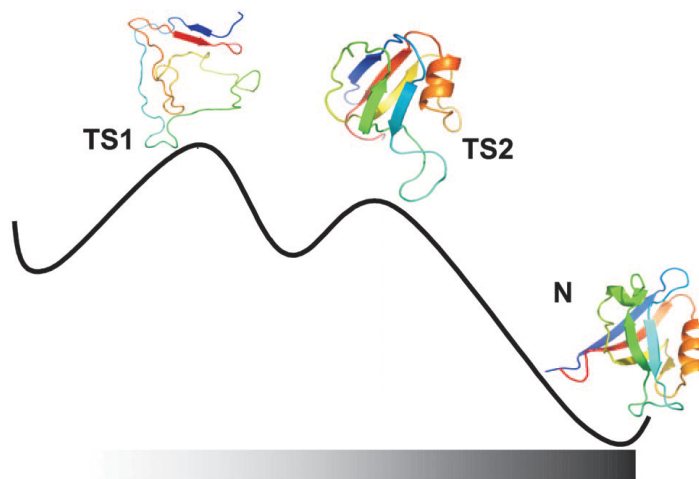


Figura 2.9. Diagrama de energía libre durante el plegamiento del dominio PDZ2-PTP-BL. En los dos máximos de la curva están representadas las estructuras tanto del primer estado de transición (TS1) como del segundo (TS2). El estado nativo (N) se representa a continuación del segundo estado de transición. (figura tomada de (Gianni et al., 2007)).

En el estudio del estado de transición del dominio PDZ3-PSD95 se observó igualmente que el estado de transición del paso de estado desplegado a estado intermedio tiene poca estructura secundaria definida, presentando interacciones entre la hélice $\alpha 2$ y las láminas $\beta 1$ y $\beta 2$, por lo que difiere estructuralmente con el equivalente al dominio PDZ2-PTP-BL (Calosci et al., 2008). Por otro lado, la estructura del segundo estado de transición del dominio PDZ3-PSD95 es prácticamente idéntica en ambos dominios, por lo que los mecanismos de plegamiento difieren en las etapas iniciales, pero convergen en las etapas finales. Esto puede implicar que la topología nativa ejerza un control del mecanismo de plegamiento en los estados cercanos al estado nativo, pero permite más libertad en estados alejados de éste. Este postulado se confirmó en dos estudios posteriores de estados de transición de dominios PDZ, realizados con el permutante circular del dominio PDZ2-PTP-BL (Ivarsson et al., 2009), que se ha descrito dos párrafos arriba y que se

puede observar en la Figura 2.8, y con el dominio PDZ de la proteasa D1 del alga *Scenedesmus obliquus* (Gianni et al., 2010), que además tiene las características de ser un permutante circular natural de un dominio PDZ. En ambos trabajos, nuevamente, y aunque tengan una conectividad distinta entre elementos de estructura secundaria, el primer estado de transición difiere en estructura entre los otros dominios estudiados, pero en el segundo la estructura es prácticamente la misma.

Todos estos resultados, y teniendo en cuenta su flexibilidad funcional, además de su tamaño relativamente grande para ser dominios modulares (unos 100 aminoácidos), indican que es poco probable que el dominio PDZ3-PSD95, y los dominios PDZ en general, por analogía con los paralelismos observados en los experimentos cinéticos, no presente otra conformación, distinta de la nativa, que sea marginalmente estable en condiciones de equilibrio.

2.5 La proteína Erbin

La otra proteína cuyo dominio PDZ ha sido estudiado en esta Tesis Doctoral ha sido la proteína Erbin. Ésta pertenece a la familia de proteínas LAP, que contiene dominios LRR (*Leucine-Rich Repeat*, repetición rica en leucina, en inglés) y dominios PDZ. Este tipo de proteínas, que se expresan en una amplia variedad de tejidos, entre los cuales están el cerebro y el epitelio, contienen 16 dominios LRR en la zona N-terminal y de 0 a 4 dominios PDZ en su región C-terminal (Ko & Kim, 2007). Erbin, en concreto, posee únicamente un dominio PDZ en su estructura y fue identificada al estudiar las proteínas que interactúan con los distintos receptores del factor de crecimiento epidérmico (Borg et al., 2000). Estos receptores, denominados ERBB/HERR, y entre los que están las variantes ERBB/HERR2, ERBB/HERR3 y ERBB/HERR4, se activan mediante su fosforilación por la acción de enzimas quinasa de tirosina, lo que provoca el inicio de la cascada de señalización dentro de la célula. A través de su dominio PDZ, la proteína Erbin interactúa con el receptor ERBB/HERR2, lo cual además le da el nombre (***ERBB/HERR2** Interacting protein*, proteína que interactúa con el receptor ERBB/HERR2).

Esta interacción tiene mucha relevancia en las neuronas puesto que la proteína ERBB/HERR2 actúa como receptor de las neurogulinas, cuyo papel es esencial para la transcripción de genes que codifican los receptores del neurotransmisor acetilcolina (Ozaki et al., 1997; Yang et al., 1998; Rieff et al., 1999). Aún no se conoce exactamente el mecanismo de localización y regulación de estas moléculas, pero se ha demostrado que la proteína Erbin interactúa a través de su extremo C-terminal con la proteína PSD95 (Huang et al., 2001), lo cual lleva a pensar que la interacción entre estas dos proteínas juega un papel fundamental en la localización de las neurogulinas y uno de los receptores sobre los que actúan.

Por otro lado, se ha observado una sobre-expresión del receptor ERBB2/HER2 en cánceres de tipo epitelial, como el cáncer de mama, de ovario o de pulmón (Klapper et al., 1999; Tao et al., 2009). Puesto que la proteína Erbin es la que permite la localización de este receptor a través de la unión con su dominio PDZ, se han estudiado en profundidad

las propiedades de unión de este dominio por el valor terapéutico que tendría bloquear su interacción con el extremo C-terminal de ERBB2/HER2, de ahí que se conozcan de forma muy precisa las razones estructurales y moleculares de su especificidad de unión.

2.5.1 Características de la especificidad de unión del dominio PDZ-Erbin

Como ya se ha mencionado, las causas de que los dominios PDZ presenten diferentes especificidades se deben fundamentalmente a su capacidad para reconocer diferentes aminoácidos en las cuatro últimas posiciones del ligando peptídico, que según la numeración citada se corresponden a las posiciones 0, -1, -2 y -3. En el mapa de especificidad de los dominios PDZ en que se obtuvo esta información mediante librerías peptídicas usadas frente a más de 150 dominios PDZ naturales (Tonikian et al., 2008), también se utilizó como modelo de dominio PDZ el dominio PDZ-Erbin en el que se mutaron varios restos de distintos elementos estructurales, lo que permitió avanzar en el conocimiento de cuáles de estos restos son los que determinan la especificidad del dominio para cada una de las posiciones del ligando.

Así, los contactos que determinan la especificidad en las posiciones 0 y -2 están localizados en una determinada región del dominio (Figura 2.10). En la posición 0 del ligando, que es el extremo C-terminal de la molécula diana, aparte de la interacción descrita del grupo carboxilato del residuo presente en dicha posición con los protones amida del motivo GLGF del dominio, es muy importante la interacción de la cadena lateral del aminoácido presente en el ligando con las cadenas laterales de los residuos con los que está en contacto, que son el sexto residuo presente en el giro que hay entre las láminas $\beta 1$ y $\beta 2$ ($\beta 1:\beta 2-6$), el primer residuo de la lámina $\beta 2$ ($\beta 2-1$) y el noveno y décimo residuos de la hélice $\alpha 2$ ($\alpha 2-9$, $\alpha 2-10$). Estas interacciones tienen generalmente un carácter marcadamente apolar. De hecho, en el mapa de especificidad citado, en la posición 0 del ligando únicamente hay aminoácidos apolares y de cadena lateral grande, con la única excepción de la cisteína, presente en los perfiles de especificidad de cuatro dominios PDZ, los cuales forman una de las 16 clases de especificidad. Los más frecuentes en las interacciones presentes en la naturaleza son la valina, la leucina y la isoleucina (Tonikian et al., 2008), siendo la valina el óptimo para el dominio PDZ-Erbin (Skelton et al., 2003). Al igual que ocurre en la posición 0, la especificidad en la posición -2 también depende de una región del dominio cuyos restos forman contactos directos con el residuo presente en dicha posición. Los dos residuos que se ha demostrado hasta ahora que alteran la especificidad son el primer resto ($\alpha 2-1$) y el quinto resto ($\alpha 2-5$) de la hélice $\alpha 2$ del dominio (Figura 2.10). Cuando hay una histidina en la posición $\alpha 2-1$ en los dominios PDZ sus perfiles de especificidad presentan una treonina o una serina en la posición -2 (Tonikian et al., 2008), puesto que se forma un puente de hidrógeno entre el hidrógeno del grupo imidazol de la histidina y el grupo hidroxilo de las cadenas laterales de la treonina o de la serina, como de hecho ocurre en el dominio PDZ-Erbin (Appleton et al., 2006).

Finalmente, la especificidad de unión en las posiciones -1 y -3 depende tanto de contactos locales, como de otras regiones del dominio que están más alejadas

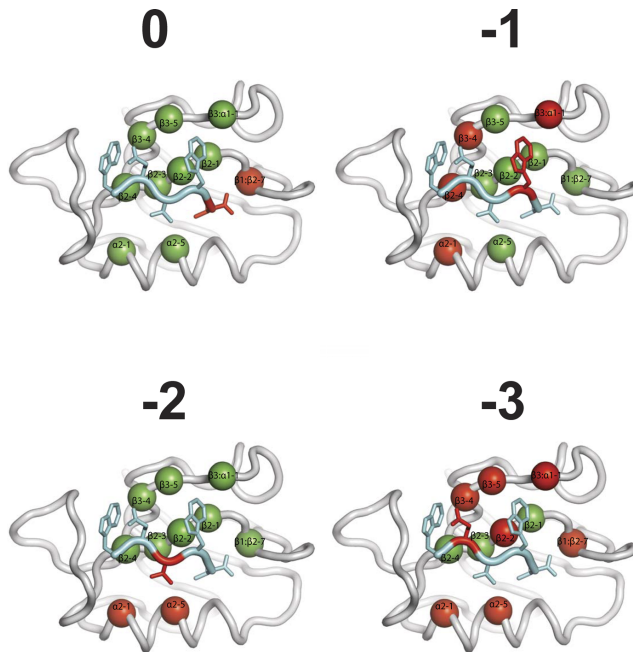


Figura 2.10. Dominio PDZ-Erbin unido al péptido de alta afinidad TGWETWV. En cada representación las posiciones del dominio PDZ en las que las mutaciones afectan la especificidad para cada posición del péptido están coloreadas en rojo y las que no, están en verde. Las posiciones del ligando están numeradas en negro y en cada una de las representaciones el aminoácido presente en la correspondiente posición está coloreado en rojo (figura tomada de Tonikian et al., 2008).

especialmente del sitio de unión (Figura 2.10). Se vio que la cadena lateral del aminoácido de la posición -1 del ligando interacciona únicamente con la cadena lateral del primer aminoácido del giro presente entre la lámina β_3 y la hélice α_1 ($\beta_3:\alpha_1-1$), y que más bien forma contactos con la cadena principal de las láminas β_2 y β_3 del dominio (Weiss et al., 2000). Igualmente ocurre con el aminoácido de la posición -3 del ligando, que forma contactos directos con el segundo resto de la lámina β_2 (β_2-2) y los restos cuarto y quinto de la lámina β_3 del dominio (β_3-4 y β_3-5). Sin embargo, tal y como se aprecia en la Figura 2.10, la especificidad de unión en las posiciones -1 y -3 depende a su vez de otros restos que no están localizados en el sitio de unión a través de efectos conformacionales indirectos, por lo que el sitio de unión del dominio PDZ ha de considerarse como el dominio PDZ en su conjunto (Tonikian et al., 2008).

Por otro lado, el conocimiento de las causas de la especificidad del dominio PDZ-Erbin condujo al Dr. Sidhu y su laboratorio a usarlo como “progenitor” de una familia sintética de dominios para un estudio *in vitro* de los mecanismos de la evolución (Ernst et al., 2009). En dicho estudio se creó una librería artificial de dominios PDZ mediante la metodología de expresión en fagos a partir del dominio PDZ-Erbin en el que se mutaron varias posiciones que tienen contacto directo con el ligando. Se seleccionaron de ellas las aproximadamente 300 variantes que mantenían mejor la estructura propia de un dominio PDZ y de las que se obtuvo un perfil de especificidad mediante librerías peptídicas. Se

observó que presentaban 14 clases de especificidad, siete de las cuales además coinciden con otras tantas clases de especificidad de los dominios naturales. Esto significa que se pueden simular de forma sintética familias de dominios, y que éstas parece que son el producto de la evolución de un progenitor que es estructuralmente robusto, pero al que las mutaciones en su secuencia primaria le permiten presentar distintas especificidades, lo cual le lleva a realizar las distintas funciones celulares.

Finalmente, a partir de 15 de los ligandos óptimos de otras tantas variantes de esta nueva familia sintética de dominios PDZ se estudió si se podían volver a seleccionar las mismas variantes de dicha familia sintética. Se usó para ello el dominio PDZ-Erbin para generar una librería de variantes con la metodología de expresión en fagos en que se mutaron de nuevo las mismas posiciones que tienen contacto directo con el ligando. Se observó que para un ligando dado podían existir varias variantes de dominios PDZ que podían unirlos con buena afinidad (Ernst et al., 2010). Al seleccionarse más variantes que la original, este estudio llevó a establecer que, mediante la generación de nuevas variantes del dominio PDZ-Erbin, su unión a los péptidos frente a los cuales fueron seleccionadas se da con afinidad alta, aunque de forma poco selectiva.

Todos estos estudios hacen del dominio PDZ-Erbin un modelo arquetípico de dominio PDZ para estudios de especificidad de unión de los dominios PDZ, por lo que es obvio pensar que este dominio es el candidato ideal para abordar estudios de especificidad más generales de estos dominios como los que se han realizado en esta Tesis Doctoral.

CAPÍTULO 3

MATERIALES Y MÉTODOS

3.1 Clonación de las dos construcciones del tercer dominio PDZ de la proteína PSD95

Actualmente es posible sintetizar, manipular y disponer de cantidades ilimitadas de cualquier material genético de interés. Además podemos producir alteraciones en el material genético incluyendo cambios en la secuencia (mutaciones), eliminación de parte de la misma (supresiones) o introducción de fragmentos nuevos en la misma (inserciones). Todo esto ha permitido el nacimiento de lo que se conoce como *Ingeniería de Proteínas*. Uno de los métodos más usados en este campo es la *Reacción en Cadena de la Polimerasa* (*Polymerase Chain Reaction* en inglés, PCR) (Mullis et al., 1987; Creighton, 1993), que nos permite producir las alteraciones deseadas en la secuencia de DNA que se use como molde, diseñando los oligonucleótidos adecuados. Así, se generan copias del nuevo DNA, al que se denomina DNA *recombinante* por proceder de la “combinación” de moléculas de DNA de distinto origen. Éste método se ha escogido para la producción de las dos construcciones del dominio PDZ3 de la proteína PSD95. Nuestro fin fue acotar el dominio a los aminoácidos que nos interesaban.

El molde usado fue un plásmido que ya contenía la proteína de estudio. Este plásmido nos lo proporcionó el Dr. José Reina, del grupo de investigación del Dr. Cayetano González (Centro de Regulación Genómica, CRG, Barcelona). El estudio de la estructura de este dominio unido a un péptido (Doyle et al., 1996) y otros estudios posteriores (Reina et al., 2002; Chi et al., 2007) se llevaron a cabo con una secuencia aminoacídica que incluía 13 aminoácidos en el extremo C-terminal de la proteína y 5 aminoácidos en el extremo N-terminal derivados del propio vector utilizado en el proceso de clonación.

Así, la clonación que se ha llevado a cabo en esta Memoria consistió en codificar únicamente el fragmento de secuencia 302-402 (numeración usada en el artículo de (Doyle et al., 1996)) del dominio PDZ3-PSD95. Se hizo así para evitar que los 13 restos del extremo C-terminal pudieran adoptar una estructura que no se encuentra en la proteína natural. Por otro lado, para la clonación del dominio PDZ3 sin su tercera hélice alfa, la construcción que se denomina en la Memoria como PDZ3-h-PSD95, se acotó a los aminoácidos 302 hasta 393 (ver Capítulo 7).

3.1.1 Clonación de los dominios PDZ3-PSD95 y PDZ3-h-PSD95

Los métodos de clonación vienen descritos con detalle en varias Tesis Doctorales del Departamento (Cobos, 2002; Candel, 2008), por lo que no se ha considerado necesario detallar el protocolo en esta Memoria. Para obtener sendas construcciones, las reacciones de PCR se realizaron en un termociclador modelo Prog02D de la casa PROGENE. Se comprobó que el resultado fue positivo mediante un gel de agarosa al 2%. La identificación de una única banda también permitió determinar que el producto de reacción fue el único obtenido. Este último se purificó en un siguiente paso siguiendo el protocolo de purificación de productos de PCR comercial de QIAGEN (*PCR Purification Kit, Qiaquick*).

Posteriormente, el inserto purificado se digirió con las endonucleasas NcoI y HindIII. Cabe destacar que durante el diseño de los oligonucleótidos la metionina inicial (ATG) está incluida en la diana de restricción de NcoI de tal forma que deja una guanina (G) que desplaza el marco de lectura de los tripletes de pares de bases del dominio PDZ. Por esta razón, se decidió añadir otras dos guaninas a la secuencia de dicha diana de restricción, de forma que, a continuación del triplete de la metionina inicial, pueda leerse el triplete GGG, que *E. Coli* traduce como una glicina, aminoácido que tiene el menor impacto posible en las características del dominio PDZ que se quería producir.

Para la ligación del inserto al vector de expresión pBAT4 (EMBL Core Purification Facility), se usó un protocolo comercial de ligación de la casa comercial ROCHE (*Rapid Ligation Kit*). El producto de la ligación (vector con el inserto unido) se introdujo, mediante un proceso de transformación, en bacterias *E-coli* de la cepa XL1Blue. Esta cepa permite obtener un número elevado de copias del plásmido que procede de la ligación. Finalmente, el plásmido se purificó mediante una batería comercial de QIAGEN.

3.1.2 Ensayos de expresión

El plásmido puro se transformó a su vez en células de la cepa BL21-DE3, que tienen una alta capacidad de expresión de proteína. Se seleccionaron a continuación algunas de las colonias obtenidas en esta transformación para hacer ensayos de expresión bajo distintas condiciones.

Así, se añadió a un tubo de ensayo estéril 5 mL de medio de cultivo Laura Bertani (LB), con el antibiótico ampicilina ($50 \mu\text{g}\cdot\text{mL}^{-1}$) e IPTG ($40 \mu\text{g}\cdot\text{mL}^{-1}$), el agente inductor que provoca la sobreexpresión de la proteína después de inocular las colonias en el mismo medio de cultivo. En otro tubo independiente, la inducción con IPTG se realizó cuando la densidad óptica, medida a 600 nm, alcanzó un valor de 0.6. En cualquiera de los dos casos se recogieron muestras transcurridas 4 horas desde la inducción y a la mañana siguiente, para establecer la posible toxicidad de la proteína.

Los niveles de expresión de las muestras de los ensayos se determinaron mediante electroforesis en gel de poliacrilamida al 16% en presencia de SDS (SDS-PAGE) (Cobos, 2002). Los resultados indicaron que no existen diferencias significativas en los niveles de expresión asociados a diferentes tiempos de inducción, con lo que, con el fin de facilitar el proceso experimental de purificación de la proteína, se decidió llevar a cabo la inducción al inicio del crecimiento celular. A continuación, se purificó el plásmido de la colonia de mayor expresión, enviando el DNA puro al servicio de secuenciación del Instituto de Parasitología y Biomedicina López Neyra para confirmar que el plásmido escogido contenía la secuencia de bases que codifica la secuencia aminoacídica correcta de nuestra proteína.

3.2 Expresión y purificación de las proteínas

3.2.1 Expresión y purificación de las dos construcciones del dominio PDZ3 de la proteína PSD95

Las células, conteniendo el plásmido correspondiente, o bien el del dominio PDZ3-PSD95 o bien el del dominio PDZ3-h-PSD95, se crecieron en varios litros de medio de cultivo LB, induciendo desde el comienzo con IPTG, como se describe en el párrafo anterior. Al cabo de 15 horas dichos cultivos celulares se centrifugaron en una centrífuga Sorvall, a 4°C y 7000 rpm. Posteriormente el pellet se resuspendió en tampón Tris 50 mM pH 7.5. La lisis de las células se realizó con una Prensa de French a 1000 psi y posteriormente se centrifugó en una ultracentrífuga Beckman a 4°C y 30000 rpm durante 30 minutos. El pH del sobrenadante se ajustó al valor de 3 (donde precipitan la mayoría de las proteínas de *E. coli*) y con un gel de electroforesis al 16% se comprobó que la proteína es soluble a dicho pH. A continuación se precipitó en presencia de sulfato amónico al 75% con agitación a 4°C y, una vez añadido todo el sulfato se dejó durante toda la noche sin la agitación.

La última etapa de la purificación consistió en una cromatografía de exclusión molecular (ver apartado 2.8), en la que el dominio se obtuvo con una pureza mayor del 95%. Con este método de purificación el rendimiento final fue tal que se obtuvieron unos 50 mg de proteína pura por cada litro de cultivo en el caso del dominio PDZ3-PSD95, mientras que unos 10 mg en el caso del dominio PDZ3-h-PSD95.

Con el fin de confirmar que las proteínas purificadas corresponden a ambas construcciones del dominio PDZ3 de la proteína PSD95, se obtuvieron sus respectivos espectros de masas. Los resultados proporcionados por distintos laboratorios (el Centro de Instrumentación Científica de la Universidad de Granada, y el servicio de espectrometría de masas del Centro de Investigaciones Biológicas (CIB), en Madrid) mostraron que la masa difería en 131 daltons con respecto a la esperada, por lo que se decidió secuenciar también el extremo N-terminal de la proteína en el servicio de Química de Proteínas del CIB. Así se comprobó que las proteínas pierden la metionina inicial de la secuencia, hecho que parece darse con cierta frecuencia en el caso de la expresión en bacterias *E-coli* (Hirel et al., 1989; Dalboge et al., 1990).

3.2.2 Expresión y purificación de los péptidos unidos a la proteína S-Transferasa de Glutación (GST)

Las células se crecieron en 250mL de medio de cultivo LB, añadiendo IPTG cuando la densidad óptica, medida a 600 nm, alcanzó un valor de 0.6. Al cabo de 15 horas a 16°C dichos cultivos celulares se centrifugaron en una centrífuga Sorvall, a 4°C y 7000 rpm. Posteriormente el pellet se resuspendió en tampón PBS (*Phosphate buffered saline*, tampón salino de fosfato). La lisis de las células se realizó mediante ultrasonificación con una amplitud del 40% durante 10 minutos utilizando pulsos alternantes de 5 segundos

funcionando el ultrasonidos y otros 5 segundos parado, para evitar un excesivo calentamiento de la muestra. Posteriormente se centrifugó en el rotor ss34 de la centrífuga Sorvall a 4°C y 18000 rpm durante 20 minutos. El sobrenadante se cargó en una columna con 300 μ L de resina Sepharose 4B de Glutación (GE Healthcare) previamente equilibrada con PBS y se llevaron a cabo sucesivos lavados con PBS, PBS + NaCl 150 mM y PBS de nuevo. Por último, la proteína se eluyó con Tris 50mM pH 8 y glutación 10mM. Los péptidos fusionados a GST se obtuvieron con una pureza mayor del 90% determinado esto último mediante electroforesis en gel de poliacrilamida al 16% en presencia de SDS (SDS-PAGE). Con este método de purificación el rendimiento final fue tal que se obtuvieron unos 3 mg de péptido unido a GST.

3.3 Preparación de tampones y de disoluciones de proteína

Las muestras empleadas para la realización de los experimentos recogidos en esta Memoria se dializaron en una membrana de diálisis de SPECTRA-POR con un peso molecular de corte entre 6000 y 8000 Daltons frente al tampón correspondiente. Las disoluciones tampón se prepararon a partir de las formas ácida y básica del propio tampón, comprobando y corrigiendo en su caso, el pH al valor deseado con un pHmetro (CRISON). Las muestras dializadas se centrifugaron en una microcentrífuga para eliminar las partículas en suspensión. La concentración de las muestras se determinó espectrofotométricamente en un espectrofotómetro ultravioleta-visible usando un coeficiente de extinción molar de 2985 M⁻¹·cm⁻¹ para el dominio PDZ3-PSD95 y 1425 M⁻¹·cm⁻¹ para el dominio PDZ3-h-PSD95, ambos medidos a 278 nm, establecidos por el método de Gill y von Hippel (Gill & von Hippel, 1989).

3.4 Calorimetría Diferencial de Barrido (DSC)

La calorimetría diferencial de barrido (*Differential Scanning Calorimetry*, DSC) es una técnica que permite determinar el cambio de capacidad calorífica del sistema que tiene lugar cuando se induce su desplegamiento mediante el cambio continuo y a velocidad constante de la temperatura. Durante el experimento se mide la diferencia de capacidad calorífica entre dos células prácticamente idénticas que contienen la disolución de la proteína de estudio y como referencia el tampón en el que ésta se dializa previamente. En nuestro caso, los experimentos se llevaron a cabo en dos microcalorímetros, un VP-DSC y VP-Tubo capilar-DSC, ambos de MICROCAL INC.

Antes de iniciar el barrido hasta alta temperatura con la proteína, ambas células se rellenaron con tampón de diálisis con el fin de registrar un primer barrido, a partir del cual se determinó la *línea base instrumental*. Es necesario llevar a cabo varios barridos sucesivos para asegurar la reproducibilidad de dicha línea base. A continuación, se rellenó la célula de muestra con la disolución de proteína y se realizó un barrido de temperatura entre 2 y 35°C con el fin de eliminar la presencia de posibles artefactos en la traza experimental debidos a efectos de llenado de las células o incluso a la presencia de

estados de agregación presentes en la muestra. Seguidamente, sin extraer el contenido de ninguna de las células, se realizaron dos barridos hasta altas temperaturas, el segundo de los cuales se utilizó para comprobar la reversibilidad del proceso de desplegamiento.

3.4.1 Cálculo de la capacidad calorífica molar parcial de la proteína

Los ficheros generados por el programa de adquisición de datos del calorímetro contienen el valor de una señal de voltaje (en mV), que es el que se suministra a la resistencia para mantener la velocidad de barrido constante cuando se produce algún efecto térmico, por ejemplo, la desnaturalización de una proteína, ya que se produce una diferencia de temperatura entre las células del calorímetro. Así, el voltaje suministrado a la célula de menor temperatura que mantiene en cualquier momento la diferencia de temperatura entre ambas células igual a cero, es proporcional a las diferencias de capacidad calorífica entre las células en función de la temperatura. Los ficheros se analizaron con el programa *MicroCal Origin for DSC v 4.1* en el que se llevó a cabo de forma automática la conversión de la señal de unidades de mV a unidades de capacidad calorífica (en $\text{cal}\cdot\text{C}^{-1}$ o $\text{J}\cdot\text{C}^{-1}$), así como la corrección del tiempo de respuesta del calorímetro (corrección dinámica del termograma), en función del valor de la constante de calibración del instrumento y del tiempo de respuesta seleccionado para los experimentos. Posteriormente se resta la línea base instrumental del termograma obtenido con la muestra de proteína y se normalizan los datos según la concentración de proteína y el volumen efectivo de la célula, expresando la diferencia de capacidad calorífica aparente entre la disolución de macromolécula y el disolvente, $\Delta C_{p,ap}(T)$ en términos de capacidad calorífica molar en $\text{kJ}\cdot\text{K}^{-1}\cdot\text{mol}^{-1}$. Finalmente, la capacidad calorífica molar parcial de la proteína se obtiene según:

$$C_p = \left[C_{p,s}(T) \cdot \frac{V_p(T)}{V_s(T)} PM + \Delta C_{p,ap}(T) \right] \quad (3.4.1)$$

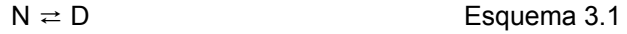
donde $C_{p,s}(T)$ es la capacidad calorífica parcial del disolvente, $V_p(T)$ y $V_s(T)$ son los correspondientes volúmenes específicos parciales de la proteína en disolución y del disolvente respectivamente y PM es el peso molecular de la proteína. El volumen específico parcial que se usa generalmente como promedio para las proteínas globulares es $V_p=0.73 \text{ mL}\cdot\text{g}^{-1}$.

3.4.2 Análisis de las trazas calorimétricas según la termodinámica de equilibrio.

Del análisis de la capacidad calorífica molar parcial de la proteína como una función de la temperatura se pueden obtener los parámetros termodinámicos del proceso de desplegamiento térmico, para lo que se aplica un modelo de equilibrio, considerando que el proceso es reversible.

3.4.2.1 Modelo de equilibrio de dos estados

El modelo de equilibrio de dos estados es el modelo más sencillo, según el cual, durante el proceso de desplegamiento de la proteína sólo existen dos estados poblados en el equilibrio, el estado nativo (N) y el desplegado (D). Dicho equilibrio puede representarse mediante el siguiente esquema:



La constante de equilibrio aparente para el proceso de desplegamiento viene dada por:

$$K_{N-D} = \frac{[D]}{[N]} \quad (3.4.2)$$

La función de partición del sistema se define como:

$$q = 1 + K_{N-D} \quad (3.4.3)$$

A partir de la función de partición pueden calcularse las poblaciones de los estados N y D:

$$f_N = \frac{1}{q} \quad (3.4.4)$$

$$f_D = \frac{K_{N-D}}{q} \quad (3.4.5)$$

La entalpía molar parcial del sistema tomando el estado nativo como referencia será:

$$\langle H \rangle = f_N \cdot H_N + f_D \cdot H_D = H_N + f_D \cdot \Delta H_{N-D} = H_N + \frac{K_{N-D}}{q} \cdot \Delta H_{N-D} \quad (3.4.6)$$

donde ΔH_{N-D} es la diferencia de entalpía de desplegamiento, $\Delta H_{N-D} = H_D - H_N$. De esta manera, la capacidad calorífica molar parcial, C_p , se define como:

$$C_p = \left[\frac{\partial \langle H \rangle}{\partial T} \right]_P = C_{pN} + \frac{\Delta H_{N-D}}{(1 + K_{N-D})^2} \left[\frac{\partial K_{N-D}}{\partial T} \right]_P + \frac{K_{N-D}}{(1 + K_{N-D})} \left[\frac{\partial \Delta H_{N-D}}{\partial T} \right]_P \quad (3.4.7)$$

donde C_{pN} es la capacidad calorífica molar parcial del estado nativo. Así, las dependencias de K_{N-D} y ΔH_{N-D} con la temperatura pueden expresarse mediante las siguientes relaciones:

$$\frac{\partial \Delta H_{N-D}}{\partial T} = \Delta C_{pN-D} \quad (3.4.8)$$

$$\frac{\partial \Delta S_{N-D}}{\partial T} = \frac{\Delta C_{pN-D}}{T} \quad (3.4.9)$$

$$K_{N-D} = \exp\left(-\frac{\Delta G_{N-D}}{T}\right) \quad (3.4.10)$$

en las que ΔS_D , ΔG_D y ΔC_p son las diferencias de entropía, de energía de Gibbs y de capacidad calorífica entre los estados D y N, respectivamente. Según los trabajos de Privalov y colaboradores, la función que describe el cambio de capacidad calorífica de desplegamiento de una proteína no es constante, sino que presenta una dependencia no lineal de la temperatura. Se puede considerar que la función que describe la capacidad calorífica molar parcial del estado nativo, $C_{pN}(T)$, es lineal (Privalov et al., 1989), aunque por el contrario, la función capacidad calorífica de una proteína desplegada se puede estimar a partir de su secuencia aminoacídica (Makhatadze & Privalov, 1990; Privalov & Makhatadze, 1990) en la que la función $C_{pD}(T)$, que representa la capacidad calorífica molar parcial del estado desnaturalizado puede describirse como un polinomio de segundo grado.

Así, se obtienen las siguientes ecuaciones para C_{pD} , C_{pN} y ΔC_{pN-D} :

$$C_{pN} = a + b \cdot T \quad (3.4.11)$$

$$C_{pD} = c + d \cdot T + e \cdot T^2 \quad (3.4.12)$$

$$\Delta C_{pN-D} = (c - a) + (d - b) \cdot T + e \cdot T^2 \quad (3.4.13)$$

Para obtener las dependencias de ΔH_{N-D} y ΔS_{N-D} con la temperatura se integran las ecuaciones 3.4.8 y 3.4.9:

$$\Delta H_{N-D}(T) = \Delta H_{N-D}(T_{N-D}) + \int_{T_{N-D}}^T \Delta C_{pN-D} \cdot dT \quad (3.4.14)$$

$$\Delta S_{N-D}(T) = \Delta S_{N-D}(T_{N-D}) + \int_{T_{N-D}}^T \frac{\Delta C_{pN-D}}{T} \cdot dT \quad (3.4.15)$$

donde $\Delta H_{N-D}(T_{N-D})$ y $\Delta S_{N-D}(T_{N-D})$ son los cambios de entalpía y entropía de desplegamiento a la temperatura de la transición (T_{N-D}). Dicha temperatura es aquella en la que f_D se iguala a 0.5 y el valor de ΔG_{N-D} se anula, haciendo que K_{N-D} adquiera el valor de 1. En ese punto, el cambio de entropía a la temperatura T_{N-D} , $\Delta S_{N-D}(T_{N-D})$, se puede obtener según la siguiente ecuación:

$$\Delta S_{N-D} = \frac{\Delta H_{N-D}}{T_{N-D}} \quad (3.4.16)$$

Si se integran las ecuaciones de $\Delta H_{N-D}(T)$ y de $\Delta S_{N-D}(T)$ con respecto a la temperatura, se obtienen las siguientes expresiones:

$$\Delta H_{N-D} = \Delta H_{N-D}(T_{N-D}) + (c - a) \cdot (T - T_{N-D}) + \frac{(d - b)}{2} \cdot (T^2 - T_{N-D}^2) + \frac{1}{3} \cdot e \cdot (T^3 - T_{N-D}^3) \quad (3.4.17)$$

$$\Delta S_{N-D} = \frac{\Delta H_{N-D}}{T_{N-D}} + (c - a) \cdot \ln\left(\frac{T}{T_{N-D}}\right) + (d - b) \cdot (T - T_{N-D}) + \frac{1}{2} \cdot e \cdot (T^2 - T_{N-D}^2) \quad (3.4.18)$$

$$\Delta G_{N-D} = \Delta H_{N-D} - T \cdot \Delta S_{N-D} \quad (3.4.19)$$

De esta forma, puede expresarse la capacidad calorífica molar parcial mediante la siguiente función:

$$C_p = C_{pN} + \frac{(\Delta H_{N-D})^2}{RT^2} \cdot \frac{K_{N-D}}{(1 + K_{N-D})^2} + f_D \cdot \Delta C_{pN-D} \quad (3.4.20)$$

Cuando se ha aplicado el modelo de dos estados en los estudios llevados a cabo en esta Memoria se ha utilizado el programa Sigma Plot 2000 (Jandel Co) para hacer los ajustes por mínimos cuadrados no lineales de las curvas de C_p obtenidas. El resultado de dicho análisis permite obtener los valores de los parámetros T_{N-D} , $\Delta H_{N-D}(T_{N-D})$ y ΔC_{pN-D} .

3.4.2.2 Modelo de equilibrio de tres estados con asociación del intermedio

El modelo sigue el esquema:



Sin embargo, para establecer las ecuaciones que describen este modelo se puede redefinir el esquema general del mismo cambiando el orden de los dos equilibrios presentes. Esto es válido teniendo en cuenta que en los principios que rigen la termodinámica los equilibrios no dependen del orden secuencial de los procesos que tienen lugar en cada una de las etapas definidas, sino que dependen únicamente de los estados inicial y final. En función de cómo se establezcan esos equilibrios, la información que darán los modelos matemáticos será de un tipo u otro, aunque perfectamente compatible, de ahí que el criterio establecido se base en la obtención de unos parámetros termodinámicos más intuitivos y comparables con los obtenidos previamente para otras proteínas y dominios.

Así, podemos establecer el equilibrio de la forma $N \rightleftharpoons (1/n) I_n \rightleftharpoons D$, siendo K_{N-I_n} la constante de la primera etapa y K_{I_n-D} la de la segunda. K_{N-I_n} (y, en definitiva el valor de ΔG_{N-I_n}) resultaría de los procesos acoplados del desplegamiento parcial y asociación, mientras que K_{I_n-D} se asociaría al proceso de disociación con desplegamiento total. Es posible también plantear el equilibrio de la forma $N \rightleftharpoons D \rightleftharpoons (1/n) I_n$, en cuyo caso la constante de la primera etapa, K_{N-D} , corresponde a la constante de desplegamiento y es comparable a las que se pueden obtener de otros modelos. Por otro lado, K_{D-I_n} , la constante de la segunda etapa, sería la constante del proceso de asociación que, de esta manera, puede ser evaluada independientemente. Por esta razón se ha considerado más conveniente llevar a cabo el análisis utilizando este segundo planteamiento.

Como, tal y como se ha comentado, los equilibrios sólo dependen de los estados inicial y final, la información termodinámica contenida en los equilibrios implicados en el esquema general debe ser la misma, lo que conlleva una equivalencia entre los valores de las constantes y parámetros termodinámicos derivados, independientemente del orden en el que se consideren las etapas dentro de dicho esquema. En este caso se demuestra que:

$$K_{N-D} = K_{N-I_n} \cdot K_{I_n-D} \quad (3.4.21)$$

$$K_{D-In} = \frac{1}{K_{In-D}} \quad (3.4.22)$$

Una vez determinadas qué especies coexisten en equilibrio, pueden definirse matemáticamente las poblaciones de dichos estados conformacionales. De esta forma, si se le llama f_N , f_D y f_I a las poblaciones de los estados nativo, desplegado e intermedio, respectivamente, y P a la concentración total de proteína, se deduce que:

$$f_N = \frac{[N]}{P} \rightarrow [N] = f_N \cdot P \quad (3.4.23)$$

$$f_D = \frac{[D]}{P} \rightarrow [D] = f_D \cdot P \quad (3.4.24)$$

$$f_I = \frac{n[I_n]}{P} \rightarrow [I_n] = f_I \cdot \frac{P}{n}; \text{ como } f_N + f_I + f_D = 1 \rightarrow [I_n] = (1 - f_N - f_D) \cdot \frac{P}{n} \quad (3.4.25)$$

Las constantes de equilibrio se pueden definir, por tanto, en función de las poblaciones, de la estequiometría de asociación n y de la concentración de proteína P como:

$$K_{N-D} = \frac{[D]}{[N]} = \frac{f_D}{f_N} \quad (3.4.26)$$

$$K_{D-In} = \frac{[I_n]^{1/n}}{[D]} \rightarrow K_{D-In}^n = \frac{[I_n]}{[D]^n} = \frac{P \cdot (1 - f_N - f_D)}{n \cdot f_D^n \cdot P^n} = \frac{1 - f_N - f_D}{n \cdot f_D^n \cdot P^{n-1}} \quad (3.4.27)$$

Para determinar las poblaciones se calcula el cociente entre constantes de equilibrio,

$$\frac{K_{D-In}^n}{K_{N-D}} = \frac{\left(1 - \frac{f_D}{K_{N-D}} - f_D\right) \cdot \frac{f_D}{K_{N-D}}}{n \cdot f_D^n \cdot P^{n-1} \cdot f_D}$$

reordenando queda que:

$$n \cdot K_{D-In}^n \cdot K_{N-D} \cdot P^{n-1} \cdot f_D^n + (1 + K_{N-D}) \cdot f_D - K_{N-D} = 0 \quad (3.4.28)$$

De esta forma se obtiene una ecuación cuya incógnita es f_D . Sabiendo que $f_N = \frac{f_D}{K_{N-D}}$ y que la suma de todas las poblaciones conformacionales debe ser igual a la unidad en todo momento ($f_N + f_I + f_D = 1$) se pueden obtener las poblaciones de dichos estados conformacionales y así deducir toda la información termodinámica del sistema (tal como se detalla más adelante).

El problema radica en que la ecuación (3.4.28) en la que aparece f_D como incógnita no se puede resolver de forma directa y, además, presenta n soluciones diferentes, varias de ellas imaginarias. Para resolver esta situación puede recurrirse a métodos numéricos

incluidos en los programas de cálculo de uso más común (como Origin 7.0 y SigmaPlot 2000).

Así, una buena aproximación podría ser el método numérico de Newton que permite resolver ecuaciones de la forma $g(x_n)=0$. La idea en la que se basa dicho método consiste en comenzar el proceso de iteración con un valor cercano al cero, denominado punto de arranque. Tal y como se observa en la Figura 3.1, la recta tangente que pasa por dicho punto de arranque, x_n en la Figura, cuando toma el valor 0 y corta por tanto el eje de abscisas, en ese punto de corte el valor de x_{n+1} constituirá, generalmente, una aproximación mejor a la solución de la ecuación. Se puede obtener fácilmente el corte de la recta tangente con el eje de abscisas, mediante la función derivada $g'(x_n)$ y a continuación se iguala a cero para posteriormente despejar el valor a partir de una ecuación lineal. Después se puede aplicar un método iterativo, empleando cada vez el nuevo valor de x_{n+1} obtenido. En nuestro caso se han llevado a cabo hasta diez iteraciones, suficientes para establecer el valor de f_D de forma suficientemente precisa.

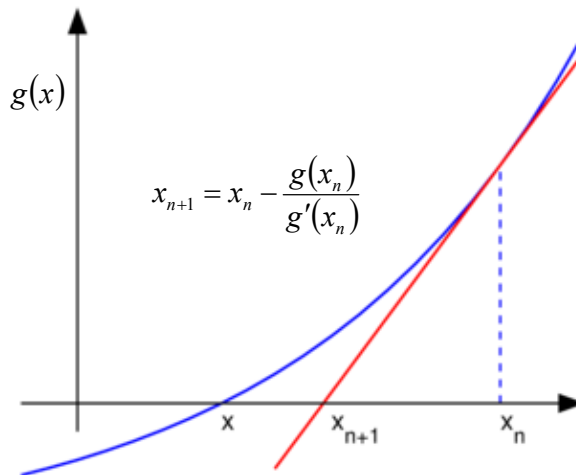


Figura 3.1. Ilustración de una iteración del método de Newton (la función $g(x)$ se muestra en azul y la línea de la tangente en rojo). x_{n+1} es una aproximación mejor que x_n para la raíz x de la función g .

Para completar el modelo con sus fórmulas matemáticas, se definen las capacidades caloríficas de los tres estados conformacionales en función de la temperatura:

$$C_{pN} = a + b \cdot T, \text{ del estado nativo} \tag{3.4.11}$$

$$C_{pD} = c + d \cdot T, \text{ del estado desnaturalizado} \tag{3.4.29}$$

$$C_{pIn} = h + i \cdot T, \text{ del estado intermedio} \tag{3.4.30}$$

A continuación se puede definir la entalpía de ambos procesos, ya que ésta es la derivada del cambio de capacidad calorífica respecto de la temperatura (se toman como temperaturas de referencia T_D y T_I a las que corresponden las entalpías ΔH_{Dm} y ΔH_{Im} respectivamente de cada equilibrio):

$$\Delta H_{N-D} = \Delta H_{N-D}(T_{N-D}) + ((c-a) \cdot (T - T_{N-D})) + \left(\frac{(d-b) \cdot (T^2 - T_{N-D}^2)}{2} \right) \quad (3.4.31)$$

$$\Delta H_{D-Im} = \Delta H_{D-Im}(T_{D-Im}) + ((h-c) \cdot (T - T_{D-Im})) + \left(\frac{(i-d) \cdot (T^2 - T_{D-Im}^2)}{2} \right) \quad (3.4.32)$$

Se obtienen también los cambios de entropía de manera similar, considerando que $dS = dH/T = C_p \cdot dT/T$; llegando a la siguiente expresión:

$$\Delta S_{N-D} = \Delta S_{N-D}(T_{N-D}) + \left((c-a) \cdot \ln\left(\frac{T}{T_{N-D}}\right) \right) + ((d-b) \cdot (T - T_{N-D})) \quad (3.4.33)$$

$$\Delta S_{D-Im} = \Delta S_{D-Im}(T_{D-Im}) + \left((i-c) \cdot \ln\left(\frac{T}{T_{D-Im}}\right) \right) + ((h-d) \cdot (T - T_{D-Im})) \quad (3.4.34)$$

siendo en este caso:

$$\Delta S_{N-D}(T_{N-D}) = \frac{\Delta H_{N-D}(T_{N-D})}{T_{N-D}} \quad (3.4.16)$$

$$\Delta S_{D-Im}(T_{D-Im}) = \frac{\Delta H_{D-Im}(T_{D-Im})}{T_{D-Im}} + R \cdot \ln\left(\frac{2 \cdot (1/2)^{1/n} \cdot P^{1/n}}{n^{1/n} \cdot P}\right) \quad (3.4.35)$$

Sabiendo que $\Delta G = \Delta H - T \cdot \Delta S$ y que $\Delta G = -RT \ln K$, se obtienen las constantes de ambos equilibrios y por tanto sus poblaciones. Por último, se define el cambio de entalpía global del proceso de desplegamiento como una sumatoria de cada una de las entalpías de los diferentes estados, ponderadas por su correspondiente población, luego:

$$\Delta H = \Delta H_{N-D} \cdot f_D + (\Delta H_{N-D} + \Delta H_{D-Im}) \cdot f_I \quad (3.4.36)$$

y la capacidad calorífica del proceso global, que es la que se mide experimentalmente en el calorímetro, como la capacidad calorífica del estado nativo más la variación de entalpía respecto de la temperatura a lo largo del proceso, $C_p = C_{pN} + d\Delta H/dT$. Esta función fue, por tanto, la que se utilizó para el análisis no lineal por mínimos cuadrados a los datos experimentales, C_p vs T , que se obtienen del experimento calorimétrico. Los parámetros a determinar a partir de dichos análisis serían, a priori, ΔH_{N-D} y ΔH_{D-Im} entalpías de desplegamiento y de asociación, respectivamente; y T_{N-D} y T_{D-Im} , temperaturas a las que las constantes de equilibrio del desplegamiento y de asociación se igualan a la unidad. También se han realizado análisis en esta Memoria en los que se ha dejado n , la

estequiometría del intermedio, como un valor flotante. Cuando se hicieron análisis globales en los que se usaron varias trazas que se realizaron a distintas concentraciones de proteína, se estableció una concentración de referencia de 100 μM , que correspondería a un experimento hipotético realizado a dicha concentración.

3.4.2.3 Modelo de equilibrio de cuatro estados con asociación del intermedio

El modelo presenta en total cuatro estados conformacionales en el equilibrio. Aparte de los estados nativo y desplegado, hay un estado intermedio entre ambos, el cual a su vez se asocia con una estequiometría concreta, siendo el cuarto posible estado conformacional. El modelo sigue el siguiente esquema:



Así, hay en total tres procesos en equilibrio. Por un lado, está el paso del estado nativo al intermedio $\text{N} \rightleftharpoons \text{I}$, cuya constante es K_{N-I} ; por otro, el desplegamiento completo del dominio, $\text{I} \rightleftharpoons \text{D}$, cuya constante es K_{I-D} y finalmente está el equilibrio de asociación, $\text{I} \rightleftharpoons (1/n) \text{I}_n$, cuya constante se va a denominar K_{I-I_n} . Al considerar de esta forma los equilibrios, se pueden comparar los valores termodinámicos obtenidos por los análisis, ya que la constante K_{I-I_n} está describiendo un proceso de asociación. De la misma manera, el producto de las constantes $K_{N-I} \cdot K_{I-D}$ equivale a una constante de desplegamiento típica de dos estados, el paso del nativo al desplegado.

Una vez determinadas qué especies coexisten en equilibrio, pueden definirse matemáticamente las poblaciones de dichos estados conformacionales. De esta forma, se pueden calcular las poblaciones de los estados nativo, intermedio, intermedio asociado y estado desplegado (las cuales se denominan f_N , f_I , f_{I_n} y f_D respectivamente), en función de la concentración total de proteína, P , mediante las expresiones 3.4.23 y 3.4.24, junto con las siguientes:

$$f_I = \frac{[I]}{P} \rightarrow [I] = f_I \cdot P \quad (3.4.37)$$

$$f_{I_n} = \frac{n[I_n]}{P} \rightarrow [I_n] = f_{I_n} \cdot \frac{P}{n} \quad (3.4.38)$$

Como en todo momento, la suma de las poblaciones de los estados conformacionales ha de ser la unidad ($f_N + f_I + f_D + f_{I_n} = 1$), se puede deducir la siguiente expresión:

$$[I_n] = (1 - f_N - f_D - f_I) \cdot \frac{P}{n} \quad (3.4.39)$$

Las constantes de equilibrio se pueden definir, por tanto, en función de las poblaciones, de la estequiometría de asociación n y de la concentración de proteína P como:

$$K_{N-I} = \frac{[I]}{[N]} = \frac{f_I}{f_N} \quad (3.4.40)$$

$$K_{I-D} = \frac{[D]}{[I]} = \frac{f_D}{f_I} \quad (3.4.41)$$

$$K_{I-In} = \frac{[I_n]^{1/n}}{[I]} \quad (3.4.42)$$

Si en esta última expresión sustituimos el valor de la concentración del intermedio asociado se obtiene lo siguiente:

$$K_{I-In} = \frac{[I_n]^{1/n}}{[I]} = \frac{\left(\frac{f_{In} \cdot P}{n}\right)^{(1/n)}}{f_I \cdot P} \rightarrow K_{I-In} = \frac{f_{In}^{1/n} \cdot P^{1/n}}{f_I \cdot n^{1/n} \cdot P}$$

Al reordenar esta expresión, nos quedaría:

$$f_{In}^{1/n} = K_{I-In} \cdot f_I \cdot n^{1/n} \cdot \frac{P}{P^{1/n}} \rightarrow f_{In} = K_{In}^n \cdot f_I^n \cdot n \cdot P^{n-1}$$

Ya que $f_{In} = (1 - f_N - f_D - f_I)$, y teniendo en cuenta que $f_N = f_I / K_{N-I}$ y que $f_D = K_{I-D} \cdot f_I$, nos queda la siguiente ecuación con una única incógnita, f_I :

$$n \cdot K_{I-In}^n \cdot K_{N-I} \cdot P^{n-1} \cdot f_I^n + (1 + K_{N-I} + K_{N-I} \cdot K_{I-D}) \cdot f_I - K_{N-I} = 0 \quad (3.4.43)$$

Con la resolución de esta ecuación, similar a la 3.4.28, que se obtenía para el modelo de tres estados con asociación del intermedio, se puede deducir toda la información termodinámica del sistema.

Tal y como ocurre en la Sección anterior, para obtener el valor de f_I hay que resolver una ecuación cuya resolución no se puede hacer de forma directa. Para resolver esta situación análoga se ha acudido al método de Newton, explicado en dicha Sección.

Para completar el modelo con sus fórmulas matemáticas, se definen las capacidades caloríficas de los cuatro estados conformacionales en función de la temperatura según las ecuaciones:

$$C_{pN} = a + b \cdot T, \text{ del estado nativo} \quad (3.4.11)$$

$$C_{pD} = c + d \cdot T, \text{ del estado desnaturalizado} \quad (3.4.29)$$

$$C_{pIn} = h + i \cdot T \text{ , del estado intermedio asociado} \quad (3.4.30)$$

$$C_{pI} = p + q \cdot T \text{ , del estado intermedio} \quad (3.4.43)$$

A continuación definen las entalpías de todos los procesos, teniendo en cuenta que esta función termodinámica es la derivada del cambio de capacidad calorífica respecto de la temperatura. Para ello se toman como temperaturas de referencia T_{N-I} , T_{I-D} y T_{I-In} a las que corresponden las entalpías $\Delta H_{N-I}(T_{N-I})$, $\Delta H_{I-D}(T_{I-D})$ y $\Delta H_{I-In}(T_{I-In})$, respectivamente:

$$\Delta H_{N-I} = \Delta H_{N-I}(T_{N-I}) + ((p - a) \cdot (T - T_{N-I})) + \left(\frac{(q - b) \cdot (T^2 - T_{N-I}^2)}{2} \right) \quad (3.4.44)$$

$$\Delta H_{I-D} = \Delta H_{I-D}(T_{I-D}) + ((p - c) \cdot (T - T_{I-D})) + \left(\frac{(q - d) \cdot (T^2 - T_{I-D}^2)}{2} \right) \quad (3.4.45)$$

$$\Delta H_{I-In} = \Delta H_{I-In}(T_{I-In}) + ((h - p) \cdot (T - T_{I-In})) + \left(\frac{(i - q) \cdot (T^2 - T_{I-In}^2)}{2} \right) \quad (3.4.46)$$

Se obtienen también los cambios de entropía de manera similar, considerando que $dS = dH/T = C_p \cdot dT/T$; llegando a las siguientes expresiones:

$$\Delta S_{N-I} = \Delta S_{N-I}(T_{N-I}) + \left((p - a) \cdot \ln\left(\frac{T}{T_{N-I}}\right) \right) + ((q - b) \cdot (T - T_{N-I})) \quad (3.4.47)$$

$$\Delta S_{I-D} = \Delta S_{I-D}(T_{I-D}) + \left((p - c) \cdot \ln\left(\frac{T}{T_{I-D}}\right) \right) + ((q - d) \cdot (T - T_{I-D})) \quad (3.4.48)$$

$$\Delta S_{I-In} = \Delta S_{I-In}(T_{I-In}) + \left((h - p) \cdot \ln\left(\frac{T}{T_{I-In}}\right) \right) + ((i - q) \cdot (T - T_{I-In})) \quad (3.4.49)$$

En este caso, $\Delta S_{N-I}(T_{N-I})$, $\Delta S_{I-D}(T_{I-D})$ y $\Delta S_{I-In}(T_{I-In})$ se definen como:

$$\Delta S_{N-I}(T_{N-I}) = \frac{\Delta H_{N-I}(T_{N-I})}{T_{N-I}} \quad (3.4.50)$$

$$\Delta S_{I-D}(T_{I-D}) = \frac{\Delta H_{I-D}(T_{I-D})}{T_{I-D}} \quad (3.4.51)$$

$$\Delta S_{I-In}(T_{I-In}) = \frac{\Delta H_{I-In}(T_{I-In})}{T_{I-In}} + R \cdot \ln\left(\frac{2 \cdot (1/2)^{1/n} \cdot P^{1/n}}{n^{1/n} \cdot P}\right) \quad (3.4.52)$$

Sabiendo que $\Delta G = \Delta H - T \cdot \Delta S$ y que $\Delta G = -RT \ln K$, se obtienen las constantes de ambos equilibrios y por tanto sus poblaciones. Por último, se define el cambio de entalpía global del proceso de desplegamiento como una sumatoria de cada una de las entalpías de los diferentes estados, ponderadas por su correspondiente población, luego:

$$\Delta H = \Delta H_{N-I} \cdot f_I + (\Delta H_{N-I} + \Delta H_{I-D}) \cdot f_D + (\Delta H_{I-D} + \Delta H_{I-In}) \cdot f_{In} \quad (3.4.53)$$

La capacidad calorífica del proceso global, que es la que se mide experimentalmente en el calorímetro es la capacidad calorífica del estado nativo más la variación de entalpía respecto de la temperatura a lo largo del proceso, $C_p = C_{pN} + d\Delta H/dT$. Esta función fue, por tanto, la que se utilizó para el análisis no lineal por mínimos cuadrados a los datos experimentales, C_p vs T , que se obtienen del experimento calorimétrico. Los parámetros a determinar a partir de dichos análisis son los valores de entalpía $\Delta H_{N-I}(T_{N-I})$, $\Delta H_{I-D}(T_{I-D})$ y $\Delta H_{I-In}(T_{I-In})$ y las respectivas temperaturas T_{N-I} , T_{I-D} y T_{I-In} . También se han realizado análisis en esta Memoria en los que se ha dejado n, la estequiometría del intermedio, como un valor flotante. Cuando se hicieron análisis globales en los que se usaron varias trazas que se realizaron a distintas concentraciones de proteína, se estableció una concentración de referencia de 100 μM , que correspondería a un experimento hipotético realizado a dicha concentración.

3.5 Dicroísmo Circular (CD)

La espectroscopia de dicroísmo circular (*Circular Dichroism*, CD) es una de las técnicas más empleadas en el estudio de las propiedades conformacionales de proteínas y ácidos nucleicos en disolución (Johnson, 1990). El fenómeno del dicroísmo circular se basa en la diferencia de absorción de la luz circularmente polarizada hacia la derecha e izquierda por parte de las moléculas asimétricas que son ópticamente activas, también llamadas quirales. Cada uno de los tipos de luz circularmente polarizada (a derecha e izquierda) sigue la ley de Lambert-Beer. Según esto, un espectro de dicroísmo para una muestra de un compuesto dado puede definirse inicialmente mediante la diferencia entre las absorbancias que presenta para las dos formas de luz polarizada, en función de la longitud de onda, λ :

$$\Delta Abs(\lambda) = Abs_I(\lambda) - Abs_D(\lambda) = [\varepsilon_I(\lambda) - \varepsilon_D(\lambda)] \cdot c \cdot l \quad (3.4.54)$$

$$\Delta Abs(\lambda) = \Delta \varepsilon \cdot c \cdot l \quad (3.4.55)$$

donde Abs_I y Abs_D se refieren a las absorbancias correspondientes a la muestra para la luz circularmente polarizada a la izquierda y a la derecha respectivamente, c es la concentración de la muestra expresada en $\text{mol} \cdot \text{L}^{-1}$, l es la longitud del paso de luz de la cubeta en cm, y ε_I y ε_D los coeficientes de extinción molar en $\text{L} \cdot \text{mol}^{-1} \cdot \text{cm}^{-1}$ para cada tipo de luz.

Las bandas de un espectro de dicroísmo circular pueden ser negativas o positivas, dependiendo de qué tipo de luz se absorba más, y sólo se dan a longitudes de onda correspondientes a las bandas de absorción. Actualmente, los instrumentos que se utilizan

para realizar esta espectroscopia reciben el nombre de espectropolarímetros y miden de forma directa la diferencia de absorbancia $\Delta Abs(\lambda)$. Cuando la luz polarizada plana incide sobre una muestra, las dos componentes circularmente polarizadas que la componen se absorben en diferente magnitud. La señal es proporcional a la excentricidad de la elipse que caracteriza la luz polarizada resultante. La elipticidad, θ , es una medida angular de esta excentricidad y constituye la forma original para detectar la medida. La mayoría de los instrumentos comerciales están calibrados en unidades de elipticidad (θ). Generalmente los espectros se expresan en términos de elipticidad molar, $[\theta]$, en grados·dmol⁻¹·cm². Su relación con la diferencia entre los coeficientes de extinción molar viene dada por la relación:

$$[\theta] = 3298 \cdot \Delta \varepsilon \quad (3.4.56)$$

que incluye un factor de 100 debido a la conversión de la concentración en mol·L⁻¹ a dmol·cm⁻³.

En todos los experimentos de dicroísmo circular recogidos en esta Memoria las medidas de elipticidad se han normalizado generalmente a unidades de elipticidad molar media por residuo ($[\theta]$) en (grado·dmol⁻¹·cm²), para lo cual se utiliza la siguiente ecuación:

$$[\theta] = \frac{\theta}{\frac{C}{PM} \cdot l \cdot 10 \cdot n} \quad (3.4.57)$$

donde θ es la elipticidad medida experimentalmente en milésimas de grado, C es la concentración de la proteína en mg·mL⁻¹, l es paso de luz de la célula en cm, PM es el peso molecular de la proteína en g·mol⁻¹ y n es el número de residuos de la proteína. El factor de 10 es el resultado de convertir las unidades de concentración de mg·mL⁻¹ a dmol·cm⁻³ y la elipticidad a grados.

Mediante esta espectroscopía podemos obtener mucha información acerca de las propiedades conformacionales de las moléculas biológicas, más concretamente, podemos determinar si una proteína se encuentra plegada, y en tal caso, podemos caracterizar su estructura secundaria (ultravioleta lejano), estructura terciaria (ultravioleta cercano) y clasificarla dentro de una familia estructural. Asimismo, se puede estudiar la estabilidad conformacional de una proteína (estabilidad térmica, frente a cambios de pH o en presencia de agentes desnaturalizantes) y cómo esta estabilidad se altera dependiendo de las condiciones experimentales.

3.5.1 Dicroísmo circular en la zona del UV-lejano (170-250 nm)

Los espectros en esta región son muy sensibles a los cambios de estructura secundaria de las proteínas, mientras que la contribución de la estructura terciaria a la señal registrada en estas longitudes de onda suele ser pequeña. A estas longitudes de onda el cromóforo es el enlace peptídico. Aunque el enlace peptídico no es intrínsecamente una estructura quiral, la interacción entre grupos peptídicos adyacentes

dispuestos en determinadas orientaciones hace que las proteínas presenten actividad óptica en la región del ultravioleta lejano. Cada elemento principal de estructura secundaria de proteínas: hélices α , láminas β , giros β y estructuras aleatorias presenta un espectro de dicroísmo característico (Compton & Johnson, 1986). En la Figura 3.2 se representan los espectros típicos de cada uno de ellos y se observa cómo estos espectros difieren entre sí lo suficiente como para determinar el contenido de estas estructuras en la proteína. Las hélices α presentan un espectro muy característico en la región del ultravioleta lejano con un doble mínimo a 222 y a 208-210 nm y un máximo más intenso a 191-193 nm. Las proteínas con estructura β suelen presentar un único pico negativo y un único pico positivo cuyas intensidades son menores que las de las hélices. El pico positivo es de mayor intensidad que el negativo y aparece entre 190 y 200 nm, mientras que el mínimo se encuentra normalmente entre 210 y 225 nm. Las posiciones e intensidades de estas dos bandas varían mucho en función de la muestra.

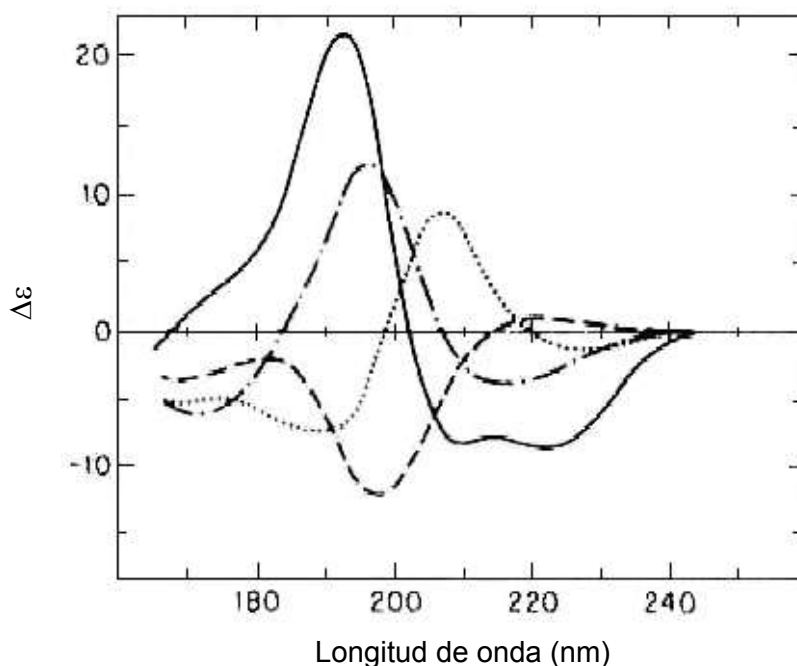


Figura 3.2. Espectros representativos de los cuatro tipos principales de estructuras secundarias desde 178 hasta 260 nm: Hélice alfa (—), hoja beta (—●—●—●—), giro beta (●●●●●●) y otra estructuras aleatorias (-----), tomado de la referencia (Brahms, 1980).

Hay que recordar, sin embargo, que en el ultravioleta lejano también absorben los aminoácidos aromáticos y que, aunque sus cadenas laterales no son quirales, se comportan como tales cuando así lo propicia su entorno proteico. Por esta razón el espectro de algunas proteínas aparece dominado por contribuciones peculiares debidas a estos residuos. Con esta salvedad, la actividad óptica de los enlaces peptídicos es de gran utilidad para conocer el contenido en estructura secundaria de las proteínas. Así, cada

proteína tiene un espectro de dicroísmo en el ultravioleta lejano que es, aproximadamente, una combinación lineal de los espectros característicos de las estructuras secundarias y no regulares que contiene (Woody, 1994).

3.5.2 Obtención de espectros de dicroísmo circular

Los experimentos se han llevado a cabo en un espectropolarímetro JASCO modelo J-715 equipado con un control de temperatura del compartimento de células. En varias tesis doctorales del grupo de investigación Biofísica y Biotecnología Molecular de la Universidad de Granada (Sadqi, 2000; Cobos, 2002) se han explicado ampliamente todos los fundamentos y características del instrumento, por lo que nos limitamos a recomendar estas referencias para un conocimiento más detallado. Una vez preparada la disolución de proteína, como se ha indicado en la Sección 3.3, ésta se introduce en una cubeta de cuarzo de paso de luz adecuado, que se escoge en función de la concentración de la muestra y del intervalo de longitud de onda seleccionado para las medidas que optimizan la relación señal/ruido de la muestra. Para eliminar las contribuciones del tampón y del instrumento se registra asimismo el espectro con el tampón de diálisis (blanco) que se resta del espectro de la proteína.

Los espectros que se han obtenido en esta Memoria, en la región del ultravioleta lejano, se han recogido en el intervalo de longitudes de onda entre 200 y 260 nm utilizando los siguientes parámetros que proporcionan una buena relación señal-ruido: resolución 0.5 nm, ancho de banda 0.5 nm, tiempo de respuesta 1 segundo y velocidad de barrido 50 nm·min⁻¹. Se acumulan 3 espectros consecutivos obteniéndose finalmente un espectro promedio del total de espectros realizados. El paso de luz de la cubeta ha de ser 0.1 cm para los experimentos a baja concentración (30 μM) y de 0.01 cm para los de alta concentración (720 μM).

Cuando se llevó a cabo la desnaturalización térmica de la proteína, la longitud de onda seleccionada fue 218 nm (θ_{218}), aquella en la que la diferencia de señal observada entre nuestra proteína nativa y desnaturalizada es máxima (ver Capítulo 5). Los barridos de temperatura se han realizado calentando a una velocidad de barrido de 1°C·min⁻¹, con una resolución de 0.1°C y un tiempo de respuesta de 4 segundos. Antes de comenzar el barrido de temperatura, se equilibra la muestra en la cubeta de medida durante 10 minutos a 2°C, registrándose después un espectro a dicha temperatura. Una vez alcanzados los 98°C, se adquiere entonces el espectro correspondiente del estado desplegado, y se enfría rápidamente hasta 2°C. Se han obtenido también espectros a los 10 minutos, a la hora y las 24 horas después de enfriar hasta dicho valor de temperatura con el objeto de comprobar la reversibilidad del proceso de desplegamiento.

3.5.3 Análisis de las curvas de desplegamiento térmico según la termodinámica de equilibrio

El desplegamiento térmico seguido por CD ha sido analizado en esta Memoria con el modelo de equilibrio de dos estados. Para ello es necesario definir en primer lugar las líneas base de los estados nativo y desnaturalizado. Estos dos estados se caracterizan por tener una variación de elipticidad con la temperatura prácticamente lineal, definida por dos funciones lineales, $[\theta]_N$ y $[\theta]_D$, para los estados nativo y desnaturalizado, respectivamente según las ecuaciones:

$$[\theta]_N = \theta_{N,0} + m_N \cdot T \quad (3.4.58)$$

$$[\theta]_D = \theta_{D,0} + m_D \cdot T \quad (3.4.59)$$

donde $\theta_{N,0}$ y m_N , por una parte y $\theta_{D,0}$ y m_D , por otra, representan las respectivas ordenadas en el origen y pendientes. El observable experimental elipticidad en función de la temperatura viene dado por la ecuación:

$$[\theta] = [\theta]_N \cdot f_N + [\theta]_D \cdot f_D \quad (3.4.60)$$

en que las poblaciones f_N y f_D , están determinadas por el valor de una constante de equilibrio entre los dos estados definidos K_{N-D} , definida de forma análoga a la de la ecuación 3.4.2 descrita en la Sección 3.4.2.1, para el modelo de equilibrio de dos estados. Así, se puede expresar f_D en función de dicha constante según la expresión:

$$f_D = \frac{[D]}{[D] + [N]} = \frac{K_{N-D}}{1 + K_{N-D}} \quad (3.4.61)$$

De esta forma, teniendo en cuenta las ecuaciones 3.4.60 y 3.4.61, se puede llegar a la ecuación:

$$[\theta] = \frac{[\theta]_N + [\theta]_D \cdot K_{N-D}}{1 + K_{N-D}} \quad (3.4.61)$$

Finalmente, teniendo en cuenta las relaciones termodinámicas 3.4.10, 3.4.14 y 3.4.15:

$$K_{N-D} = \exp\left(-\frac{\Delta G_{N-D}}{T}\right) \quad (3.4.10)$$

$$\Delta H_{N-D}(T) = \Delta H_{N-D}(T_{N-D}) + \int_{T_{N-D}}^T \Delta C_{pN-D} \cdot dT \quad (3.4.14)$$

$$\Delta S_{N-D}(T) = \Delta S_{N-D}(T_{N-D}) + \int_{T_{N-D}}^T \frac{\Delta C_{pN-D}}{T} \cdot dT \quad (3.4.15)$$

se puede obtener la información termodinámica del sistema aplicando todas estas ecuaciones mediante un proceso de ajuste no lineal de mínimos cuadrados sobre las curvas experimentales obtenidas, asumiendo el caso más sencillo en que el valor de $\Delta C_{p_{N-D}}$ es independiente del cambio de temperatura.

3.6 Dispersión dinámica de la luz (DLS)

La dispersión dinámica de la luz (*Dynamic Light Scattering*, DLS) es una técnica óptica que da información sobre el tamaño molecular de las partículas presentes en una disolución. El fundamento de esta técnica consiste en que la intensidad de la luz dispersa por una disolución fluctúa rápidamente con el tiempo. Esta fluctuación se debe a que, gracias al movimiento Browniano de las partículas, se producen variaciones en la distancia entre ellas, lo que provoca que haya interferencias entre las ondas de luz dispersadas por cada una de las partículas. Precisamente, mediante un instrumento de dispersión dinámica de la luz se obtiene la velocidad de esa fluctuación.

Esto se expresa matemáticamente mediante una función de correlación, que se obtiene midiendo la similitud entre una señal adquirida en un tiempo y la misma señal medida un cierto tiempo después. Más en concreto, hay dos funciones de correlación correspondientes a dos propiedades de la disolución, una de las cuales se obtiene a partir del flujo de fotones, $i(t)$ (ecuación 3.4.56) y la otra es la transformada de Fourier del campo magnético propio de la radiación electromagnética, $B(t)$ (ecuación 3.4.57).

Estas dos funciones de correlación se combinan para dar una función final, que tiene en cuenta la diferencia de fase que hay entre ellas (Pecora, 1986). Esta función final de correlación consiste en una exponencial simple cuando hay un único tamaño de partícula en la disolución, o una exponencial múltiple cuando hay heterogeneidad de tamaños.

$$G^2(\tau) = \langle i(t)i(t+\tau) \rangle = \int_0^{\infty} i(t)i(t+\tau)dt \quad (3.4.62)$$

$$G^1(\tau) = \langle B(t)B^*(t+\tau) \rangle \quad (3.4.63)$$

donde G^1 y G^2 representan las dos funciones de correlación, $i(t)$ representa el flujo de fotones y $B(t)$ es el campo magnético en un tiempo dado, siendo τ el tiempo de retardo entre las adquisiciones de ambas propiedades (i y B) a dos tiempos distintos. Cuando $\tau = 0$, G^1 y G^2 están completamente en fase, por lo que la función de correlación final es máxima. Cuando τ aumenta, G^1 y G^2 dejan de estar en fase, por lo que el valor de la función disminuye.

Con este análisis de la relación entre las fluctuaciones de intensidad y el tiempo, se puede obtener el coeficiente de difusión traslacional de las partículas, D . A partir de este coeficiente es posible calcular el radio hidrodinámico aparente de las partículas, suponiendo que éstas son esféricas, mediante la ecuación de Stokes-Einstein:

$$D = \frac{k_B \cdot T}{6 \cdot \pi \cdot \eta \cdot R_H} \quad (3.4.64)$$

en la que k_B es la constante de Boltzmann, T es la temperatura absoluta, η es la viscosidad de la disolución y R_H es el radio hidrodinámico de la partícula. Así, mediante una calibración basada en patrones de peso molecular conocido, se puede obtener información del peso molecular de la partícula. Esta calibración ya viene dada con el software del instrumento.

Las medidas de dispersión dinámica de la luz se llevaron a cabo en un instrumento DynaPro MS-X (Wyatt Technology Corporation, Santa Barbara, CA, USA) equipado con un control Peltier de temperatura. Para saber cómo funciona un aparato de DLS, ya ha habido una Tesis Doctoral del Departamento en el que se ha descrito con más detalle (Varela, 2010).

Para hacer las medidas, las disoluciones de proteína y los tampones se centrifugaron en una microcentrífuga y se filtraron a través de un filtro Anotop de $0.02 \mu\text{m}$ (Whatman plc, Brentford, Middlesex, UK) para eliminar cualquier partícula que pudiese interferir en la medida. Las medidas se hicieron en una cubeta de cuarzo (con $30 \mu\text{L}$ de capacidad). Una vez tomada la medida a la temperatura de 20°C , se hizo un barrido hasta 60°C , temperatura en la que se dejó midiendo hasta la saturación de señal, ya que cuando hay procesos de agregación la señal va aumentando hasta saturación. En el caso en que esto no ocurrió, se dejó el tiempo que se estimó oportuno. A continuación se bajó la temperatura hasta el valor de 20°C y se tomaron medidas desde un minuto después hasta un día después de haber alcanzado dicha temperatura para ver si se recuperaba el estado inicial. Antes de empezar el experimento, se ajustó la potencia del láser a las características de la muestra, para evitar una pronta saturación del instrumento.

Para recoger y analizar los datos se usó el programa Dynamics V6 (Wyatt Technology Corporation, Santa Barbara, CA, USA), que permite obtener las distribuciones de los radios hidrodinámicos en la disolución durante el proceso de agregación. Además, el peso molecular de una partícula puede estimarse a partir de datos de DLS usando curvas de calibración establecidas a partir de patrones de pesos moleculares conocidos frente a sus radios hidrodinámicos. Sin embargo, debido a que el volumen específico de una partícula depende de su conformación terciaria, no se ha desarrollado aún una curva de calibración universal. Sin embargo, sí que hay disponibles numerosas curvas, correspondientes cada una a un tipo de macromolécula. De esta forma, hay que asumir para la muestra estudiada un tipo de conformación. Las curvas de calibración de que dispone el programa Dynamics son *Proteínas*, que tiene forma globular; *Pululanos*, que son polisacáridos lineales; *Ficoles*, que son polisacáridos muy ramificados y *Dendrímeros*, que son polímeros esféricos con una densidad que aumenta con la distancia radial desde el centro de la partícula. En el caso de esta Memoria se usó el de *Proteínas*.

3.7 Absorbancia de Congo Red (CR)

El Congo Red (CR) es un tinte que une de una forma relativamente específica a fibras amiloides. El espectro de absorbancia del CR cambia al unirse a una fibra amiloide, dándose un desplazamiento del máximo de absorbancia hacia longitudes de onda

mayores. El cambio espectral que se produce después de la unión se debe al apilamiento de varias moléculas de CR en una determinada disposición a lo largo de las fibras.

La unión a CR depende de la estructura del amiloide, y lo hace específicamente a la conformación en la lámina beta. Aquellas fibras amiloides a las que le faltan o contienen únicamente una proporción mínima de lámina beta no se tiñen con CR. Sin embargo, la estructura en lámina beta tampoco es suficiente para la unión con CR, puesto que no se conoce ninguna proteína globular soluble rica en lámina beta que una CR.

Aunque la unión de CR a fibras es un hecho bastante aceptado y con muchas referencias en bibliografía (Puchtler & Sweat, 1965; Klunk et al., 1999; Nilsson, 2004), el hecho de que el CR une también a determinadas conformaciones parcialmente desplegadas de varias proteínas, debe usarse con cierta cautela como una prueba de presencia de fibras amiloides in vitro. Por otro lado, es bien conocido que el CR es un inhibidor de formación de fibras amiloides, se pueden encontrar ejemplos en bibliografía (Klunk et al., 1989; Caughey et al., 1993; Lorenzo & Yankner, 1994), por lo que no se puede usar para seguir la formación de las fibras a tiempo real (Khurana et al., 2001).

3.8 Fluorescencia de Thioflavina T y ANS

La fluorescencia es un fenómeno que consiste en la emisión de radiación electromagnética cuando se pasa de un estado excitado producido por la absorción de radiación electromagnética al estado fundamental. Aparte de la fluorescencia intrínseca de proteínas, debida a los aminoácidos fenilalanina, tirosina y triptófano, es posible estudiar cambios conformacionales de proteínas midiendo la fluorescencia de sondas o compuestos fluorescentes que interactúan específicamente con ciertos estados o conformaciones de la proteína.

En esta Memoria se han usado dos de estos compuestos, la Thioflavina T (ThT) (Sigma), que interactúa específicamente con fibras amiloides aumentando considerablemente su fluorescencia (LeVine III & Ronald, 1999), y el ácido 8-Anilino-naftalén-sulfónico (ANS) (Fluka), que es un compuesto aromático que interactúa específicamente con regiones hidrofóbicas expuestas al disolvente dando lugar a un aumento de su fluorescencia.

La estrategia para los experimentos de fluorescencia realizados en esta Tesis Doctoral ha sido paralela a la seguida con los experimentos de DLS. Se realizó inicialmente un barrido de temperatura desde 20°C hasta 60°C de una disolución de proteína en presencia de la sonda fluorescente correspondiente, para dejar la muestra a continuación incubando a 60°C hasta saturación de la señal debida a la formación de agregados o bien hasta que se estimase conveniente. A continuación se bajó la temperatura hasta el valor de 20°C y se registraron varios espectros desde un minuto después hasta tres días después de haber alcanzado dicha temperatura para ver si se recuperaba la fluorescencia inicial.

Estos experimentos se han realizado en un espectrofluorímetro Cary Eclipse equipado con un sistema Peltier de control de temperatura (Varian Inc.). Este aparato está compuesto de una lámpara de Xenon-Argón, una cámara de muestra y un detector de

diodos. Los caminos ópticos de excitación y de detección están en una disposición ortogonal, lo cual asegura el mínimo paso de luz de excitación a la zona de detección. A su vez, dispone de monocromadores en ambos caminos ópticos para seleccionar las longitudes de onda o región del espectro de interés.

3.8.1 Ensayos con ThT

Los ensayos con ThT se realizaron para medir la cinética de formación de agregados (Nilsson, 2004). La fluorescencia de ThT se considera aproximadamente proporcional a la masa de fibras amiloides, y es una de las sondas más usadas en el estudio de la cinética de formación de fibras (LeVine lii & Ronald, 1999).

Para los experimentos, la proteína dializada en el tampón correspondiente se mezcló con una disolución de ThT preparada en el mismo tampón para tener las concentraciones finales deseadas de proteína (30 μM o 720 μM) y de ThT (12,5 μM). La cubeta usada fue de 1 cm de paso de luz. Los espectros de emisión se tomaron entre 460 y 600 nm (con un ancho de banda de 5 nm), siendo la longitud de onda de excitación de 440 nm (ancho de banda de 10 nm). Tal y como se ha mencionado, se realizó un barrido de temperatura desde 10°C hasta 60°C y luego se mantuvo a esa temperatura durante varias horas, registrando espectros cada media hora, siendo 485 nm la longitud de onda seleccionada para seguir la cinética.

También se registraron espectros de muestras con tiempos más largos de incubación, desde pocos días hasta un mes. Para ello, después de la incubación de las muestras, que se realizó tal y como se describe en la Sección 3.9, se tomaron 10 μL de la muestra y se llevaron hasta un volumen final de 1 mL, con las mismas condiciones de tampón y una concentración de ThT de 12,5 μM . Los espectros se registraron en las mismas condiciones que cuando se hizo el barrido de temperatura.

3.8.2 Ensayos con ANS

El ANS es un compuesto que se une a zonas no polares de proteínas a través de su grupo anilinaftaleno, por lo cual es bastante usado para identificar y caracterizar estados parcialmente plegados en proteínas (Azuaga et al., 2002), (Buchner & Kiefhaber, 2005), ya que normalmente tienen expuestos restos hidrofóbicos que en la forma nativa no lo están.

Como en el caso de los ensayos con ThT, la proteína dializada en el tampón correspondiente se mezcló con una disolución de ANS en el mismo tampón para tener las concentraciones finales deseadas de proteína (30 μM o 720 μM) y de ANS (25 μM). La cubeta usada también fue de 1 cm de paso de luz. Los espectros de emisión se tomaron entre 400 y 600 nm (con un ancho de banda de 5 nm), siendo la longitud de onda de excitación de 370 nm (ancho de banda de 2,5 nm). Se llevó a cabo un barrido de temperatura desde 10°C hasta 60°C y luego se mantuvo a esa temperatura durante varias

horas, tomando espectros cada media hora, siendo 465 nm la longitud de onda seleccionada para seguir la cinética.

3.9 Microscopía electrónica de transmisión (TEM)

La microscopía electrónica de transmisión (*Transmission Electron Microscopy*, TEM) se basa en la interacción de un rayo de electrones con una muestra en la que se quieren visualizar partículas de tamaños que van del rango de algunos nm hasta varios μm . Esta técnica se usa ampliamente para obtener información morfológica sobre los agregados que aparecen durante el proceso de formación de amiloides.

Para que sea posible su visualización, las muestras que se deseen ver con este microscopio deben tener un tamaño que permita a los electrones atravesarlas de manera similar a cómo la luz es transmitida a través de la muestra en una microscopía óptica convencional. Ya que la longitud de onda de los electrones es mucho más pequeña que la de la luz, la resolución típica para imágenes de TEM es mejor en varios órdenes de magnitud que aquella de un microscopio óptico. Esto significa que con uno de estos microscopios se pueden revelar hasta los más pequeños detalles de la estructura interna, siendo posible en algunos casos ver detalles a nivel atómico.

La energía de los electrones determina su grado relativo de penetración en una muestra específica o bien influencia el tamaño a partir del cual la información útil puede obtenerse. Esa energía viene determinada por el voltaje de la fuente que los genera; los microscopios electrónicos convencionales suelen tener un voltaje de 80 a 200kV.

La dosis de electrones debe mantenerse baja, para reducir el daño producido a la muestra biológica debido a la radiación. Así, el contraste de una micrografía de muestras biológicas suele ser muy pobre y, por lo tanto, para mejorarlo se requiere una tinción negativa de las muestras. Las tinciones negativas están compuestas de metales pesados (como el acetato de uranilo, que ha sido el usado en esta Memoria) que dispersan los electrones y son opacos a estos, lo que resulta en una imagen oscura del tinte depositado alrededor de la fibra mostrando así la morfología de ésta.

La descripción de un microscopio electrónica ya ha sido hecha en profundidad en otra Tesis Doctoral del Departamento (Varela, 2010). En el caso de las micrografías recogidas en esta Memoria, se hicieron en un microscopio electrónico Zeiss 902 de un voltaje de 80 kV, que está ubicado en el Centro de Instrumentación Científica de la Universidad de Granada.

Las muestras observadas en esta Memoria se incubaron a las distintas temperaturas (60°C o 37°C, en función de los casos) en un horno, cuya temperatura se controló con un termómetro de mercurio. Para la observación de las distintas etapas del proceso a lo largo del tiempo, se incubaron a los distintos intervalos de tiempo estudiados (desde 2 horas hasta un mes) unos 100 μL de disolución de proteína a 720 μM en tampones con una concentración de 50 mM, bien fosfato a pH 7.5, o bien glicina a pH 3.

La preparación inmediata para medir en el microscopio consistió en tomar 20 μL de muestra previamente incubada a la temperatura deseada y situar encima una rejilla

recubierta de formvar (un polímero de polivinilo para sostener capas finas de muestra). Ésta se dejó a temperatura ambiente durante 4 minutos, para que se adsorbiesen los distintos componentes de la muestra. La rejilla se lavó a continuación dos veces con agua destilada y se trató durante un minuto con acetato de uranilo al 1% (peso-volumen). Después, las rejillas se secaron durante 5 minutos a 37°C y se observaron en el microscopio. Las imágenes se digitalizaron para poder evaluar los cambios morfológicos de los agregados durante el proceso de formación de fibras.

3.10 Cromatografía analítica de exclusión molecular (SEC)

La cromatografía de exclusión molecular (*Size Exclusion Chromatography*, SEC) es una cromatografía de partición utilizada para separar moléculas de diferentes tamaños. El principio básico consiste en que las moléculas se reparten entre el disolvente y una fase estacionaria de porosidad definida. El proceso de separación se consigue utilizando un gel poroso empaquetado en una columna (fase estacionaria) y rodeado de disolvente (fase móvil). Si consideramos una mezcla de moléculas de diferente tamaño que atraviesan el sistema, lo que ocurrirá es que las moléculas menores penetrarán en un mayor número de poros de la fase estacionaria y, por tanto, se desplazan más lentamente a través de la columna al tener un recorrido mayor; las moléculas mayores serán eluidas en primer lugar de la columna por este motivo (ver Figura 3.3). Para establecer de forma cuantitativa estas diferencias de velocidad dentro de la columna se suelen usar dos parámetros indistintamente, que están relacionados entre sí por la velocidad de flujo. Por un lado, el tiempo de retención, que es la diferencia de tiempo entre que se ha eluido la muestra y el que se ha inyectado; por otro, el volumen de retención, que es el volumen al que se eluye la muestra después de haberla inyectado. En esta Memoria se ha usado el volumen de retención. Así, el logaritmo neperiano del peso molecular de la partícula es directamente proporcional al volumen al que es retenida dicha partícula.

Este tipo de cromatografía, por consiguiente, permite la separación de proteínas en función de su peso molecular. Por tanto, es muy útil para el estudio de los equilibrios de asociación de macromoléculas en los que esté involucrado un cambio apreciable del tamaño molecular. Un requisito fundamental en este sentido es que el posible equilibrio entre ellas sea suficientemente lento en comparación con el tiempo de la separación cromatográfica; en caso contrario, obtendríamos un solo pico, ancho, con un tiempo de retención intermedio entre el peso molecular de las especies en equilibrio (ver Capítulo 5).

Los experimentos de cromatografía de exclusión se realizaron para estudiar los posibles estados de asociación del intermedio de plegamiento (ver Capítulos 5 y 7). La columna usada fue una Superdex 75 HR 10-30 (Pharmacia), el flujo con el que se hicieron los experimentos fue de $0,5 \text{ mL} \cdot \text{min}^{-1}$ y la cantidad de muestra que se inyectó para cada prueba fue de unos $100 \text{ } \mu\text{L}$. Para las pruebas a 60°C la columna termostatózó con un horno específico para columnas de la casa comercial Pharmacia. El cromatógrafo usado (Fast Protein Liquid Chromatography, FPLC) tiene conectado un ordenador con un programa que

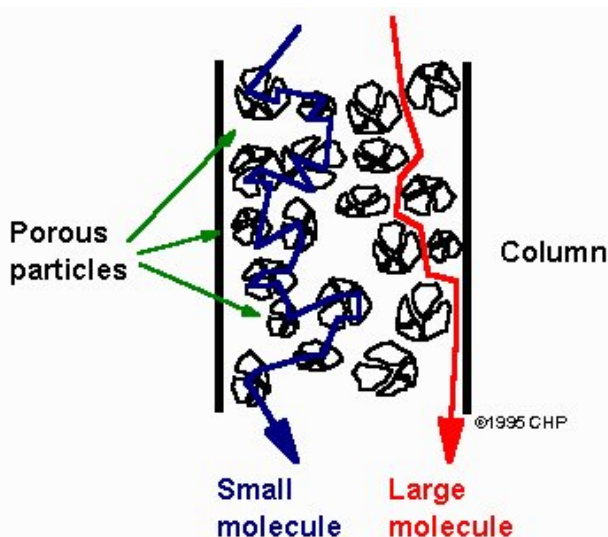


Figura 3.3. Representación esquemática del funcionamiento de una columna de exclusión molecular. La mezcla avanza a través de la columna a través de los poros de la resina. Las moléculas más grandes atravesarán menos poros, y así eluyen más rápidamente, mientras que las más pequeñas tendrán un recorrido mayor, teniendo por tanto un volumen de retención también mayor.

controla el instrumento con el que se obtuvieron de forma automática los volúmenes de retención.

Para poder estimar los pesos moleculares de las distintas especies, se llevó a cabo una calibración de la columna obteniendo los volúmenes de retención, V_R , de distintos patrones cuyos pesos moleculares, PM , están ya previamente establecidos. V_0 es el volumen de exclusión total, correspondiente al volumen de retención del azul dextrano, que tiene un peso molecular tan elevado (del orden de 10^6 Da), que no atraviesa ningún poro de la columna, por lo que, en su elución, el volumen de retención es el que se corresponde únicamente al recorrer la columna. En la Tabla 3.1 se muestran los resultados de la calibración a 60°C , con los que se construyó la recta de calibración representando el logaritmo neperiano de los pesos moleculares de los patrones, $\ln(PM)$, frente a V_R/V_0 (Figura 3.4).

Figura 3.1. Volúmenes de retención a los que eluyen los patrones de proteína cuyos pesos moleculares están bien caracterizados en la columna Superdex 75 HR 10-30 termostatzada a 60°C .

PATRÓN	PM (Da)	V_R	V_R/V_0	$\ln(PM)$
Azul dextrano	$2 \cdot 10^6$	7.47	-	-
Albúmina de sero bovino	66000	8.88	1.18	11.09
Pepsina	35000	9.77	1.31	10.46
Mioglobina	17000	12.07	1.62	9.74
Citocromo C	12400	12.63	1.69	9.42
Apoptotina	6500	14.72	1.97	8.78

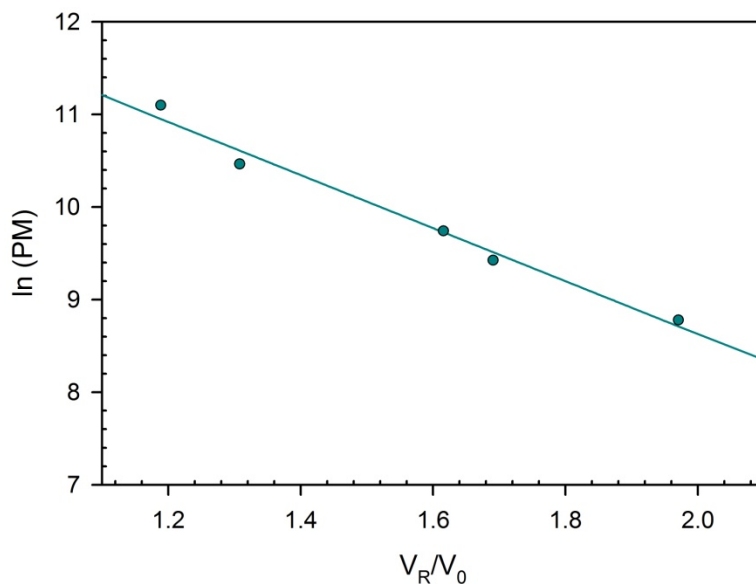


Figura 3.4. Representación de los valores del logaritmo neperiano de los pesos moleculares de los patrones de calibración frente a la relación sus volúmenes de retención con el volumen de exclusión total en la columna Superdex 75 HR 10-3. Los puntos son los datos experimentales, mientras que la línea corresponde a la recta que mejor se ajusta a los puntos, cuya ecuación es $\ln(PM) = 14.36 - 2.86 \cdot (V_R/V_0)$.

3.11 Calorimetría isotérmica de titulación (ITC)

La calorimetría isotérmica de titulación (*Isothermal Titration Calorimetry*, ITC) es una técnica que nos permite medir directamente el calor asociado a un proceso de interacción, el cual, a presión constante, equivale a la entalpía de este proceso (Wiseman et al., 1989). Si el efecto térmico neto es lo suficientemente grande, se pueden determinar parámetros tales como la constante de equilibrio de asociación, K_d , el cambio de entalpía de unión, ΔH_d , así como la estequiometría de la interacción, n . Como consecuencia, se pueden calcular la energía de Gibbs y la entropía de formación del complejo (ΔG_d , ΔS_d). Además, si se realizan experimentos a diferentes temperaturas es posible determinar el cambio en la capacidad calorífica del proceso, ΔC_{pd} .

Los experimentos de titulación realizados en esta Memoria se hicieron a 25°C, utilizando un microcalorímetro isotérmico de titulación MCS-ITC (Microcal INC, Northampton, MA). Las disoluciones de proteína ya dializadas se filtraron, se desgasificaron para evitar la formación de burbujas y se equilibraron a la temperatura experimental. Debido a la baja afinidad de la interacción, en el caso del experimento del dominio PDZ3-h-PSD95 con el péptido de secuencia KKETAV, los experimentos de titulación se realizaron inyectando sucesivamente volúmenes variables (perfil de volúmenes variable) de la disolución de ligando (de 2 μL a 13 μL) para poder definir mejor la curva de titulación. Para generar este perfil se emplearon unas plantillas de simulación desarrolladas por uno de los miembros de nuestro grupo de investigación. Un parámetro muy útil es el producto de la constante de afinidad, K_d , por la concentración total de

macromolécula, $[M]_T$, cuyo valor debe estar entre 1 y 100, aunque el valor óptimo está entre 10 y 100, el cual es determinante para optimizar la forma de la curva isoterma de unión, que debe ser lo más sigmoideal posible (Wiseman et al., 1989). Dicha curva se obtiene representando el calor por mol de ligando añadido tras cada inyección frente al cociente de la concentración de ligando total dividido entre la concentración de macromolécula total (considerando en este caso que el ligando está en la jeringa de inyección y la macromolécula en la célula de reacción). Otros parámetros utilizados para la optimización son el volumen de inyección y la concentración de ligando, los cuales, junto con la optimización del producto anterior, deben conseguir que se alcance la saturación elevada (es recomendable por encima del 80%), que indicaría que casi la totalidad de proteína que contiene la célula calorimétrica se encuentra unida al ligando (Jelesarov & Bosshard, 1999; Velazquez-Campoy & Freire, 2006; Velazquez-Campoy et al., 2004).

De estos experimentos se obtiene el termograma correspondiente a la interacción. Del mismo modo, es necesario obtener el termograma de dilución para lo que se titula el ligando sobre la disolución tampón en ausencia de proteína utilizando el mismo perfil de volúmenes. En ambos casos, el calor desarrollado tras cada inyección de péptido se mide a partir de la integral de la señal calorimétrica. Una vez corregida la línea base de ambos experimentos, la isoterma de unión neta para la interacción proteína-ligando se obtiene restando para cada inyección los calores de dilución a los calores del experimento principal. La constante de afinidad, K_d , y el cambio de entalpía aparente de unión, ΔH_d , se determinan analizando las isotermas de unión resultantes mediante ajuste no lineal de mínimos cuadrados al modelo de un conjunto de n sitios idénticos e independientes cuya ecuación de ajuste es:

$$\frac{1}{V_c} \frac{\Delta Q}{\Delta [L]_T} = \frac{\Delta H_d}{2} \left[1 - \frac{1 + [M]_T \theta - nK_d [M]_T}{\sqrt{1 + K_d [M]_T \theta + nK_d [M]_T - 4nK_d^2 [M]_T^2 \theta}} \right] \quad (3.4.65)$$

donde V_c es el volumen de la célula calorimétrica, $\Delta Q/\Delta [L]_T$ variación de calor respecto a la variación en la concentración total de ligando y $\theta = [L]_T / [M]_T$. El desarrollo matemático de dicha ecuación ya se ha descrito en alguna Tesis Doctoral realizada en el Departamento (Martin-García, 2009) y está implementado en el módulo para ITC del programa ORIGIN 7 (*Microcal Software Inc.*), con el cual se realizaron los análisis de los datos. El péptido sintético de secuencia KKETAV fue adquirido de la empresa Peptide 2.0.

3.12 Cristalización de Proteínas

La cristalización de moléculas en disolución es un fenómeno de equilibrio químico cuya cinética y termodinámica están controladas por diferentes características físicas (temperatura, gravedad, superficie, viscosidad, constante dieléctrica o vibraciones) y químicas (pH, fuerza iónica, iones metálicos o detergentes) del disolvente y del soluto. Bajo ciertas condiciones, el sistema origina una fase sólida a partir de la disolución que disminuye la energía libre del sistema (núcleo crítico o un agregado no específico) (Weber & Anfinsen, 1991). Esta disminución de energía libre es mayor cuanto más ordenada sea

esta fase sólida segregada. Estos conocimientos se han aplicado a la cristalización de proteínas, lo que ha permitido conseguir grandes avances en el campo.

Generalmente, se ha establecido que las disoluciones de proteína para cristalización deben cumplir los siguientes requisitos: a) una concentración elevada, b) una pureza tan alta como sea posible, al menos del 95%, c) homogeneidad y estabilidad en las condiciones de cristalización. Éste último requisito es de gran importancia, ya que la heterogeneidad puede aparecer como resultado de la manipulación experimental (grados distintos de oxidación, proteólisis, saturación parcial con ligandos, agregación en grado variable, desnaturalización parcial, deamidación, etc.).

La estrategia general para la cristalización de proteínas se basa en reducir la solubilidad de la macromolécula hasta que ésta alcanza el estado de sobresaturación. Se han desarrollado diferentes técnicas para llevar una disolución subsaturada de proteína a la región sobresaturada donde la nucleación y el crecimiento cristalino son posibles. Las técnicas más empleadas son la de difusión en vapor y la contradifusión en capilares (García-Ruiz & Carter, 2003).

En la técnica de difusión en vapor, la disolución de proteína se mezcla con la disolución de agente precipitante (en una gota sentada o colgante sobre una superficie) para que la mezcla se sitúe en la región sobresaturada que permita el crecimiento del cristal. En un reservorio cerrado, habrá un equilibrio entre la gota y la disolución precipitante, en el que, al haber una menor concentración de agente precipitante en la gota, ésta va perdiendo volumen para que se igualen las concentraciones de la gota y de la disolución del reservorio. En esta Memoria se ha utilizado el método de gota colgante, en el que la gota queda adsorbida en la cubierta que tapa herméticamente el reservorio (Figura 3.5).

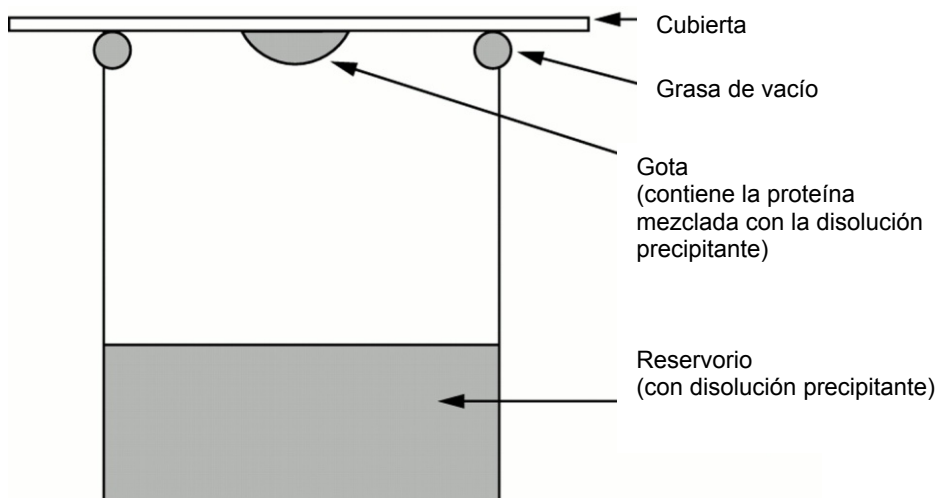


Figura 3.5. Dibujo esquemático del método de gota colgante.

En cuanto a la técnica de contradifusión en capilares, se utiliza un capilar en el que se pone una disolución de proteína y una cámara con el precipitante, en un tampón con unas condiciones de pH y fuerza iónica determinadas. La difusión del agente precipitante (mayor coeficiente de difusión de la proteína) en la cámara de la proteína originará un gradiente continuo de sobresaturación hasta alcanzar condiciones de equilibrio dependientes del espacio y del tiempo a lo largo de todo el capilar. El gradiente de sobresaturación generado mostrará distintos resultados a lo largo del capilar, que consistirán en la disolución clara sin ningún precipitado hasta la posible formación de cristales (Figura 3.6).

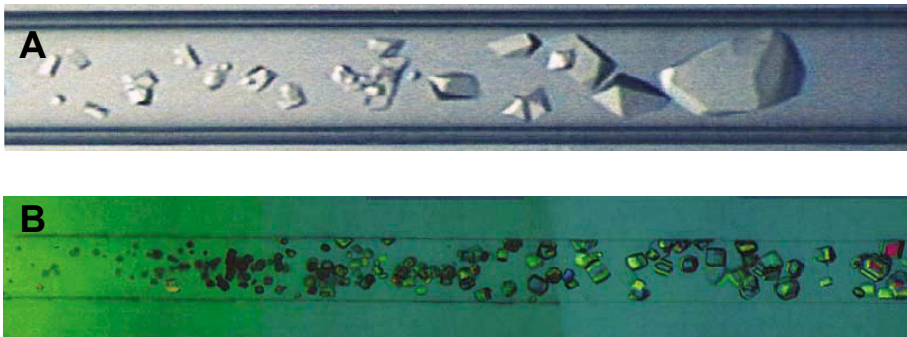


Figura 3.6. Dos ejemplos de cristales crecidos por el método de contradifusión por capilar, en el que el gradiente de sobresaturación va desde izquierda a derecha. Los cristales son de las proteínas (A) dehidroquinasa y (B) lisozima (Figura tomada de (García-Ruiz et al., 2003)).

3.13 La metodología de expresión en fagos (*Phage Display*)

La metodología de expresión en fagos (*phage display*, en inglés) se basa en el hecho de que la expresión de péptidos, proteínas o fragmentos de anticuerpos puede tener lugar en la superficie de un virus. Existen más métodos de co-expresión en los que se ha co-expresado una molécula de interés en ribosomas, en levaduras, en células de insectos... (Kehoe & Kay, 2005).

Para la expresión en fagos, se introduce una secuencia de DNA en el genoma de un virus, que hace que una secuencia aminoacídica se exprese fusionada a la de las proteínas de la cápside del virus. Ya que las proteínas de la cápside se exponen en la superficie del fago, se consigue una interrelación entre fenotipo y genotipo muy ventajosa, pues esta correlación permite fácilmente seleccionar dentro de una variedad los fagos más afines a determinadas moléculas diana, conociéndose su secuencia aminoacídica sencillamente con la secuenciación de su genoma.

Esta correlación genotipo-fenotipo junto con que hoy día se pueden hacer librerías de péptidos o proteínas de una forma muy eficiente y barata con una muy elevada variabilidad (más de 10^9 variantes) explican la potencialidad de esta tecnología. Esta es la razón de que su uso se haya extendido tanto, desbancando en varios casos a otras

tecnologías que han demostrado dar buenos resultados, como el cribado de péptidos unidos a membranas o el uso de levaduras alteradas genéticamente de tal forma que permiten detectar múltiples interacciones (el sistema *Yeast Two Hybrid*).

3.13.1 El bacteriófago M13

La partícula vírica que se usa con más frecuencia es el bacteriófago M13 (Clackson & Lowman, 2004), que ha sido el usado en esta Memoria. Este bacteriófago tiene una forma alargada (Figura 3.7), y consiste básicamente en un DNA circular de cadena monocatenaria que está encapsulado con un armazón de componentes proteicos. La proteína que se encuentra de forma mayoritaria es la proteína pVIII, de la que contiene unas 2700 copias. Además, cada extremo del fago está cubierto con dos tipos de proteínas: en uno de los extremos hay cinco copias de las proteínas pVII y pIX y en el otro extremo hay otras cinco copias de las proteínas pIII y pVI. Por otro lado, el virus puede presentar distintos tamaños, su longitud depende de la longitud de su genoma; genomas más largos generan fagos más largos, mientras que si son más pequeños, el tamaño será menor. Debido a estas características estructurales, se pueden añadir hasta 12000 bases al DNA del fago natural sin alterar su empaquetamiento.

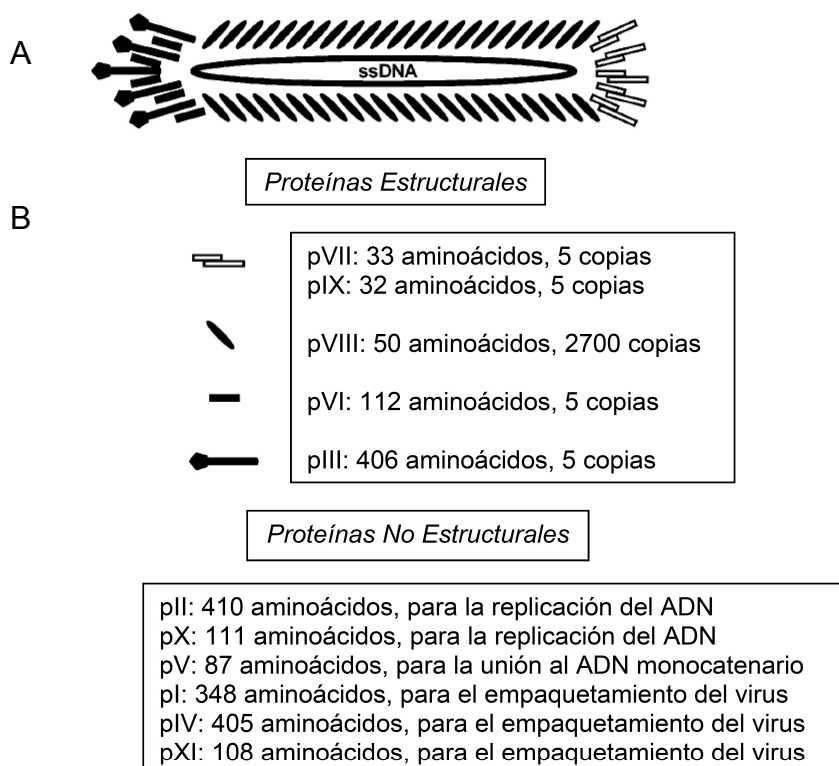


Figura 3.7. Panel A) Representación esquemática del fago M13. Panel B) Tabla-resumen con las proteínas que codifica su DNA monocatenario y la función que tienen en su ciclo de vida.

Por otro lado, el fago tiene un ciclo de vida (Figura 3.8) en el que primeramente se une a la bacteria a la que infecta (*E. coli*) a través del pilus bacteriano. Después de la unión, el pilus se repliega, llevando al fago a la superficie de la célula (paso 1 de la Figura 3.8). De esta forma se permite la interacción de la proteína pIII del fago con la proteína TolA de la membrana bacteriana. Gracias a esta interacción, el DNA monocatenario del fago puede introducirse en la célula bacteriana. Una vez allí, es convertido a bicatenario por la maquinaria de la bacteria (paso 2 de la Figura 3.8).

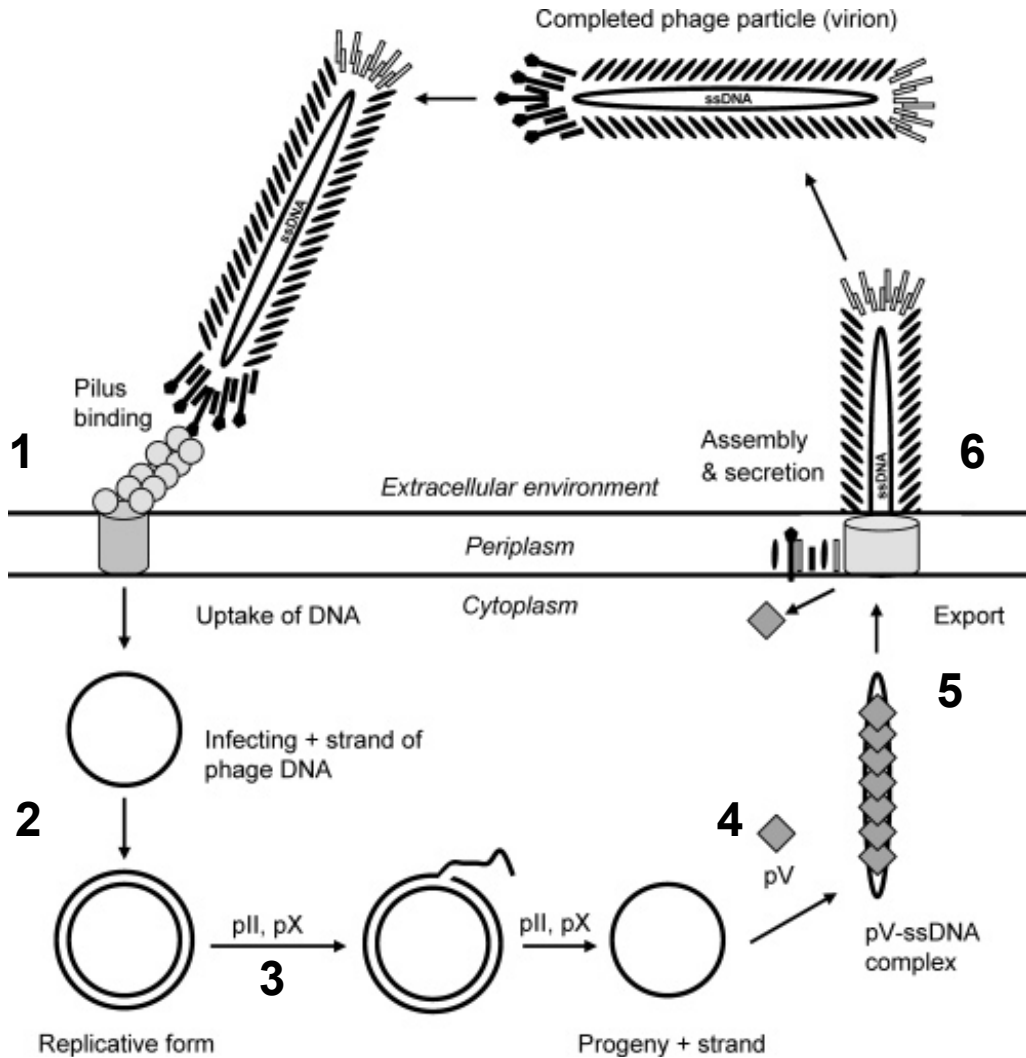


Figura 3.8. Ciclo de vida del fago M13. Una vez introducido en la bacteria mediante la unión al pilus F (paso 1), se forma el DNA bicatenario, que rápidamente va a ser replicado por la maquinaria de la bacteria (paso 2). Una vez que se han producido suficientes copias (paso 3), el DNA monocatenario se une a la proteína pV (paso 4), cuya función es doble: impedir la formación de DNA bicatenario y favorecer la interacción con la maquinaria de secreción (paso 5). Finalmente el fago es expulsado al exterior de la bacteria sin su muerte ni lisis (paso 6). Figura tomada de (Kehoe & Kay, 2005)

Este DNA viral en su forma bicatenaria provoca una cascada de procesos que llevan a la proliferación del virus (Clackson & Lowman, 2004). El primer paso es la replicación del DNA. Para ello se requiere la acción de dos proteínas no estructurales del fago, la pII y la pX, que a su vez permiten también la formación del DNA viral en su forma monocatenaria (paso 3 de la Figura 3.8). Para que ésta no pase a bicatenaria, conforme se va generando la cadena de DNA, se va uniendo a la proteína pV (paso 4 de la Figura 3.8). En un paso posterior, con el fin de ir ensamblando la partícula vírica, este conglomerado DNA-proteína pV interacciona con la maquinaria de exportación, que es un complejo supramacromolecular que está formado por las proteínas víricas pI, pIV, pIX y una tiorredoxina de la bacteria. Esta maquinaria va simultáneamente empaquetando y secretando los nuevos fagos, usando las proteínas de recubrimiento que ya han sido previamente insertadas en la membrana interna de la bacteria (paso 5 de la Figura 3.8). Finalmente, los fagos son liberados de la célula, sin destruirla ni provocar su lisis (paso 6 de la Figura 3.8).

3.13.2 Co-expresión de proteínas y péptidos fusionados a una proteína de la cápside

Las cinco proteínas de la cápside del fago (pIII, pVI, pVII, pVIII y pIX) se han usado para fusionar una secuencia de interés (Kehoe & Kay, 2005); sin embargo, las dos que se han empleado con más frecuencia son las proteínas pIII y pVIII. Ambas se utilizan para fusionar péptidos y únicamente la pIII se utiliza para co-expresar proteínas. Esto es así debido a los requerimientos estéricos derivados de la estructura del fago, ya que esta proteína pIII únicamente presenta unas cinco copias por fago. Sin embargo, como ya se ha comentado, hay unas 2700 copias de la pVIII, lo que provoca que haya mucha mayor dificultad en el correcto empaquetamiento del virus si esta proteína tiene fusionada otra proteína de un cierto tamaño.

Por otro lado, cuando se co-expresan péptidos fusionados a la proteína pVIII, hay que tener en cuenta el efecto de multiplicidad. Al co-expresarse tantas copias del mismo péptido, puede ocurrir que el fago esté interaccionando con varias moléculas diana a la vez (Figura 3.9), lo que provocaría su selección debido a un efecto tipo quelato. De esta forma, su afinidad estaría sobrevalorada por esa unión múltiple. Por esto, el uso de librerías peptídicas en pVIII es muy adecuado para la identificación de ligandos de baja afinidad para moléculas diana problemáticas. Sin embargo, a la vez, hace que sea difícil la discriminación entre ligandos de alta y baja afinidad. Finalmente, este fenómeno no ocurre cuando un péptido se co-expresa con la proteína pIII, ya que al haber muchas menos copias, este efecto de multiplicidad es mucho menor.

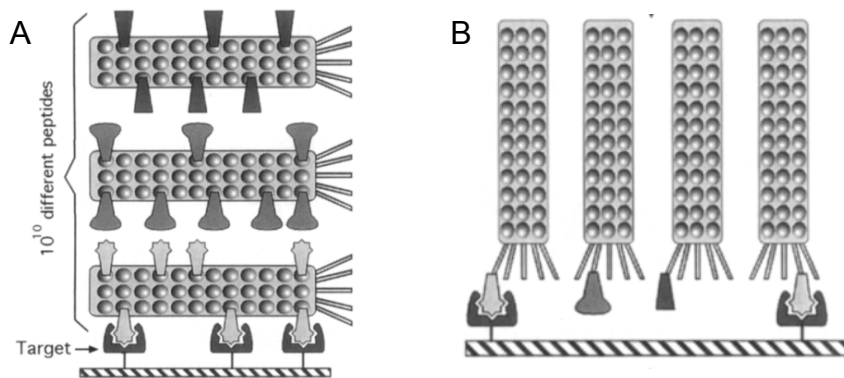


Figura 3.9. Efecto de multiplicidad. Panel A) Co-expresión en *pVIII*, que se representa con círculos dentro del cuerpo del fago, habiendo muchas copias dentro de cada fago. Panel B) Co-expresión en *pIII*, que está situada en el extremo del fago como líneas, siendo cinco el número total de copias en cada fago. Figura tomada de (Sidhu et al., 2000).

3.13.3 Procedimiento experimental

Hay varios tipos de vectores y distintas metodologías posibles para la co-expresión de secuencias aminoacídicas en cualquiera de las proteínas del fago (Sidhu, 2005). Para la construcción de librerías de fagos, también existen varias metodologías. A continuación se describe la que se ha usado en la elaboración de esta Tesis Doctoral, mediante el método de mutagénesis de Kunkel y el uso de fagémidos. También se describe la forma de seleccionar las variantes óptimas de librerías de fagos que unen con afinidad a las moléculas diana frente a las cuales están siendo seleccionadas, así como la detección de las variantes seleccionadas a partir de librerías de fagos que unen con mayor afinidad a su molécula diana mediante un ensayo inmunoabsorbente ligado a enzimas (ELISA, en inglés). Finalmente, se describe el proceso de secuenciación de dichas variantes.

3.13.3.1 Mutagénesis de Kunkel

Para la co-expresión de secuencias aminoacídicas en cualquiera de las proteínas de la cápside del fago, en esta Tesis Doctoral se han usado fagémidos (*phagemid*, en inglés), es decir, plásmidos que contienen los componentes requeridos para empaquetar el DNA viral dentro de las partículas víricas (Clackson & Lowman, 2004). Junto con el fagémido, se ha de usar un *fago asistente* (*helper phage*, en inglés) para aportar la maquinaria vírica que permita la replicación del virus. Este fago asistente es un fago que posee determinadas mutaciones que reducen la eficacia de empaquetamiento con sus proteínas, permitiendo así empaquetar preferencialmente las proteínas que se codifican del fagémido.

Más concretamente, para co-expresar péptidos o proteínas usando un sistema de fagémidos, las bacterias que contienen el fagémido se infectan con el fago asistente. Una vez producida la infección, el genoma del fago asistente inicia la síntesis de todas las

proteínas naturales del fago. Entonces, la síntesis de la proteína de interés fusionada a la proteína de la cápside del fago se inicia también, ya que el DNA del fagémido contiene un promotor previo a la secuencia de nucleótidos que codifica ambas proteínas fusionadas, de tal forma que se van incluyendo entre la mezcla de las proteínas víricas que ya se encuentra en la célula. Cuando comienza el empaquetamiento del fago, los virus se empaquetan usando preferentemente el DNA del fagémido y su proteína de fusión, que se diferencian con las correspondientes del fago asistente en que contienen señales de empaquetamiento plenamente funcionales.

Para construir una librería de fagos que co-expresen múltiples variantes de un péptido o proteína a partir de un fagémido, se usa generalmente el método de mutagénesis desarrollado por Kunkel (Kunkel, 1985), consistente básicamente en una mutagénesis dirigida utilizando un oligonucleótido que se une a una cadena de DNA monocatenario para ir añadiendo bases complementarias a dicha cadena con el uso de la enzima polimerasa T7. El DNA final de doble cadena que contiene las mutaciones se circulariza mediante la ligasa T4, que une los extremos romos que quedan después de la acción de la polimerasa. Este paso de mutagénesis por el que se genera la librería aleatoria es el paso más crítico en la metodología de la expresión en fagos, por lo que se han probado otros métodos alternativos, que se pueden encontrar en la literatura (Sidhu, 2005), aunque se ha demostrado que el de Kunkel es el más efectivo, de ahí que sea el más usado.

En el caso de la metodología de Kunkel, para la introducción de mutaciones aleatorias se usan oligonucleótidos en los que se introducen distintos codones que codifican distintos aminoácidos naturales, que se denominan codones degenerados. Así, por ejemplo, si se quiere disponer de los veinte aminoácidos naturales en una determinada posición del péptido o proteína a co-expresar, el codón usado sería el NNK (Tabla 3.2), en donde N es una mezcla equimolar de las cuatro bases del DNA (A/G/C/T) y K es una mezcla al 50% de timina y guanina. En la Tabla 3.2 se pueden ver los codones degenerados posibles que se pueden usar para introducir un determinado tipo de aminoácido u otro (polares, hidrofóbicos, etc).

El número de posiciones mutadas variará en función del estudio que se esté realizando; sin embargo, hay que tener en cuenta que el número teórico de variantes de la librería aumentará conforme aumenta el número de mutaciones. Por ejemplo, si se mutan 5 posiciones con los 20 posibles aminoácidos naturales (codón degenerado NNK), y 3 posiciones con 8 aminoácidos hidrofílicos (codón degenerado DVT) la variabilidad teórica sería $20^5 \cdot 8^3 = 1.6 \cdot 10^9$ variantes.

El número total de mutaciones, por tanto, viene restringido por la diversidad real a la que se puede llegar experimentalmente, debido a las limitaciones de la transformación en la bacteria. Los rendimientos óptimos que se pueden obtener en el proceso son de unas 10^{10} variantes, por lo que la aleatoriedad teórica que se desee introducir no debería superar este número.

Tabla 3.2. Codones degenerados usados para la formación de librerías combinatorias de péptidos y proteínas. (Tabla tomada de (Tonikian et al., 2007)).

Triplete ^a	Descripción	Aminoácidos	Triplete de parada	Tripletes únicos ^b
NNK	Los 20 aminoácidos naturales	Los 20	TAG	32
NNC	15 aminoácidos	A, C, D, F, G, H, I, L, N, P, R, S, T, V, Y	-	16
NWW	Cargados, hidrofóbicos	D, E, F, H, I, K, L, N, Q, V, Y	TAA	16
RVK	Cargados, hidrofílicos	A, D, E, G, H, K, N, R, S, T	-	12
DVT	Hidrofílicos	A, C, D, G, N, S, T, Y	-	9
NVT	Cargados, hidrofílicos	C, D, G, H, N, P, R, S, T, Y	-	12
NNT	Mezcla	A, D, G, H, I, L, N, P, R, S, T, V	-	16
VVC	Hidrofílicos	A, D, G, H, N, P, R, S, T	-	9
NTT	Hidrofóbicos	F, I, L, V	-	4
RST	Cadenas laterales pequeñas	A, G, S, T	-	4
TDK	Hidrofóbicos	C, F, L, W, Y	TAG	6

^aLos símbolos de las bases degeneradas están representados según el código IUB (D: A/G/T; K: G/T; N: A/C/G/T; R: A/G; S: G/C; V: A/C/G; W: A/T).

^bNúmero de tripletes únicos que están contenidos en el triplete degenerado. Debido a la naturaleza del código genético, varios tripletes degenerados son redundantes. Esto significa que el número de tripletes únicos excede el número de aminoácidos únicos que son traducidos.

Para comenzar la mutagénesis, esquematizada en la Figura 3.10, se purifica el fagémido en su forma monocatenaria a partir de la cepa de *E. coli* CJ236 (Tonikian et al., 2007). Esta cepa se caracteriza por incorporar uracilo en vez de timina en el DNA. Para llevar a cabo la obtención del fagémido monocatenario, después de la transformación en ese tipo de células, se crecen cultivos de éstas que posean a la vez fago asistente. De esta forma, se generan fagos cuyo DNA monocatenario es el fagémido, pudiéndose aislar éste en su forma monocatenaria con una batería comercial de QIAGEN (*QIAprep Spin M13*).

Tal y como se observa en la Figura 3.10, este DNA monocatenario que tiene uracilo en vez de timina (dU-ssDNA en la Figura) es el que se usa como molde al que se une el oligonucleótido mutagénico. Para asegurar que la unión es adecuada, es crítico que la complementariedad de dicho oligonucleótido con el molde de DNA monocatenario sea de, al menos, 15 nucleótidos por encima y por debajo de la región mutada. Aparte de este requerimiento, la otra limitación viene dada por la longitud del oligonucleótido, ya que a partir de unas 100 bases su síntesis y purificación es mucho más complicada y está sujeta a posibles errores de secuencia.

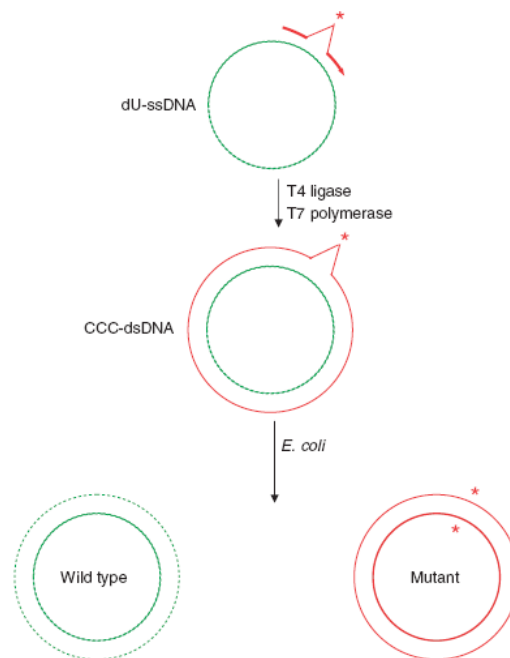


Figura 3.10. Esquema de los pasos en la mutagénesis de Kunkel. Al DNA monocatenario que tiene uracilo en vez de timina (dU-ssDNA) se une un oligonucleótido mutagénico, en rojo con la mutación señalada con un asterisco. La acción de la ligasa T4 y la polimerasa T7 permiten la formación de un DNA de doble cadena circularizado (CCC-dsDNA), el cual, al introducirlo en bacterias *E. coli*, se replica dando el DNA original (Wild type) y el mutado (Mutant). El original es destruido por la bacteria al presentar uracilo en lugar de timina (Figura tomada de (Tonikian et al., 2007)).

Cuando el oligonucleótido, representado en rojo en la parte superior de la Figura 3.10 con la mutación indicada con un asterisco, está unido con el molde, la polimerasa T7 de DNA comienza la síntesis del DNA complementario al molde, obteniéndose así una doble cadena de DNA. A continuación se usa una ligasa T4 para unir los extremos romos del DNA que quedan como consecuencia de la acción de la polimerasa. Este nuevo DNA bicatenario (CCC-dsDNA en la Figura 3.10) contiene una zona en la que no hay un buen emparejamiento de bases, que es la correspondiente a la mutación.

Para completar la construcción de la librería, ese DNA de cadena doble y cerrado se introduce en una cepa de células de *E. coli* como las XL1-Blue, la cual, al iniciar el proceso de replicación del DNA, destruye la cadena que posee uracilo al no reconocerla como propia, y se replica la que se había generado con la polimerasa, que es además la que contiene la secuencia del oligonucleótido mutagénico. Este procedimiento de mutagénesis tiene una eficacia de más del 80% (Tonikian et al., 2007).

La introducción del DNA bicatenario ha de hacerse en una cepa de *E. coli* que tenga el pilus bacteriano al que se une el bacteriófago M13 que permite su infección dentro de la bacteria. Una limitación importante para la producción de librerías con alta variabilidad ha sido la transformación eficiente de grandes cantidades de DNA en células de este tipo. Muchas veces se usa la electroporación, que se basa en que las bacterias

toman el DNA de la disolución que les rodea cuando se ven sometidas a un campo eléctrico intenso.

Para optimizar ese proceso, en el laboratorio del Dr. Sidhu, se diseñó una nueva cepa de *E. coli* para este tipo de transformaciones de alta eficacia, llamada SS320 (Sidhu et al., 2000). Esta cepa se generó hibridando los genomas de las cepas MC 1061 y XL1-Blue y seleccionando en medios con tetraciclina y estreptomina. Así, la nueva cepa mantiene la alta eficacia en la transformación que posee la cepa MC 1061 y contienen también el gen que codifica el pilus bacteriano de las XL1-Blue, que es requerido para la infección y propagación de los fagos.

3.13.3.2 Proceso de selección

Con la librería construida, se puede comenzar el proceso de selección de los fagos más afines frente a la molécula diana (ver Figura 3.11). El protocolo que se describe a continuación es una estrategia basada en microplacas, lo que permite escalarla a un alto número de moléculas diana de forma simultánea (Tonikian et al., 2007).

El primer paso es inmovilizar la molécula diana. Para ello hay bastantes estrategias, muchas de ellas basadas en que previamente se le ha unido a la molécula diana otra molécula que tiene una elevada especificidad por un determinado compuesto. Así, esa unión tan fuerte permitirá la inmovilización, presentando además una única orientación de la molécula. Otras veces, y es el caso de esta Memoria, la inmovilización se produce sencillamente por la adsorción a un material con una alta capacidad adsorbente (*Maxisorp*[®], de la casa comercial *NUNC*). Hay que bloquear las posiciones activas a las que no se ha adsorbido la molécula diana, para que no se unan fagos a la superficie de forma inespecífica. En esta Memoria el bloqueo se ha hecho con la albúmina de suero bovino, que es la que además generalmente se usa en este tipo de procesos.

Una vez inmovilizada la molécula, se añade la librería de fagos en condiciones apropiadas para la unión (paso señalado como "*Library Selection*" en la Figura 3.11). Después de un período de incubación, habrá fagos unidos específicamente y otros cuya unión haya sido inespecífica, y, por tanto, más débil. Para eliminar estos últimos se realizan una serie de lavados con PBS al que se le añade un 0,05% del detergente Tween 20 (paso señalado como "*Wash*" en la Figura 3.11).

Para continuar con la selección hay que desestabilizar el enlace entre los fagos que han quedado después de los lavados y la molécula diana (paso señalado como "*Elution*" en la Figura 3.11). De nuevo, para ello hay varias alternativas, como por ejemplo usar ligandos conocidos que sean altamente afines a la molécula diana y hagan competición con el fago. Este método puede resultar útil cuando haya algún ligando conocido; sin embargo, el método más generalizado es añadir un medio ácido, por ejemplo, ácido clorhídrico a una concentración 0,1 M, que rompe la unión y permite eluir los fagos. Cuando se usa el medio ácido, hay que llevar la disolución de nuevo a un pH neutro añadiendo una base, por ejemplo Tris a pH 11, para que los fagos estén en las condiciones óptimas para el siguiente paso.

Así, para continuar el proceso de selección se han de amplificar los fagos eluidos en la primera etapa (paso señalado como “*Amplify in Escherichia coli and Reselect*” en la Figura 3.11). Esto se hace mediante la infección con estos de un cultivo de la cepa XL1-Blue en su punto óptimo de crecimiento celular (punto en el cual la absorbancia del cultivo a 600 nm está en torno a 0.6) y se deja incubar durante una media hora. A continuación se añade el fago asistente en el proceso conocido como “super-infección”, para facilitar la infección por parte de los fagos eluidos. Finalmente, el cultivo se añade a otro medio de cultivo con un volumen mayor, para dejarlo un período largo de tiempo propagándose y así poder aislar los fagos al día siguiente. Con este nuevo conjunto de fagos, que posee muchas menos variantes que la librería y que en principio unen ya con cierta afinidad, se continúa el proceso de selección como en la primera etapa. Este proceso se puede repetir cuantas veces se consideren necesarias para obtener un enriquecimiento en aquellos fagos que unen mejor a la molécula diana.

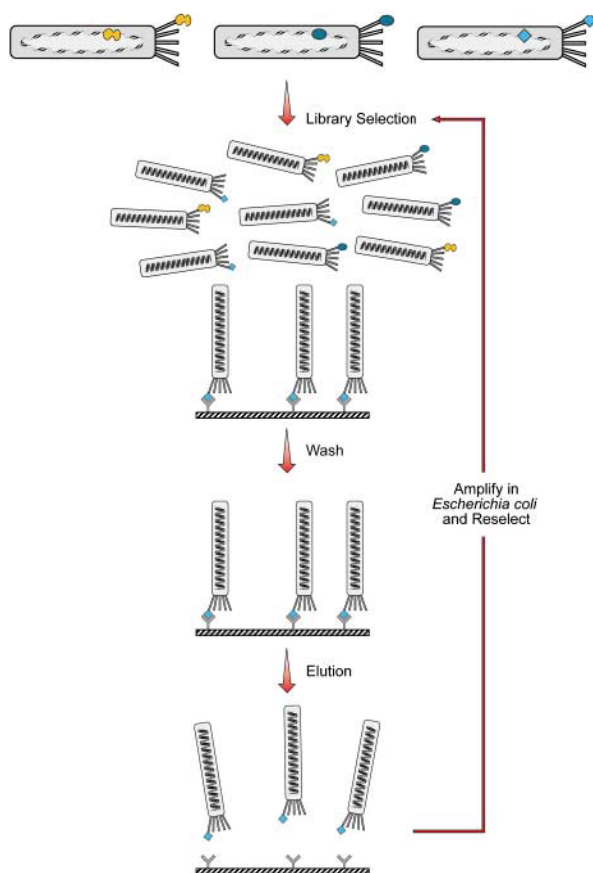


Figura 3.11. Esquema de los pasos en el proceso de selección. El conjunto de fagos iniciales, que tienen distintos polipéptidos expresados, se incuban con la molécula diana, fijada en una superficie (Library selection). Después de la unión de los fagos que unen mejor a dicha molécula, se hacen una serie de lavados para eliminar aquellos fagos que no han unido (Wash). A continuación hay que romper el enlace entre el fago y la molécula diana (Elution), para poder amplificarlos en bacterias *E. coli* (Amplify in *Escherichia coli* and Reselect) y continuar el proceso de forma iterativa hasta conseguir enriquecer el conjunto de los fagos con aquellos que unen mejor a la molécula diana (Figura tomada de (Sidhu et al., 2003)).

De hecho, para ver si realmente se produce una ratio de enriquecimiento de los fagos más afines a las moléculas diana durante el proceso de selección, se puede realizar una titulación de los fagos sembrando en placas los cultivos infectados con estos y contando el número de colonias o bien llevar a cabo un ensayo ELISA, que se describe en la siguiente Sección. La ratio de enriquecimiento indica cómo varía el número de fagos unidos a la molécula de una etapa a otra. Este número, que inicialmente se encuentra alrededor de 10^9 , correspondiente a la diversidad de la librería, ha de bajar en la primera etapa debido a que existen muchas variantes de la librería que no unen de forma específica, quedando en unos 10^5 fagos. A continuación ha de ir aumentando en las siguientes etapas, puesto que se van amplificando únicamente aquellos que unen con más afinidad. De esta forma, el número de etapas que se llevan a cabo depende de cómo vaya el enriquecimiento. Generalmente se hacen un mínimo de tres y a veces se puede llegar hasta seis etapas. Tampoco es recomendable hacer muchas etapas más puesto que pueden aparecer artefactos debidos a la amplificación de los fagos, como por ejemplo que proliferen variantes que se expresen mejor pero que no sean especialmente afines a la molécula diana (Sidhu, 2005).

En un siguiente paso, es necesario analizar dentro del conjunto de fagos cuáles de ellos unen a la molécula diana, puesto que han podido ocurrir uniones inespecíficas. Para ello hay que sembrar un cultivo de células XL1-Blue infectadas por los fagos seleccionados y obtener colonias en una placa petri. Cada una de estas colonias representa una *estirpe de fagos monoparental*, puesto que cada una de esas colonias contendría un fagémido con una única secuencia de la librería. Para ver si una estirpe de fagos une con afinidad a la molécula diana el análisis que generalmente se hace, y es el que se ha llevado a cabo en esta Memoria, es el ensayo ELISA (Sidhu et al., 2000).

3.13.3.3 Ensayo inmunosorbente ligado a enzimas (ELISA)

El ensayo inmunosorbente ligado a enzimas ELISA (Enzyme-Linked ImmunoSorbent Assay) se basa en la adición de los fagos en estudio tanto a la molécula diana como a una molécula control, que ejerce de control negativo, lo cual permite saber si la unión del fago es por la interacción específica entre el polipéptido fusionado a la proteína de la cápside del fago y la molécula diana o bien si la unión se debe a otras causas, como por ejemplo la unión a la albúmina, que, como se ha comentado, se usa muy frecuentemente para bloquear los sitios de unión que no son ocupados por la molécula diana.

Este ensayo consta de una serie de etapas que están esquematizadas en la Figura 3.11. La primera etapa, como en las selecciones de los fagos, consiste en inmovilizar tanto la molécula diana como la proteína control (Panel A de la Figura 3.12). Las estrategias de inmovilización son las mismas que las descritas para las selecciones. De nuevo en esta Memoria se eligió el uso de una superficie con una alta capacidad de adsorción, por lo que, tal y como se acaba de mencionar, después de la inmovilización se bloquean las posiciones activas que no han sido ocupadas por las moléculas diana mediante la

albúmina. Posteriormente, se añaden los fagos que se han tenido creciendo en el medio de cultivo 2YT durante al menos 16 horas (Panel B de la Figura 3.12). Se deja un período de incubación de 2 horas y se hace una serie de lavados con PBS con Tween 20 al 0.05%, para eliminar aquellos fagos que no se han unido.

A continuación se añade el anticuerpo antiM13 (Panel C de la Figura 3.12). Este anticuerpo, comercial (GE Healthcare), es específico a la proteína pVIII del fago M13. Dicho anticuerpo se deja incubando durante un intervalo de tiempo, y se vuelven a hacer lavados con PBS y Tween 20 al 0.05%, para eliminar el anticuerpo no unido. A continuación se añade la 3,3,5,5'-tetrametilbenzidina (TMB) (Pierce), que es un producto que se oxida por una enzima que está conjugada al anticuerpo antiM13, la peroxidasa de rábano común (Panel D de la Figura 3.11). Al oxidarse el TMB por la acción de dicha enzima, da un producto azul que absorbe entre 370 nm y 652 nm. La reacción enzimática se deja un tiempo variable, entre dos minutos y media hora, para que dé tiempo a que se desarrolle por completo y después se detiene añadiendo un medio fuertemente ácido, normalmente ácido fosfórico a una concentración de 1 M, lo que provoca un nuevo cambio de color, en este caso, a amarillo. Con una medida espectrofotométrica a la longitud de onda de 450 nm se obtiene un valor de absorbancia, que da una estimación de la cantidad de anticuerpo presente, que, a la vez, es la misma cantidad de fagos presentes. Al comparar los valores de absorbancia obtenidos con la molécula diana y con la molécula control, se podrá estimar de esta forma la diferencia en el número de fagos unidos a ambas moléculas, con lo cual se podrá saber si existe unión entre la molécula diana y el polipéptido fusionado a la proteína de la cápside del fago o bien si se da una unión inespecífica que se da por otras causas, o simplemente si no hay unión.

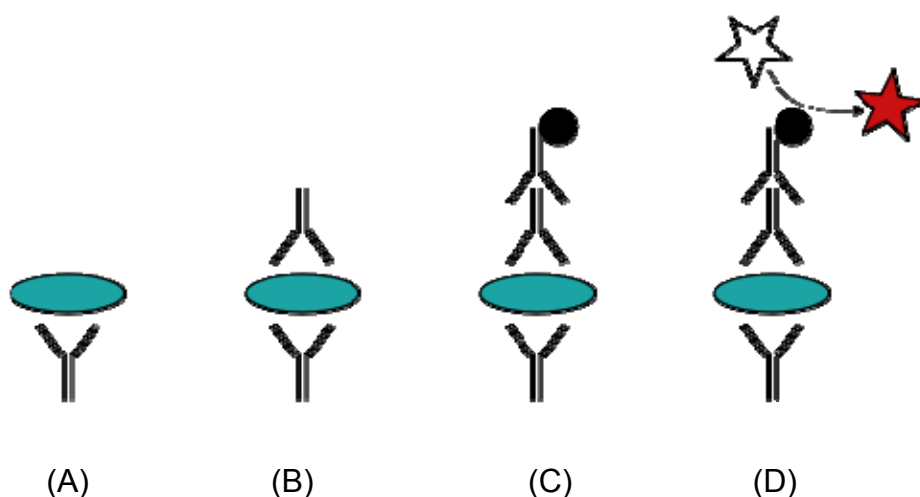


Figura 3.12. Esquema de las etapas del ensayo ELISA. En el paso A, se inmoviliza la molécula diana (y la control, paralelamente). En el B, se añaden los fagos, para, después de hacer lavados, añadir el anticuerpo del fago, que lleva conjugada una enzima (paso C). Finalmente, en el paso D está representada la adición del sustrato enzimático, que la enzima transforma en el producto final.

3.13.3.4 Secuenciación del DNA y su análisis

Después de realizar el análisis de estirpes de fagos monoparentales y seleccionar aquellos que unen con afinidad a la molécula diana mediante el ensayo ELISA, el último paso consiste en secuenciar los fagémidos que están encapsulados en dichos fagos. Estos se usan como molde para una PCR en la que se amplifica la región donde se encuentra la variabilidad. A continuación, esa PCR es secuenciada y traducida para conocer la correspondiente secuencia aminoacídica.

A veces estas secuencias aminoacídicas se usan para compilar un perfil consenso. Para generar dicho perfil, las secuencias deben alinearse entre sí. Para ello puede usarse sencillamente una hoja de cálculo; sin embargo, en el caso de que no se viese un consenso claro, se podría usar alguna de las herramientas de alineación de secuencias como *ClustalW* (Higgins et al., 1996), por ejemplo. Una vez que las secuencias han sido alineadas, se puede usar el programa Weblogo (Crooks et al., 2004) para compilar y visualizar los perfiles consenso, también llamados logos de secuencias. En estos, la altura de un aminoácido es indicativa de su frecuencia relativa en esa posición.

CAPÍTULO 4

ANÁLISIS ESTRUCTURAL POR DIFRACCIÓN
DE RAYOS X DEL TERCER DOMINIO PDZ DE
LA PROTEÍNA PSD95

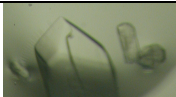
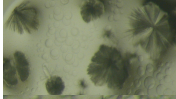
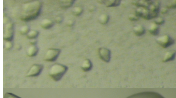
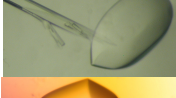
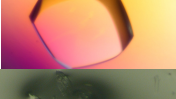
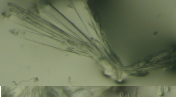
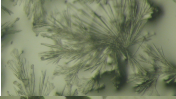
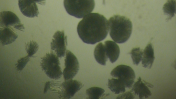

4.1 Resultados

Hasta la fecha, el tercer dominio PDZ de la proteína PSD95 ha sido uno de los dominios PDZ mejor caracterizados, y su estructura fue la primera que se resolvió tanto en presencia como en ausencia de ligandos peptídicos (Doyle et al., 1996). Después de este trabajo inicial, en la base de datos de estructuras de proteínas del *Protein Data Bank* (PDB), han ido apareciendo más estructuras resueltas de este dominio (estructuras con códigos PDB 1TP3, 1TP5 y 1TQ3). Todas estas estructuras tienen en común que contienen una horquilla β antiparalela en su extremo C-terminal (correspondiente a los residuos 403-415, según la numeración de (Doyle et al., 1996)), que no están presentes en la secuencia natural de la proteína PSD95. Por otro lado, en los estudios de plegamiento que se habían llevado a cabo con este dominio también se utilizó esa construcción que presenta la horquilla β en su extremo C-terminal que resulta como un producto de la clonación (Gianni et al., 2005; Feng et al., 2005). En este sentido, es bien conocido que el plegamiento de una proteína puede depender de todos y cada uno de los restos presentes, ya que la secuencia primaria de aminoácidos es la responsable en última instancia de la estructura y plegamiento de las proteínas. Por todas estas razones, y puesto que uno de los objetivos de esta Memoria es desarrollar los estudios del equilibrio conformacional de este dominio, decidimos clonar y purificar este dominio sin esa secuencia adicional, es decir, mantener los residuos 302-402 y eliminar los residuos 403-415 de la secuencia de (Doyle et al., 1996) que pertenecen al vector de expresión que se utilizó en los estudios mencionados y no a la proteína PSD95.

4.1.1 Cristalización del dominio PDZ3-PSD95

Puesto que todas las estructuras disponibles en el *Protein Data Bank* contienen la horquilla β adicional, se hizo más que necesario determinar la estructura cristalográfica de este dominio sin dicho artefacto. Para ello, se dializó el dominio PDZ3-PSD95 en tampón tris 10 mM pH 8 y se concentró hasta $30 \text{ mg}\cdot\text{mL}^{-1}$. A continuación, para poder encontrar condiciones de cristalización, se usaron las disoluciones de la batería comercial I de Hampton y se probaron todas las condiciones precipitantes de dicha batería. Éstas se dispusieron en una placa de 48 pocillos a los que se añadieron $200 \mu\text{L}$ de cada una de las disoluciones precipitantes y sobre los que se pusieron las cubiertas en que se añadió una gota de $2 \mu\text{L}$ de la disolución del dominio PDZ3-PSD95 mezclada con otros $2 \mu\text{L}$ de la disolución precipitante correspondiente al pocillo donde se coloca la cubierta con la gota. Para sellar el pocillo con la cubierta se usó grasa de vacío, para poder permitir el paso de agua desde la gota hacia la disolución, lo que ocurrirá hasta que la concentración de agente precipitante en la gota y en la disolución presente en el pocillo se igualen, puesto que en la gota la concentración inicial es la mitad que en el pocillo. Con este método, llamado de gota colgante y descrito en la Sección 3.12, se obtuvieron, después de varias semanas a 15°C , cristales de proteína en las condiciones 11, 20, 23, 35, 42, 43, 47 y la 48 de la batería comercial I de la casa Hampton (Tabla 4.1).

Tabla 4.1. Condiciones experimentales de cristalización obtenidas mediante la batería comercial I de Hampton para el dominio PDZ3-PSD95.

nº	pH	Composición de la disolución	Cristal
11	5.6	0.1 M Citrato sódico tribásico dihidratado, 1.0 M fosfato amónico monobásico	
20	4.6	0.2 M Sulfato amónico, 0.1 M acetate sódico trihidratado, 25% p/v Polietilenglicol 4,000	
23	7.5	0.2 M Cloruro de magnesio hexahidratado, 0.1 M HEPES sódico, 30% v/v Polietilenglicol 400	
35	7.5	0.1 M HEPES sódico, 0.8 M fosfato sódico monobásico monohidratado, 0.8 M fosfato potásico monobásico	
35*	7.5	0.1 M HEPES sódico, 0.8 M fosfato sódico monobásico monohidratado, 0.8 M fosfato potásico monobásico	
42	-	0.05 M fosfato potásico monobásico, 20% p/v Polietilenglicol 8,000	
43	-	30% p/v Polietilenglicol 1,500	
47	4.6	0.1 M Acetato sódico trihidratado, 2.0 M sulfato amónico	
48	8.5	0.1 M Tris, 2.0 M Fosfato amónico monobásico	

4.1.2 Resolución de la estructura del dominio PDZ3-PSD95 mediante difracción de rayos X

Las estructuras recogidas en esta Memoria se resolvieron a partir de los cristales obtenidos de las condiciones 20 y 35 de dicha batería comercial. Los que se obtuvieron de la condición 35, que contiene 0.1 M de HEPES sódico pH 7.5, 0.8 M de fosfato sódico monobásico mono hidratado y 0.8 M de fosfato potásico monobásico mono hidratado, se tomaron directamente del barrido de condiciones con la batería comercial I de Hampton y se añadieron a una disolución que contiene un 10% de glicerol, que sirve de crioprotector cuando se introduce el cristal en un flujo de N₂ líquido a 110 K. Las bajas temperaturas se usan para que los cristales puedan tolerar la intensidad de los rayos X. A continuación se midió con el difractómetro de rayos X que está en el Laboratorio de Estudios

Cristalográficos (LEC), en colaboración con el Dr. José Antonio Gavira, de dicho laboratorio. La resolución de la estructura y todo el procesado de los datos se llevó a cabo por la Dra. Ana Cámara, de la Universidad de Almería.

En cuanto a los cristales obtenidos mediante la condición número 20 de la batería I de Hampton, que contiene 0.2 M de sulfato de amonio, 0.1 M de acetato de amonio trihidratado pH 4.6 y un 25% p/v de PEG 4000, se crecieron mediante la técnica de contradifusión. En esta técnica al introducir un capilar que contiene la disolución de proteína en la disolución precipitante se genera mediante un proceso de difusión un gradiente de concentraciones entre las cuales está la adecuada para la cristalización de la macromolécula (García-Ruiz et al., 2003). Los cristales obtenidos de esta forma se midieron con la radiación de las Instalaciones Europeas de Radiación Sincrotrón (*European Synchrotron Radiation Facility*, ESRF). De nuevo, la resolución de la estructura la realizó la Dra. Cámara.

En la Tabla 4.2 están recogidos todos los datos provenientes de la difracción así como del refinamiento de la estructura. Las celdillas unidad de los cristales obtenidos mediante las dos condiciones son de distinta simetría. Para los obtenidos a partir de la condición 35 la simetría es tetragonal, perteneciente al grupo espacial $P4_3$, mientras que los obtenidos a partir de la condición 20 la simetría es triclínica, perteneciente al grupo espacial P1. Cabe destacar de esos datos la resolución máxima obtenida, de 1.4 Å para los que tienen simetría tetragonal y de 1.35 Å para los que tienen simetría triclínica, lo que refleja la excelente calidad de los datos. Cabe destacar que hasta ahora, todas las estructuras resueltas del dominio PDZ3-PSD95 pertenecen al mismo grupo espacial ($P4_132$) con parámetros de celdilla unidad parecidos entre sí y poseen grados más bajos de resolución, siendo 1.54 Å la máxima para la estructura de este dominio que está formando un complejo con un péptido, en el caso de la que tiene el código PDB 1TP5 y 1.89 Å en el caso en que está cristalizada sin péptido, con código PDB 1TQ3.

4.1.3 Análisis comparativo de las diferentes estructuras del dominio PDZ3-PSD95

Para comparar la cadena única de la estructura tetragonal con las cuatro cadenas presentes en la unidad asimétrica del cristal en forma triclínica, se usó el parámetro conocido con las siglas RMSD (*Root Mean Square Deviation*, la raíz de la media de la desviación cuadrática, la desviación típica), que engloba las diferencias individuales en las posiciones de cada resto. Estos valores se calcularon con el programa *Isqkab* de la macro de programas *CCP4*. Los valores de RMSD de la cadena principal de ambas estructuras son menores de 0.4 Å, pudiéndose afirmar por tanto que son prácticamente idénticos entre sí. Las diferencias más importantes se encuentran en los residuos 350-352 que pertenecen al giro que conecta la hélice $\alpha 1$ y la lámina $\beta 4$, que tienen un valor de RMSD máximo de 1.5 Å, y el residuo 384 del comienzo de la lámina $\beta 5$, que tiene un valor de RMSD máximo de 3 Å. Sin embargo, hay que tener en cuenta que estas diferencias están dentro del intervalo de aquellas encontradas cuando se comparan entre sí las cuatro cadenas

presentes en la estructura triclínica. Además, estas cuatro cadenas se disponen en un tetrámero que es casi idéntico al generado por los operadores de simetría del espacio P4₃, que es el que corresponde a la otra estructura, con un pequeño desplazamiento de las moléculas, lo que explicaría la pérdida de simetría (Figura 4.1). Por otro lado, hay diferentes moléculas de disolvente localizadas en el sitio de unión y en los contactos cristalinos en ambas estructuras, siendo ésta, por tanto, la única diferencia sustancial entre ellas.

Tabla 4.2. Colección de datos de difracción de rayos X y estadística de refinamiento para el dominio PDZ3-PSD95.

	Tetragonal			Triclínico		
Grupo espacial	P4 ₃			P1		
Dimensiones celdilla unidad a b c	42.98	42.98	46.89	43.32	42.15	47.60
Intervalo de resolución (Å)	20-1.4			90.00	90.00	92.08
Número de observaciones	16934(3138)			67847 (6858)		
Redundancia	5.88 (2.72)			2.4 (2.3)		
Completitud de los datos (%)	100 (100)			92.0 (92.9)		
R _{merge} ^a (%)	5.48 (30.8)			6.20 (41.1)		
I/σ(I)	17.05 (3.10)			16.07 (1.84)		
Coefficiente Matthews /Contenido de disolvente	1 mol/AU 1.93 / 36 %			4 mol/AU 1.93/36 %		
Refinamiento						
Residuos de proteína	98			392		
Moléculas de disolvente	97			250		
R _{factor} (%)	17.1 (24.8)			18.9 (28.2)		
R _{free} (%)	21.1 (26.3)			21.9 (29.1)		
Desviaciones de geometría ideal RMS						
Enlaces (Å)	0.024			0.022		
Ángulos (grados)	2.093			2.172		
B promedio para proteína (Å ²)	17.35			17.97		
Residuos en regiones más favorecidas del gráfico de Ramachandran (%) ^b	100			100		

Los valores entre paréntesis son para el intervalo de resolución más alta

$$^a R_{merge} = \left[\frac{\sum_h \sum_i |I_{hi} - \langle I_h \rangle|}{\sum_h \sum_i I_{hi}} \right] \cdot 100\% \text{ donde } \langle I_h \rangle \text{ es la intensidad media de la reflexión } i \text{ con intensidad } I_{ih}$$

^b Estadística obtenida mediante el programa PROCHECK

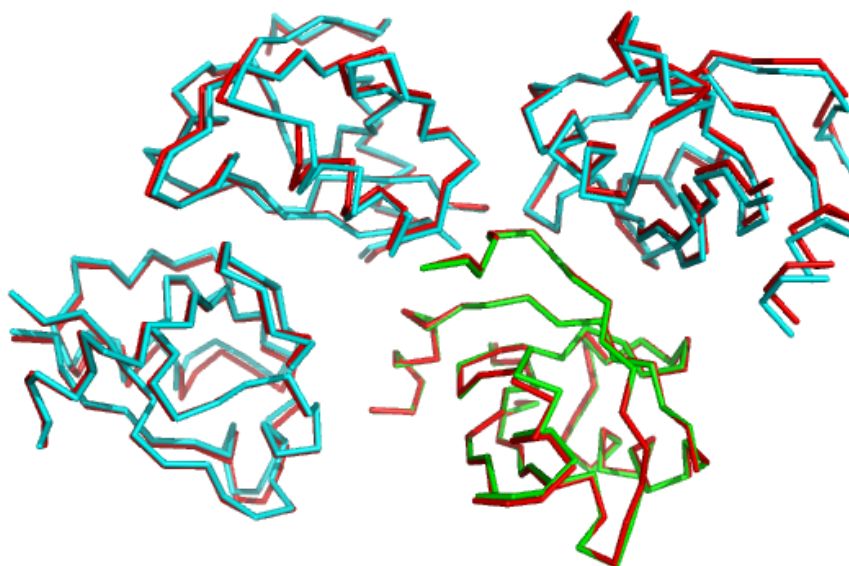


Figura 4.1. Superposición de la cadena principal de la estructura tetragonal del dominio PDZ3-PSD95 (color verde) y las cadenas del tetrámero obtenido por simetría (color cian) con la estructura triclinica (color rojo).

4.2 Discusión

4.2.1 Comparación de las estructuras del dominio PDZ3-PSD95 (restos 302-402) con las estructuras previamente publicadas (restos 302-415)

Cuando las estructuras recogidas en esta Memoria se comparan con las estructuras del dominio PDZ3-PSD95 resueltas hasta la fecha (que tienen los códigos PDB 1BE9, 1BFE, 1TP3, 1TP5, 1TQ3), se observan diferencias importantes. El conjunto de valores de RMSD que corresponde a posiciones de átomos de la cadena principal tiene unos valores de media entre 1.3 Å y 1.5 Å. Tal y como era esperable, las mayores diferencias en los valores de RMSD se encuentran en las regiones terminales de la proteína, aunque también hay diferencias significativas en los giros que conectan las cadenas $\beta 1$ y $\beta 2$ y las cadenas $\beta 2$ y $\beta 3$ (Figura 4.2).

En las estructuras previas del dominio PDZ3-PSD95, los aminoácidos presentes en el extremo C-terminal que forman la horquilla β antiparalela adicional presentan múltiples contactos con los residuos presentes en la hélice $\alpha 3$. De hecho, estos dos elementos estructurales presentan en las estructuras previas la mayor superficie de interacción en el cristal a través de enlaces de hidrógeno intra- e inter-catenarios y de puentes salinos, cubriendo un total de 700 Å² de superficie, dato que se obtuvo con el programa *PISA*, que permite obtener la superficie entre dos elementos estructurales. En las nuevas estructuras, la falta de esta secuencia adicional provoca un cambio de conformación de la hélice $\alpha 3$

(residuos 393-402, Figura 4.3), puesto que se producen cambios en los enlaces de hidrógeno y puentes salinos formados entre las cadenas laterales de los residuos.

Uno de los cambios más significativos en la hélice $\alpha 3$ es el del residuo Tyr397, que además se ha visto que es un sitio de fosforilación (Ballif et al., 2007) lo cual hace que este residuo tenga un papel importante en la regulación del dominio PDZ3-PSD95 *in vivo* en función de su accesibilidad a la acción de enzimas tipo kinasa (Petit et al., 2009). Los posibles residuos que podrían estabilizar este sitio de fosforilación mediante la formación de puentes salinos y enlaces de hidrógeno son los residuos Lys355, Ser 339 y Glu401, que son los que están cercanos a la Tyr397. En las estructuras previas, este residuo se localiza a la distancia de un enlace de hidrógeno de la cadena lateral del Glu401, pero esta interacción no se presenta en las estructuras recogidas en esta Memoria. En nuestras estructuras se comprueba que, a diferencia con las que presentan el artefacto producto de la clonación, el residuo Glu401, en vez de interactuar con el residuo Tyr397, se orienta de tal manera que forma una nueva interacción con el residuo Glu334. Como consecuencia de ello, la distancia entre el esqueleto de la hélice $\alpha 3$ y el giro que conecta las cadenas $\beta 2$ y $\beta 3$, formado por los residuos Gly329-Gly335, se acorta hasta un valor de 3 Å. Todo esto puede tener consecuencias importantes, puesto que la hélice $\alpha 3$, aunque está situada fuera del sitio activo y no hace contactos directos con el ligando peptídico, regula la unión de ligandos en este dominio (Petit et al., 2009). Además, como veremos en los Capítulos siguientes, la interacción entre los restos Glu334 y Glu401 puede tener una implicación importante en el equilibrio conformacional del dominio PDZ3-PSD95.

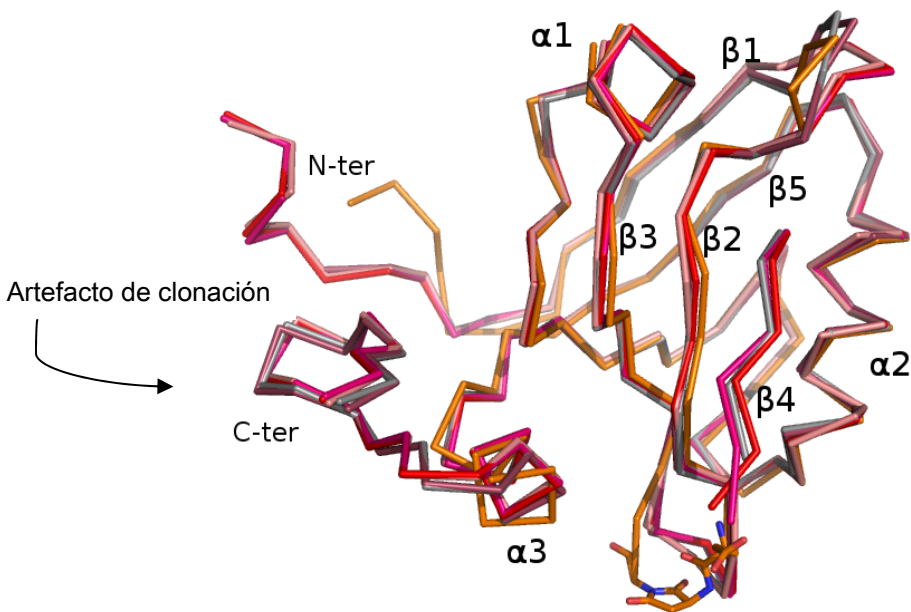


Figura 4.2. Superposición de la estructura tetragonal del dominio PDZ3-PSD95 obtenida en esta Memoria (color naranja) y las estructuras previas de este dominio: 1BE9 en rosa, 1BEF en gris, 1TP3 en magenta, 1TP5 en rojo y 1TQ3 en color morado. En el caso de las estructuras 1BE9, 1TP3 y 1TP5, en las que el dominio está formando un complejo con distintos tipos de péptidos, se ha representado la cadena principal del péptido en el mismo color de su estructura.

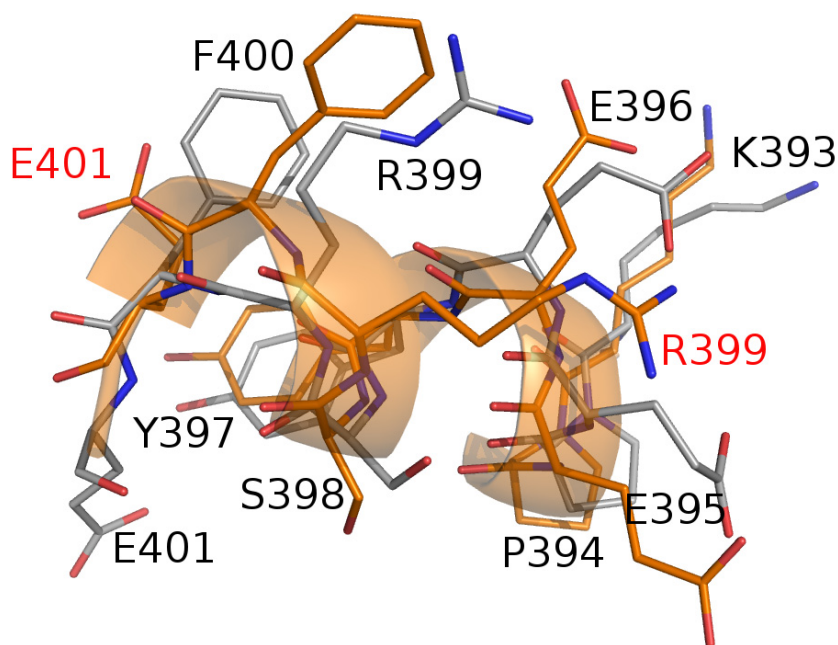


Figura 4.3. Comparación de la hélice α_3 de la estructura 1BEF (color gris), que no contiene un péptido unido con las estructuras de esta Memoria (color naranja).

4.2.2 Análisis de las consecuencias estructurales de la presencia de una succinimida en el resto aspartato de la posición 332

Tal y como se ha indicado en el análisis comparativo entre las estructuras previas del dominio PDZ3-PSD95 y las presentadas en esta Memoria, el giro que conecta las cadenas β_2 y β_3 formado por los residuos Gly329-Gly335 y que contiene la secuencia GGEDGEG, es la segunda región del dominio PDZ3-PSD95 que tiene valores de RMSD más altos que la media. De hecho, los residuos de este giro muestran orientaciones distintas a las de las cinco estructuras previas. Por otro lado, la cadena lateral del residuo Asp332 presente en dicho giro se ha modelado únicamente para la estructura previa de código PDB 1TQ3; sin embargo, la resolución mayor de las estructuras recogidas en esta Memoria (1.35 Å y 1.4 Å) ha hecho posible la identificación de un anillo tipo succinimida en la posición del residuo Asp322 en ambas estructuras (Figura 4.4).

La formación del anillo de succinimida es un paso intermedio común en los procesos de isomerización de los residuos tipo aspartato así como de deamidación de residuos asparragina en péptidos y proteínas. Su formación se da mediante el ataque nucleofílico del nitrógeno del enlace peptídico que está contiguo al carbonilo del residuo aspártico o asparragina (panel A de la Figura 4.5). Dicho ataque nucleofílico provoca la ciclación de la cadena (Clarke, 1987). El anillo succinimida formado es inestable en condiciones fisiológicas y se hidroliza para dar una mezcla de residuos tipo isoaspartato (panel B de la Figura 4.5) o aspartato (panel C de la Figura 4.5) en una relación 3:1 (Stephenson & Clarke, 1989).

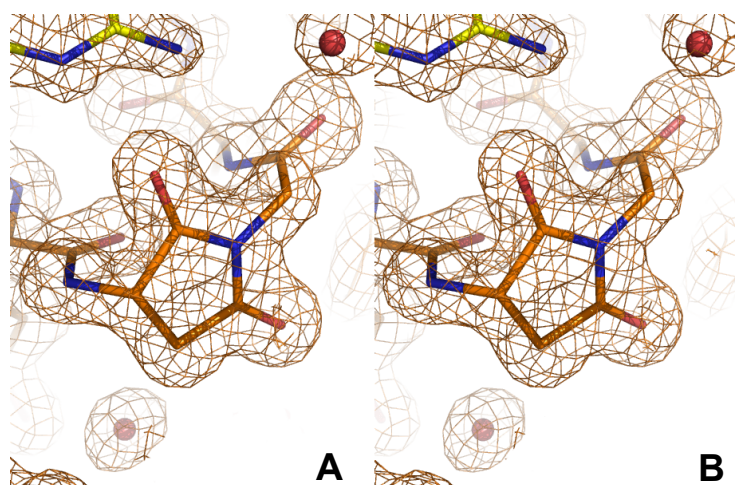


Figura 4.4. Mapa de densidad electrónica del giro que conecta las cadenas β_2 y β_3 de las dos estructuras. A y B corresponden a las estructuras tetragonal y triclínica, respectivamente.

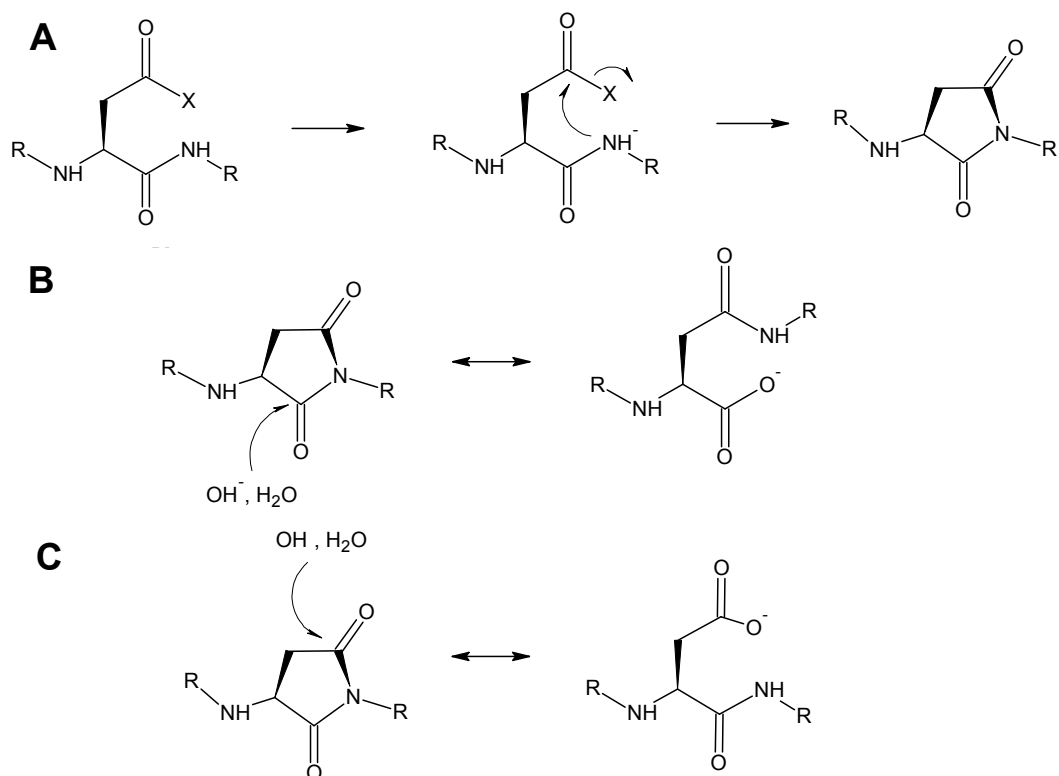


Figura 4.5. (A) Esquema de la ciclación de los residuos tipo aspartato (X corresponde al enlace $-\text{OH}$) o asparragina (X corresponde al enlace $-\text{NH}_2$) para formar el grupo succinimida. (B) Hidrólisis de la succinimida para formar el residuo tipo isoaspartato. (C) Hidrólisis de la succinimida para formar el residuo tipo aspartato.

Desde un punto de vista conformacional, la presencia de este anillo succinimida en el dominio PDZ3-PSD95 provoca la disminución de la flexibilidad del giro entre las cadenas $\beta 2$ y $\beta 3$. Un parámetro para reflejar esto es el factor B o factor de temperatura, que es una medida de la situación de un residuo en su posición que se obtiene como parte de los parámetros que caracterizan la estructura resuelta. Puesto que cualquier átomo presenta movimiento por el hecho de no estar a la temperatura del cero absoluto, no se encontrará en una posición fija en el espacio. De esta forma, cuanto más alto sea el factor B, mayor será el espacio que ocupa; es decir, será un residuo que esté menos fijo en su posición, siendo más flexible por tanto la zona en la que esté situado. Los residuos presentes en el giro $\beta 2$ - $\beta 3$, junto con los residuos del extremo N-terminal, muestran los valores de factor B más altos en todas las estructuras cristalográficas previas, lo que significa que esas dos regiones son las más flexibles del dominio. En cambio, en las estructuras recogidas en esta Memoria los valores del factor B correspondientes al giro $\beta 2$ y $\beta 3$ son menores que el valor medio del conjunto de valores del factor B, por lo que presenta una mayor rigidez que el resto del dominio.

El aumento de rigidez de esta región además favorece el proceso de cristalización, puesto que cuanto más rígida sea la estructura de la proteína, más fácil es de cristalizar (Vedadi et al., 2006; Price et al., 2009). La presencia de la succinimida favorece, por tanto, el proceso de cristalización debido a la pérdida de un grupo carboxilato, ya que esta pérdida provoca una disminución de la solubilidad de la proteína (Noguchi et al., 1998). De hecho, en nuestras estructuras del dominio PDZ3-PSD95 se observa que hay una participación directa del resto succinimida en los contactos cristalinos (Figura 4.6). Esto implica que este residuo debe estar presente en la disolución de proteína antes de que se forme el cristal, porque en el caso en que no estuviese, no podría participar en el contacto cristalino. Además, los enlaces de hidrógeno que se forman entre los residuos implicados en este contacto podrían evitar el ataque nucleofílico de moléculas de agua que iniciaría la hidrólisis de la succinimida, previniéndose de esta forma su conversión a aspartato o isoaspartato.

La presencia del anillo de succinimida se comprobó mediante espectrometría de masas. El peso molecular del dominio PDZ3-PSD95 es de 11004 Da (Figura 4.7), que corresponde al peso molecular del dominio intacto. Ahora bien, cuando uno de los cristales obtenidos mediante la condición 35 se disolvió, se obtuvo un peso molecular de 10986 Da (Figura 4.7), que corresponde a la pérdida de una molécula de agua (18 Da) debida a la ciclación. También se comprobó la ya mencionada inestabilidad de este tipo de anillo, puesto que la medida de la misma muestra transcurridas dos horas desde la disolución del cristal, dio de nuevo un peso molecular de 11004 Da (Figura 4.7).

La presencia del anillo de succinimida también podría influenciar la red alostérica del dominio PDZ3-PSD95 que modula las afinidades de unión del mismo (Petit et al., 2009), puesto que las interacciones entre los residuos polares localizados en la hélice $\alpha 3$ y los que están en el giro entre las cadenas $\beta 2$ y $\beta 3$ podrían también contribuir al posicionamiento de la hélice y, como consecuencia, modular la energética de unión. Esto convertiría a este residuo tipo succinimida en un posible elemento regulador de las

propiedades de unión de este dominio. No hay además ninguna evidencia de la presencia *in vivo* de un grupo isoaspartato o de una succinimida en la proteína PSD95, por lo que no se puede conocer si su presencia afectaría a su papel biológico; sin embargo, su influencia en el posicionamiento de la hélice $\alpha 3$ descrita en este Capítulo, junto con las consecuencias que podría tener en sus propiedades de unión, apunta a que podría ser un factor más de regulación de estas propiedades de unión. De hecho, la abundante presencia en la célula de la enzima D-aspartil/L-isoaspartil metiltransferasa (PIMT), cuyo papel es transformar residuos tipo isoaspartato en aspartato (Shimizu et al., 2005), indica que este tipo de residuos modificados podrían tener algún papel funcional. Sin embargo, son necesarias más evidencias experimentales que verifiquen el posible papel regulador de este tipo de residuos en la proteína PSD95.

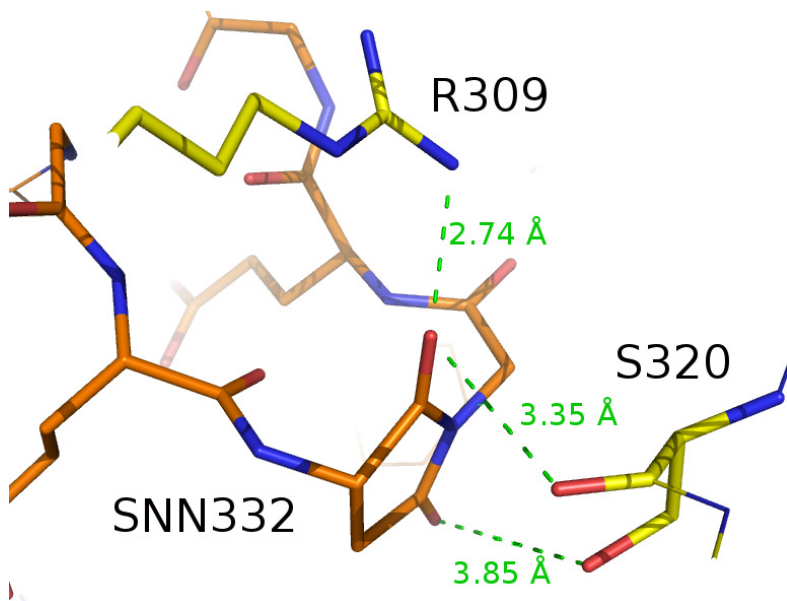


Figura 4.6. *Detalle de las interacciones del grupo succinimida con residuos de otras moléculas de PDZ3-PSD95 relacionadas mediante simetría como los residuos Arg309 y Ser320 coloreados en amarillo. Los enlaces de hidrógeno que estos residuos forman con los grupos carbonilo del grupo succinimida se representan en color verde.*

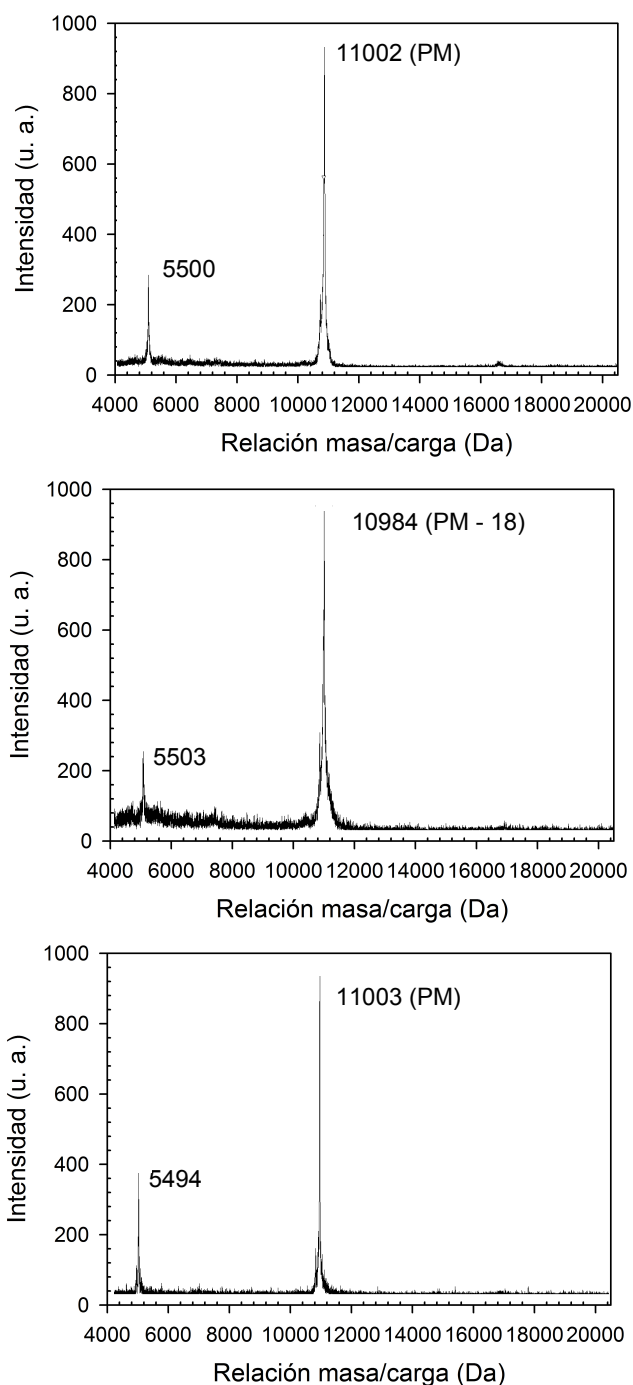


Figura 4.7. Espectros de masas del dominio PDZ3-PSD95 obtenidos antes de la formación del cristal (A), justo al disolver el cristal (B) y a las dos horas de la disolución del cristal (C). En el eje de abcisas se representa la relación carga masa (el Dalton, Da, como unidad), mientras que en el de ordenadas se representa la intensidad (en unidades arbitrarias). El pico de mayor peso molecular corresponde al dominio PDZ3-PSD95 con una carga positiva, mientras que el de menor peso molecular corresponde al dominio PDZ3-PSD95 con dos cargas positivas.

CAPÍTULO 5

ANÁLISIS TERMODINÁMICO Y CINÉTICO DEL
EQUILIBRIO CONFORMACIONAL DEL TERCER
DOMINIO PDZ DE LA PROTEÍNA PSD95 EN
CONDICIONES DE pH NEUTRO

5.1 Resultados

El estudio conformacional de los dominios modulares que componen una proteína *hub* puede facilitar la comprensión de las causas de la variada funcionalidad que dichas proteínas presentan. Para ello, en esta Tesis Doctoral se ha abordado el estudio del equilibrio conformacional del dominio PDZ3-PSD95 a pH neutro usando distintas técnicas: calorimétricas (DSC), espectroscópicas (CD, fluorescencia), ópticas (DLS, TEM) y cromatográficas (SEC).

5.1.1 Mecanismo de desplegamiento térmico del dominio PDZ3-PSD95 a pH neutro

Hasta ahora, todos los estudios de plegamiento en condiciones de equilibrio que se habían llevado a cabo con diferentes dominios PDZ han mostrado un mecanismo simple de dos estados. Esto se deducía a partir de las curvas sigmoideas obtenidas al usar técnicas espectroscópicas, tales como el CD o la fluorescencia, y seguidas tanto con desnaturalización térmica como química (Sicorello et al., 2009; Gianni et al., 2005; Feng et al., 2005; Chi et al., 2007). Para corroborar este comportamiento, se realizó un barrido de temperatura con el dominio PDZ3-PSD95 seguido por CD desde 2°C hasta 98°C a una velocidad de barrido de 1.5 K·min⁻¹.

Para hacer el barrido de temperatura seguido por CD, se dializó la proteína en tampón fosfato potásico 50 mM pH 7.5. La concentración final del dominio PDZ3-PSD95 fue de 0.3 mg·mL⁻¹. Estas condiciones se eligieron para poder comparar con los resultados obtenidos en bibliografía, en los que se siguió el desplegamiento térmico mediante CD del mismo dominio PDZ en ese tampón (Chi et al., 2007). En esta Memoria los resultados han sido equivalentes a los estudios anteriores, puesto que se observó también la curva sigmoide, que se analizó con el modelo de dos estados (Figura 5.1).

Sin embargo, esta única prueba no es concluyente. Es bien conocido que la técnica de DSC es la óptima para detectar cualquier conformación en el equilibrio que se pueda poblar durante el desplegamiento. Por ello se llevó a cabo un estudio de desplegamiento mediante esta técnica en las mismas condiciones de tampón que para el CD. En este caso, la concentración de proteína fue de 1.6 mg·mL⁻¹, puesto que la calorimetría requiere una mayor concentración de muestra que el CD. El intervalo de temperatura usado fue mayor, desde 5°C hasta 110°C, puesto que las características del instrumento de DSC así lo permiten. Se comprobó la reversibilidad del proceso mediante un segundo barrido de forma consecutiva. Se vio que la traza presenta dos transiciones y que la reversibilidad es de un 60% aproximadamente (Figura 5.2).

Para comprobar la reversibilidad del proceso se usó otra técnica complementaria, el CD. Se registraron espectros del dominio PDZ3-PSD95 a una concentración de 0.3 mg·mL⁻¹ en fosfato 50 mM pH 7.5 a 2°C antes y después de calentar hasta 98°C. El resultado fue que recupera el espectro del estado nativo en cinco minutos (Figura 5.3). Todos estos resultados indican que los tiempos de relajación de estos fenómenos son

lentos, pero no contradicen su naturaleza de equilibrio a pH neutro. Por lo tanto se puede concluir que se puede abordar el análisis del equilibrio conformacional.

Para poder aplicar un modelo de equilibrio para el análisis de los datos de DSC ha de comprobarse que las transiciones calorimétricas no dependen de la velocidad de barrido de temperatura (Lepock et al., 1992; Mayorga & Freire, 1987). Los estados conformacionales de equilibrio son independientes de las condiciones mediante las cuales se alcanzan. Entre esas condiciones está la velocidad de barrido, por tanto, si dependen de ésta, el proceso no está en el equilibrio. Para comprobarlo, se hicieron los experimentos a distintas velocidades de barrido, en concreto a $2 \text{ K}\cdot\text{min}^{-1}$, $1 \text{ K}\cdot\text{min}^{-1}$, $0.5 \text{ K}\cdot\text{min}^{-1}$ y $0.25 \text{ K}\cdot\text{min}^{-1}$, manteniendo la misma concentración de proteína y en las mismas condiciones de tampón. Las trazas calorimétricas no presentan diferencias apreciables entre ellas, puesto que sus formas son prácticamente idénticas, a la vez que no hay un desplazamiento en los máximos de las transiciones. El hecho de que no sean totalmente superponibles ha de deberse a errores debidos al instrumento (Figura 5.2).

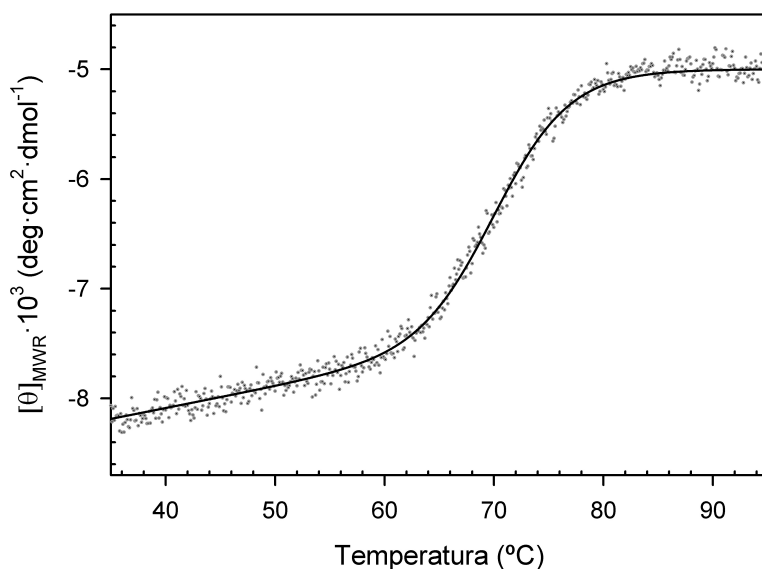


Figura 5.1. Curva de desdoblamiento térmico seguido por CD a 218 nm del dominio PDZ3-PSD95 a una concentración de $0.3 \text{ mg}\cdot\text{mL}^{-1}$ en fosfato 50 mM pH 7.5. Los puntos representan los datos experimentales, mientras que la línea es el ajuste al modelo de dos estados.

Las dos transiciones que se observan en las trazas de DSC indican sin ninguna duda la existencia de no menos de tres estados macroscópicos bien definidos durante el desdoblamiento. Por tanto tendríamos al menos el estado nativo, N; el desnaturalizado, D, y un estado intermedio, I. Además, al llevar a cabo los barridos de temperatura con distintas concentraciones de proteína, se puede observar que ambas endotermas se distancian entre sí a lo largo del eje de la temperatura conforme aumenta la concentración de proteína, sugiriendo un cierto grado de asociación del estado intermedio (Figura 5.4).

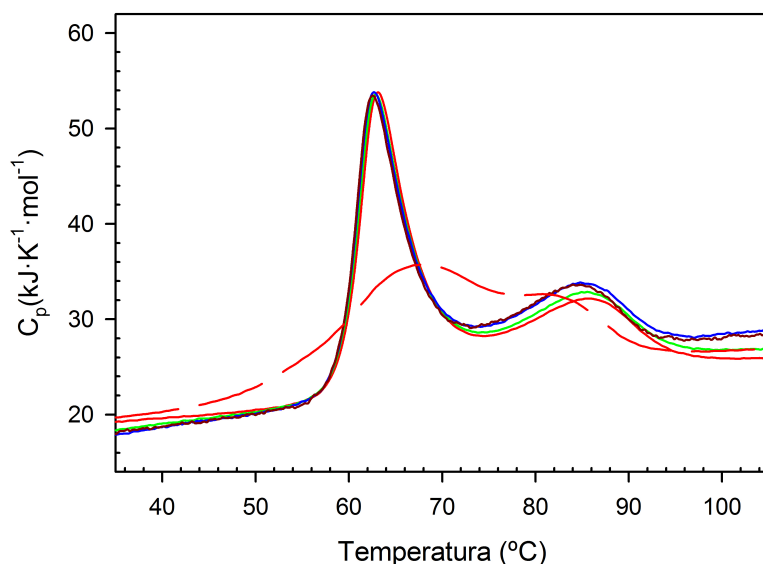


Figura 5.2. Curvas de DSC a distintas velocidades de barrido para el dominio PDZ3-PSD95 a una concentración de $1.6 \text{ mg}\cdot\text{mL}^{-1}$ en fosfato 50 mM pH 7.5 . Las velocidades son $2 \text{ K}\cdot\text{min}^{-1}$ (naranja), $1 \text{ K}\cdot\text{min}^{-1}$ (verde), $0.5 \text{ K}\cdot\text{min}^{-1}$ (azul) y $0.25 \text{ K}\cdot\text{min}^{-1}$ (marrón). El segundo barrido de la traza de $2 \text{ K}\cdot\text{min}^{-1}$ a está representado con una línea discontinua en color naranja. Para simplificar la gráfica sólo se ha mostrado el segundo barrido del experimento realizado a $2 \text{ K}\cdot\text{min}^{-1}$.

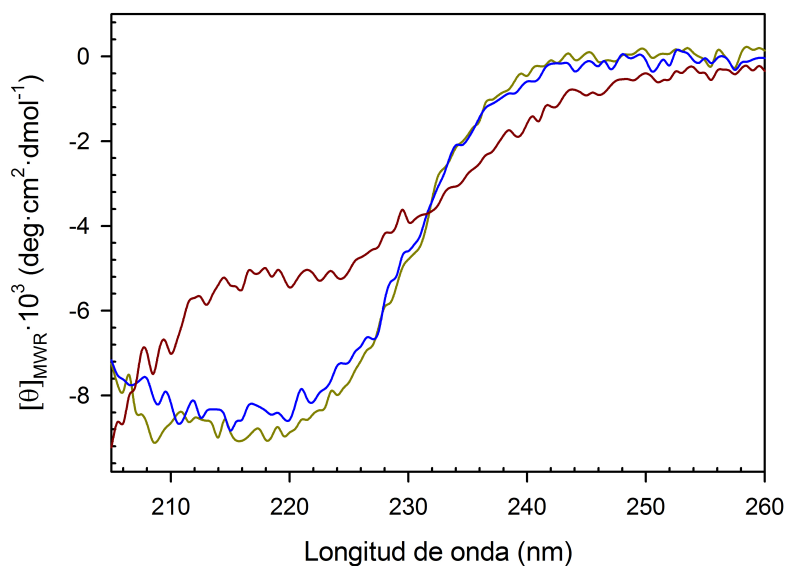


Figura 5.3. Espectros de CD del dominio PDZ3-PSD95 a una concentración de $0.3 \text{ mg}\cdot\text{mL}^{-1}$ en fosfato 50 mM pH 7.5 . En verde se representa el espectro de la proteína nativa a 2°C y en rojo el espectro a 98°C . Para comprobar la reversibilidad se enfrió la muestra dentro del dicrógrafo hasta 2°C y se registró otro espectro a los 5 minutos después de haber mantenido esta temperatura (en azul).

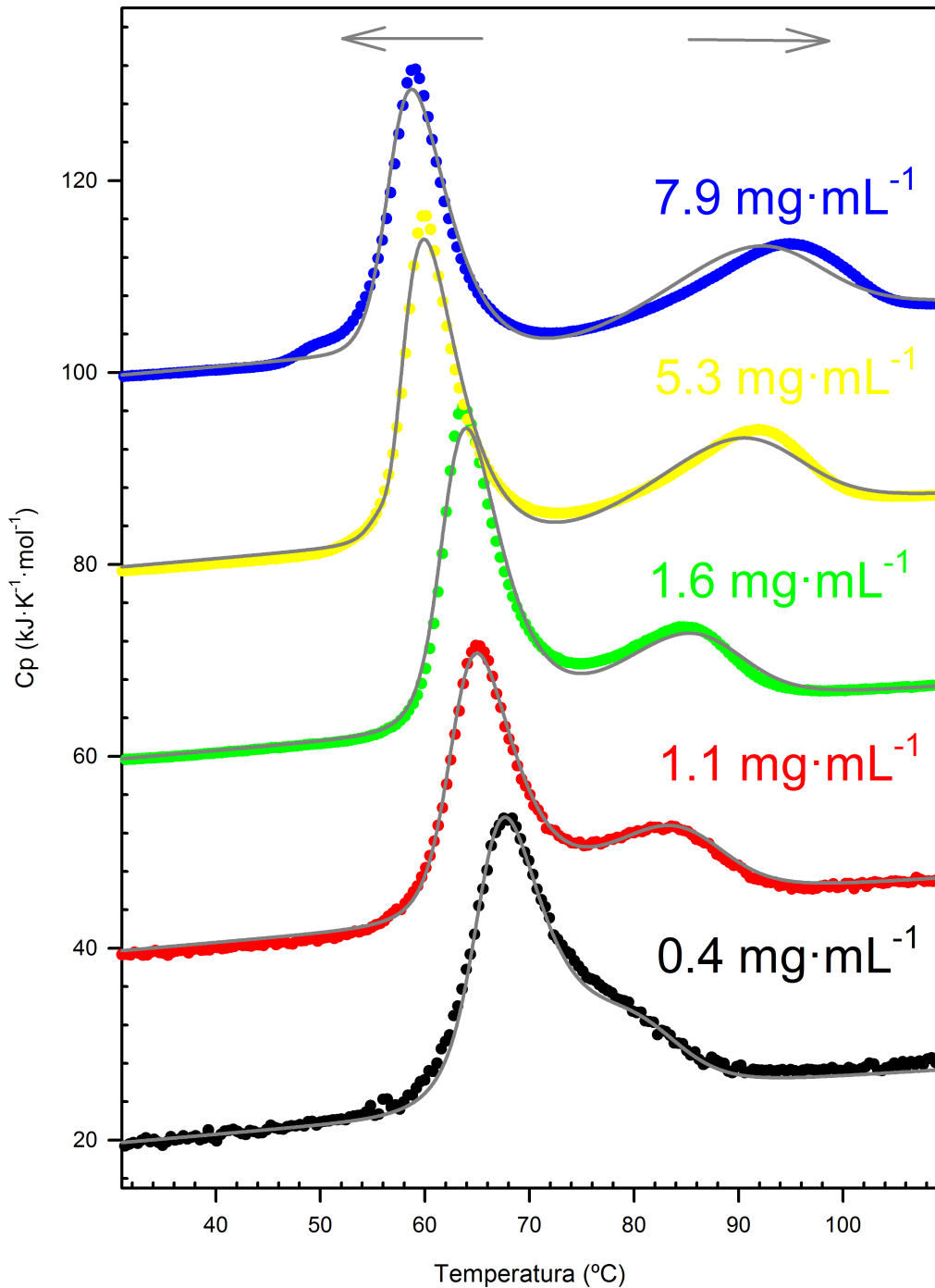


Figura 5.4. Curvas de desplegamiento térmico del dominio PDZ3-PSD95 seguidas por DSC en función de la concentración de proteína en fosfato 50 mM pH 7.5. Las flechas indican el desplazamiento de ambas endotermas con respecto al aumento de concentración. La línea en color gris representa el ajuste global al modelo $3N \rightleftharpoons I_3 \rightleftharpoons 3D$.

El efecto de la concentración de proteína en trazas de DSC ha sido descrito en bibliografía con otros ejemplos, y consiste en el aumento de la población de las especies asociadas conforme aumenta la concentración de proteína (Filimonov & Rogov, 1996; Ruiz-Sanz et al., 2004). De esta forma, el movimiento de la primera transición (que corresponde al equilibrio $N \rightleftharpoons I$) hacia menores temperaturas al aumentar la concentración indica un aumento del estado de asociación del intermedio durante el desplegamiento. Ocurre lo contrario en la segunda transición, correspondiente al equilibrio $I \rightleftharpoons D$, indicando así una disminución del estado de asociación del estado final. De esta forma, se puede establecer que el desplegamiento del dominio PDZ3-PSD95 se podría resumir mediante el esquema $nN \rightleftharpoons I_n \rightleftharpoons nD$.

De hecho, este comportamiento podría explicar la única transición aparente cuando se sigue el desplegamiento por técnicas espectroscópicas, debido a las bajas concentraciones que se usan en dichas técnicas, alrededor de $0.3 \text{ mg} \cdot \text{mL}^{-1}$ en el caso de los experimentos de CD realizados en esta Memoria. A esa concentración, el intermedio asociado no se puebla de forma suficiente como para ser detectado, cosa que está de acuerdo con la forma de la traza de DSC a la concentración de $0.4 \text{ mg} \cdot \text{mL}^{-1}$, que se parece más a una transición única que aquellas trazas en las que la concentración es mayor, donde ambas endotermas se distinguen muy bien.

5.1.2 Estudio de la formación reversible de agregados globulares y fibrilares a pH neutro a partir del estado intermedio

Para caracterizar los equilibrios de asociación-disociación postulados se llevaron a cabo experimentos de DLS en función de la temperatura, puesto que esta técnica permite conocer los radios hidrodinámicos de las partículas presentes en una disolución. Con esos valores de radios mediante el programa que controla el instrumento (Dynamics V6, ver Sección 3.6), se pueden hacer estimaciones de los pesos moleculares de dichas partículas, con lo que se pueden definir posibles grados de asociación. De esta forma, una disolución del dominio PDZ3-PSD95 a una concentración de $8 \text{ mg} \cdot \text{mL}^{-1}$ dializada en tampón fosfato 50 mM pH 7.5, se calentó desde 20°C hasta 60°C a una velocidad de barrido de $1 \text{ K} \cdot \text{min}^{-1}$. La temperatura de 60°C es la máxima que puede alcanzar el instrumento; sin embargo, es lo suficientemente alta como para asegurar que la especie I_n se puebla, de acuerdo con el análisis de las trazas de DSC (ver siguiente Sección, 5.1.3).

Según los datos del barrido de temperatura, el tamaño aproximado de la partícula presente en disolución a 20°C presenta un radio hidrodinámico de 1.8 nm, valor que se corresponde con el peso molecular de un monómero, 11 kDa. Sin embargo, cuando se alcanzan los 60°C , el tamaño aumenta hasta llegar a unos 2.6 nm, que corresponde a un peso molecular de 33 kDa, siendo por tanto un oligómero trimérico (Figura 5.5).

Una vez terminado el barrido de temperatura, se mantuvo constante la temperatura del instrumento a 60°C para seguir la evolución de la masa de las distintas especies (intensidad relativa ponderada de cada una) en función del tiempo de incubación. Se observó que conforme va aumentando el tiempo de incubación aparecen otras especies de

mayor tamaño (unos 12 nm), las cuales pasan a ser el componente mayoritario una vez que han pasado unas 2 horas (Figura 5.6). Estas partículas siguen aumentando en tamaño cuando se dejan tiempos mayores de incubación (Tabla 5.1).

A continuación, se bajó la temperatura del instrumento de nuevo a 20°C para comprobar la reversibilidad del proceso. Se observó que aparecen de nuevo las partículas de 12 nm, y que se alcanza el estado monomérico después de varias horas (Figura 5.7 y Tabla 5.2). Estos resultados demuestran la naturaleza reversible de estos fenómenos de asociación.

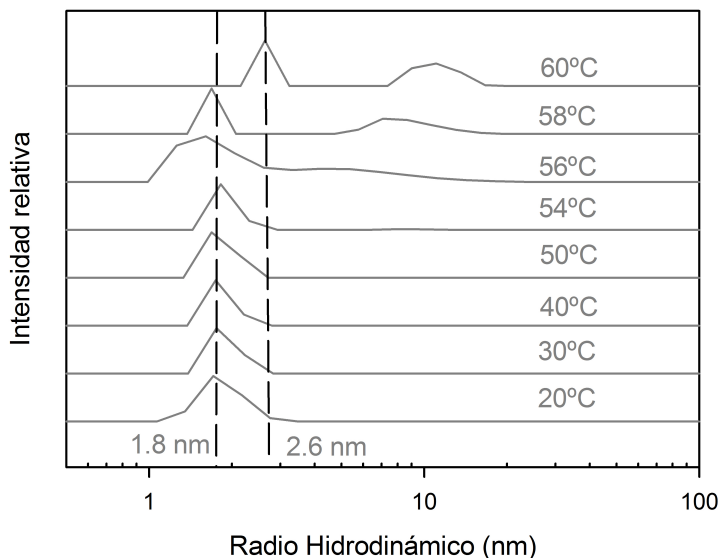


Figura 5.5. Barrido de temperatura desde 20°C hasta 60°C del dominio PDZ3-PSD95 a una concentración de $8 \text{ mg}\cdot\text{mL}^{-1}$ en fosfato 50 mM pH 7.5 seguido por DLS. La velocidad de barrido fue de $1 \text{ K}\cdot\text{min}^{-1}$.

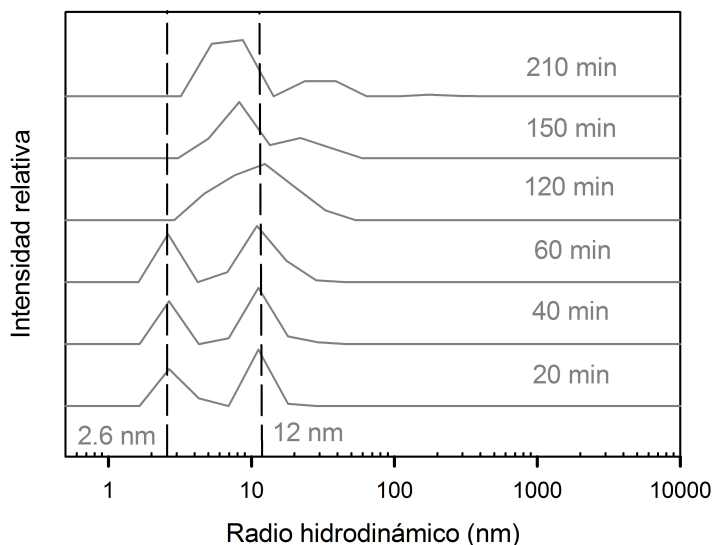


Figura 5.6. Evolución a 60°C del radio hidrodinámico en función del tiempo de las especies presentes en una disolución del dominio PDZ3-PSD95 a una concentración de $8 \text{ mg}\cdot\text{mL}^{-1}$ y pH neutro seguida por DLS.

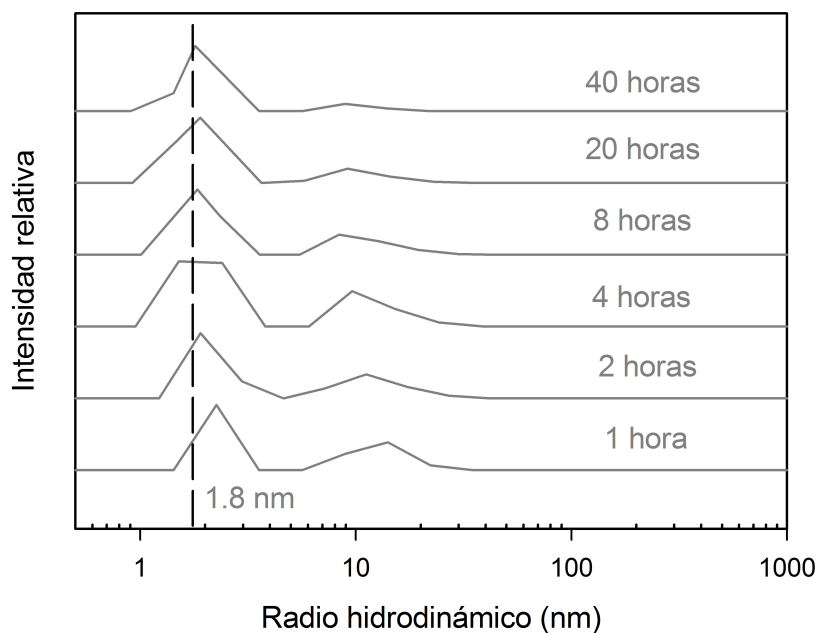


Figura 5.7. Evolución a 20°C del radio hidrodinámico en función del tiempo de las especies presentes en una disolución del dominio PDZ3-PSD95 a una concentración de $8 \text{ mg}\cdot\text{mL}^{-1}$ en fosfato 50 mM pH 7.5 seguida por DLS.

Tabla 5.1. Poblaciones de las distintas especies detectadas por DLS que presenta el dominio PDZ3-PSD95 cuando se incubaba a 60°C durante tres horas y media en tampón fosfato 50 mM pH 7.5.

Tiempo (min)	Radio (nm)	% Masa	Peso Molecular (kDa)
20	2.6	43	32
	11.4	57	1004
40	2.7	37	33
	12	63	1133
60	2.6	35	32
	12.4	65	1219
120	12.7	100	1280
150	8.3	75	473
	26.9	25	7486
210	7	77	323
	31.5	21	10612
	199.4	2	816146

Tabla 5.2. Poblaciones de las distintas especies detectadas por DLS que presenta el dominio PDZ3-PSD95 a 20°C a lo largo del tiempo después de haber incubado a 60°C durante tres horas y media en tampón fosfato 50 mM a pH 7.5.

Tiempo (horas)	Radio (nm)	% Masa	Peso Molecular (kDa)
1	2.2	57	23
	13.2	43	1408
2	2.1	63	20
	12.7	37	1279
4	1.9	69	14
	12.5	31	1253
8	1.8	72	12
	11.6	28	1054
20	1.9	81	14
	10.9	19	900
40	1.8	89	12
	8.9	11	560

Para seguir la formación de estas estructuras supramacromoleculares mediante TEM se incubaron muestras durante diferentes tiempos a 60°C, desde pocos días hasta un mes. Se puede observar en la Figura 5.8 la evolución de las distintas especies a lo largo del tiempo. De acuerdo con las evidencias mostradas mediante DLS, se pueden observar unas formas globulares de unos 10 a 20 nm con bajos tiempos de incubación (micrografía señalada con “1-3 días” de la Figura 5.8), las cuales, conforme aumenta el tiempo de incubación se van haciendo más grandes, unos 50 nm (micrografías “8 días”, “15 días (A)” de la Figura 5.8). Es difícil saber la forma tridimensional de estas partículas con el microscopio electrónico; probablemente sean anulares, por el oscurecimiento de los círculos hacia el centro. Podrían ser también esferas, aunque en principio el oscurecimiento de este tipo de partículas debería ser en la dirección contraria. Por otro lado, tal y como se ve en la micrografía señalada como “15 días (A)” de la Figura 5.8, estas partículas globulares coexisten con otras estructuras, como las protofibras o incluso con fibras, dándose el caso además de que a tiempos de incubación mayores de 15 días muestran solamente las fibras (micrografía señalada con “1 mes” de la Figura 5.8).

Estas fibras también coexisten reversiblemente con las estructuras circulares porque desaparecen de las disoluciones incubadas un mes cuando se diluyen 1:1 con tampón y se enfrían a temperatura ambiente. Además, cuando esas muestras diluidas se volvieron a incubar a 60°C durante unas 15 horas, las partículas redondeadas de 12 nm aparecieron de nuevo (micrografía señalada con “1 mes (dilución 1:1)” de la Figura 5.8). Sin embargo, la reversibilidad del proceso no parece que sea de un 100 %, puesto que algunos agregados permanecen en disolución después de la dilución (micrografía señalada con “1 mes (dilución 1:1)” de la Figura 5.8). Puede ocurrir que la irreversibilidad aumente

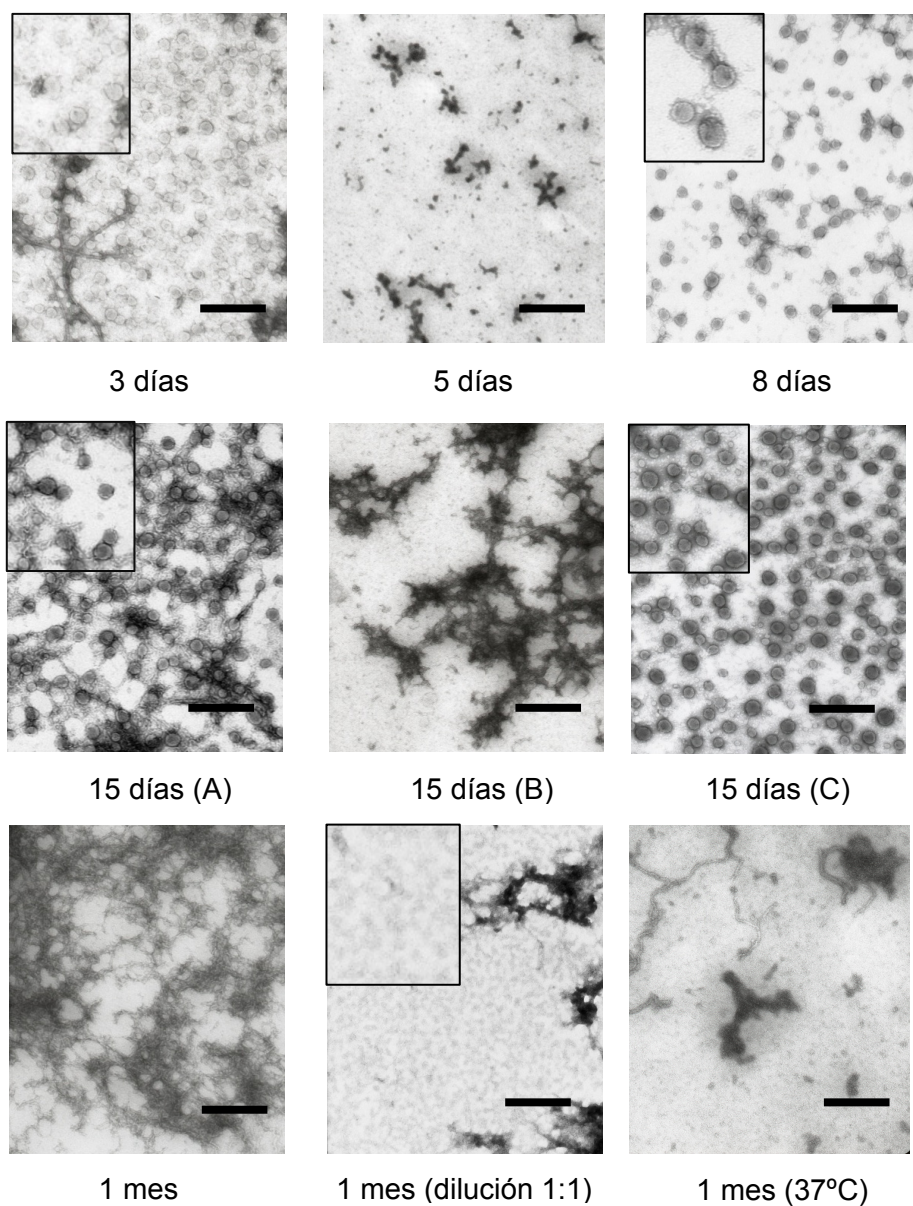


Figura 5.8. Análisis de las estructuras macromoleculares del PDZ3-PSD95 seguidas por TEM. En el pie de cada micrografía está indicado el tiempo de incubación a 60°C de la disolución de proteína a $8 \text{ mg} \cdot \text{mL}^{-1}$ en fosfato 50 mM pH 7.5, excepto para la señalada como “1 mes (37°C)”, que se incubó a 37°C. Para poder observarlas mejor, en aquellas micrografías en que aparecen las estructuras globulares se ha hecho un pequeño aumento en la esquina superior izquierda. La micrografía que está indicada con “1 mes (dilución 1:1)” fue tomada después de 15 horas de re-incubación de una dilución 1:1 con tampón de una muestra con un tiempo de incubación de un mes. La barra de escala corresponde a un tamaño de 200 nm.

conforme aumenta el tamaño de las fibras y su madurez temporal, tal y como está descrito en otros ejemplos (Calamai et al., 2005).

Por lo tanto, el intermedio oligomérico se organiza en partículas globulares y protofibras en las primeras etapas de incubación, permaneciendo además tras dos semanas de incubación, después de las cuales pasan a ser fibras. Lo más interesante de estos resultados es la coexistencia de tanto las fibras como las estructuras anulares o esféricas en la misma disolución del dominio PDZ3-PSD95 en fosfato 50 mM pH 7.5. La formación de las fibras se da a partir de las partículas circulares como muestra el microscopio (micrografía señalada con “3 días” de la Figura 5.8), puesto que éstas no se observaron cuando la muestra se incubó durante largos períodos de tiempo (micrografía señalada con “1 mes” de la Figura 5.8). Por otro lado, también se incubó durante un mes una muestra en las mismas condiciones experimentales variando únicamente la temperatura, 37°C, por ser la temperatura fisiológica. El resultado fue observar pocas fibras y muy finas (micrografía señalada con “1 mes (37°C)” de la Figura 5.8).

5.1.3 Aspectos conformacionales y energéticos del intermedio oligomérico obtenidos por el análisis del equilibrio del desplegamiento del dominio PDZ3-PSD95

Se llevó a cabo un análisis de los experimentos de DSC con el modelo de tres estados con asociación del intermedio. Al esquema inicial $nN \rightleftharpoons I_n \rightleftharpoons nD$, que se ha definido en la Sección 5.1.1, se le varió el orden para que los parámetros termodinámicos obtenidos del análisis sean más intuitivos y comparables. Según el esquema $N \rightleftharpoons D \rightleftharpoons (1/n)I_n$, los parámetros de la primera transición son los asociados a los de la desnaturalización de la proteína ($N \rightleftharpoons D$) y los de la segunda son los correspondientes a los procesos de asociación ($D \rightleftharpoons (1/n)I_n$) (para ver más detalles sobre el modelo, ver Sección 3.4.2.2).

El desplazamiento relativo de las trazas se debe únicamente a las diferentes concentraciones de proteína, siempre y cuando se mantengan constantes el resto de condiciones experimentales. Así, haciendo un ajuste global de las trazas se puede llegar a conocer la estequiometría global del equilibrio (Filimonov & Rogov, 1996; Ruiz-Sanz et al., 2004). En una primera aproximación, se consideró para el análisis matemático $n = 3$, puesto que los resultados obtenidos mediante DLS indican que el intermedio ha de ser trimérico, ya que durante los primeros veinte minutos de incubación la especie mayoritaria tiene un radio hidrodinámico de 2.6 nm, que corresponde a un peso molecular de 33 kDa, correspondiente a un trímero (Tabla 5.1). Ese período de tiempo podría considerarse un tiempo equivalente al que está en la célula del calorímetro, ya que la velocidad de barrido usada fue $1.5 \text{ K} \cdot \text{min}^{-1}$.

Para realizar el análisis se tomaron como comunes las capacidades caloríficas de cada estado conformacional así como los parámetros del proceso de disociación. Sin embargo, no fue así con las entalpías y temperaturas correspondientes a la primera parte del esquema considerado ($N \rightleftharpoons D$), las cuales se realizaron individualmente para obtener el error estándar asociado a este valor según la media de los parámetros obtenidos

individualmente. Los resultados del análisis indican que el modelo realmente está describiendo el comportamiento experimental, debido a la calidad del ajuste (Figura 5.9 y Tabla 5.3).

Por otro lado, las trazas experimentales de DSC se ajustaron también al modelo dejando como flotante el parámetro n . Los resultados de este análisis dieron un valor $n = 3.8 \pm 0.3$, dando además los mejores valores de R^2 (0.996; Figura 5.9). De hecho, el ajuste con un valor fijo de $n = 4$ converge algo mejor que aquel llevado a cabo con $n = 3$ (Figura 5.9). Se observa que hay pequeñas desviaciones en las segundas endotermas de las trazas que corresponden a concentraciones más grandes de proteína; sin embargo, el ajuste con $n = 3$ reproduce mejor la forma de la traza de $0.4 \text{ mg}\cdot\text{mL}^{-1}$. El hecho de que el ajuste de las trazas de mayor concentración sea mejor a valores de n mayores, y las de menor concentración se ajusten mejor a valores de n menores, puede deberse a que cuando la concentración de proteína aumenta, las especies oligoméricas se pueblan más. Por otro lado, el uso de los valores de $n = 2$ y $n = 6$, dio lugar a peores ajustes (Figura 5.9), tanto, que no se pudieron ajustar globalmente las cinco trazas, sino únicamente tres de ellas. Finalmente, los parámetros termodinámicos de la Tabla 5.3 han sido calculados a partir de las sesiones de ajuste usando $n = 3$, cuyos intervalos de error han sido estimados comparando los parámetros obtenidos tanto con $n = 3$ como con $n = 4$. Estos errores son un poco más altos y posiblemente sean más realistas que los errores estándar del ajuste.

Para obtener más evidencias experimentales de estos procesos de asociación, se llevaron a cabo experimentos de SEC en fosfato 50 mM pH 7.5 con la columna termostaticada a 60°C de muestras de PDZ3-PSD95 a una concentración de $8 \text{ mg}\cdot\text{mL}^{-1}$ que fueron incubadas a 60°C durante distintos períodos de tiempo (Figura 5.10). Aparece un pico ancho, probablemente debido a la presencia de una mezcla de tamaños, aunque los máximos que aparecen corresponden a la especie monomérica, agregados grandes que salen en el volumen de exclusión y, de nuevo, a una especie trimérica.

Tabla 5.3. Parámetros termodinámicos del desplegamiento térmico del dominio PDZ3-PSD95 en tampón fosfato 50 mM pH 7.5 obtenidos del análisis de experimentos de DSC analizados según el modelo $3N \rightleftharpoons I_3 \rightleftharpoons nD$.

T_{N-D} ($^\circ\text{C}$)	$\Delta H_{N-D}(T_{N-D})$ ($\text{kJ}\cdot\text{mol}^{-1}$)	$\Delta G_{N-D}(298)$ ($\text{kJ}\cdot\text{mol}^{-1}$)	T_{D-In} ($^\circ\text{C}$)	$\Delta H_{D-In}(T_{D-In})$ ($\text{kJ}\cdot\text{mol}^{-1}$)	$\Delta G_{D-In}(343)$ ($\text{kJ}\cdot\text{mol}^{-1}$)
70.4 ± 0.5	335 ± 20	39 ± 6	79.2 ± 1.2	-130 ± 20	25 ± 5

Los intervalos de error han sido calculados tal y como indica el texto. Los valores de las magnitudes termodinámicas del estudio han sido estimados con una $P_{\text{ref}} = 100 \mu\text{M}$. Las funciones de capacidad calorífica obtenidas del ajuste fueron $C_{pN} = -9.23 + (0.095^\circ\text{T})$; $C_{pD} = 2.91 + (0.064^\circ\text{T})$; $C_{pIn} = -99.0 + (0.347^\circ\text{T})$ en $\text{kJ}\cdot\text{mol}^{-1}$.

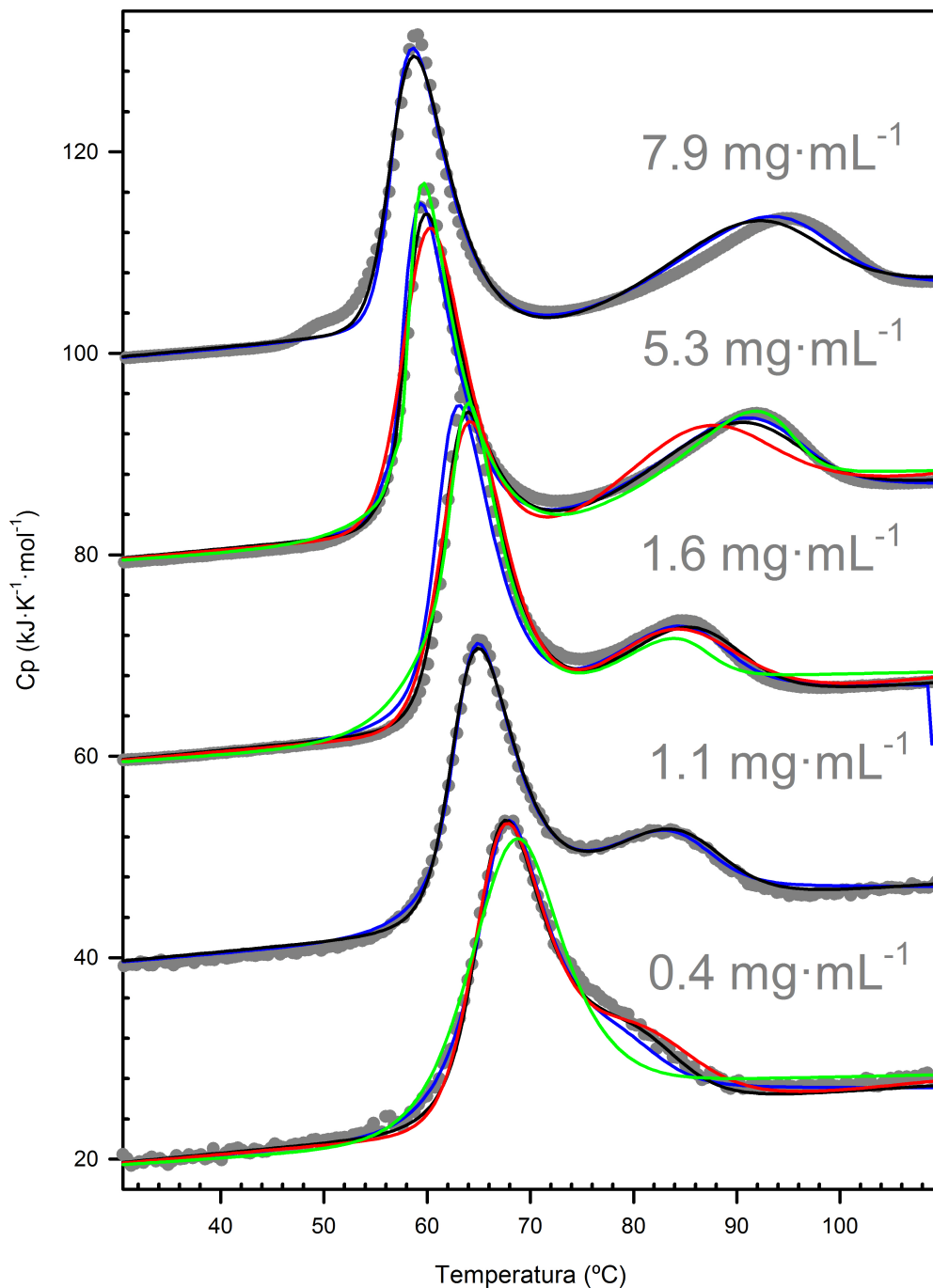


Figura 5.9. Curvas de desplegamiento térmico del dominio PDZ3-PSD95 seguidas por DSC en función de la concentración de proteína en fosfato 50 mM pH 7.5 analizadas según el modelo $nN \rightleftharpoons I_n \rightleftharpoons nD$ usando distintas estequiometrías. Los puntos experimentales son los círculos grises, mientras que los ajustes globales al modelo se representan mediante líneas de distintos colores en función del estado de asociación, n . Los colores son rojo para $n=2$, negro para $n=3$, azul para $n=4$ y verde para $n=6$.

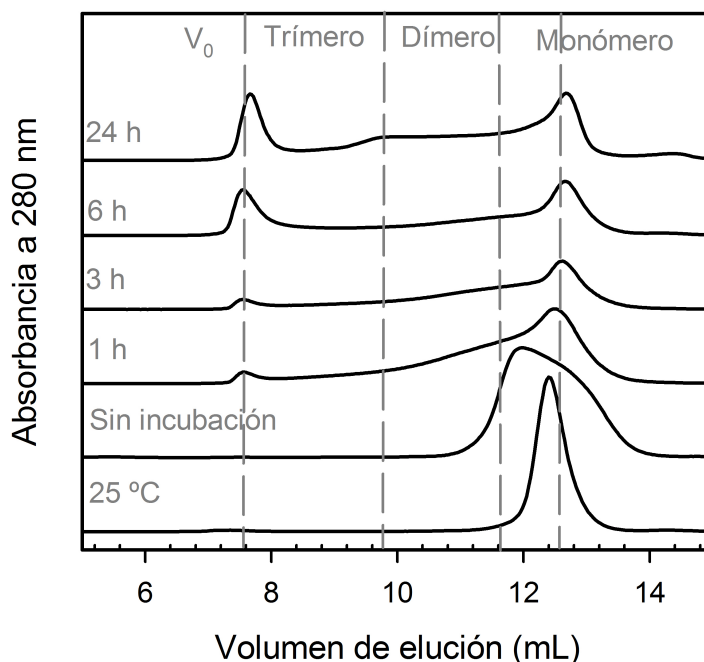


Figura 5.10. Distintas muestras del dominio PDZ3-PSD95 a una concentración de $8 \text{ mg}\cdot\text{mL}^{-1}$ en fosfato 50 mM pH 7.5 fueron incubadas a 60°C a distintos intervalos de tiempo (indicados en la Figura) e inyectados a continuación en la columna Superdex 75 HR 10/30 termostatzada a su vez a 60°C . También se inyectó una muestra con la columna a 25°C (primera curva). Las líneas verticales corresponden al volumen de elución de las especies monoméricas, diméricas y triméricas del dominio PDZ3-PSD95, siendo V_0 el volumen de exclusión de la columna (ver Sección 3.10 del Capítulo 3).

La hipotética naturaleza trimérica de las especies oligoméricas ya ha sido descrita en otras ocasiones (Quist et al., 2005; Güthe et al., 2004), y se puede explicar estructuralmente mediante un análisis computacional de la tendencia que tiene la secuencia del dominio PDZ3-PSD95 a dar agregación tipo β . Para ello se ha usado el algoritmo TANGO desarrollado por Serrano y colaboradores (Fernandez-Escamilla et al., 2004), cuyos cálculos están mostrados en la Figura 5.11. En ellos se muestra que la tendencia es prácticamente nula en todo el dominio con la excepción de los residuos 340 a 350 (que pertenecen a los motivos $\beta 3$ y $\alpha 1$) y los residuos 385-392 (que están localizados en la lámina $\beta 5$ y la hélice $\alpha 3$). La disposición tridimensional de estos residuos en la estructura resuelta a alta resolución a partir de la difracción de Rayos X (ver Capítulo 4) muestra dos superficies principales localizadas en caras opuestas a la estructura globular, que puede permitir a cada monómero interactuar con otros dos monómeros adicionales (Figura 5.12). Por lo tanto, es posible que el aumento en la exposición al disolvente de estas regiones con propensión a la agregación durante el desplegamiento pueda hacer que los monómeros del dominio PDZ3-PSD95 desarrollen cierta tendencia a auto-asociarse en trímeros.

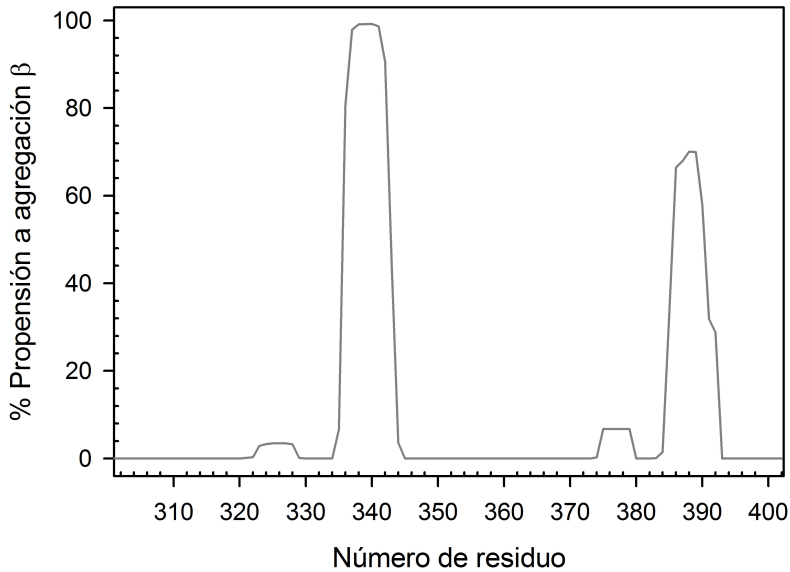


Figura 5.11. Gráfica obtenida a partir del programa TANGO al introducir la secuencia de aminoácidos del dominio PDZ3-PSD95 en un tampón hipotético a una concentración de 50 mM y pH 7.5. En la ordenada se representa el tanto por ciento de tendencia a agregación β , mientras que en la abscisa está el número de residuos desde el 302 al 402 (numeración de (Doyle et al., 1996)).

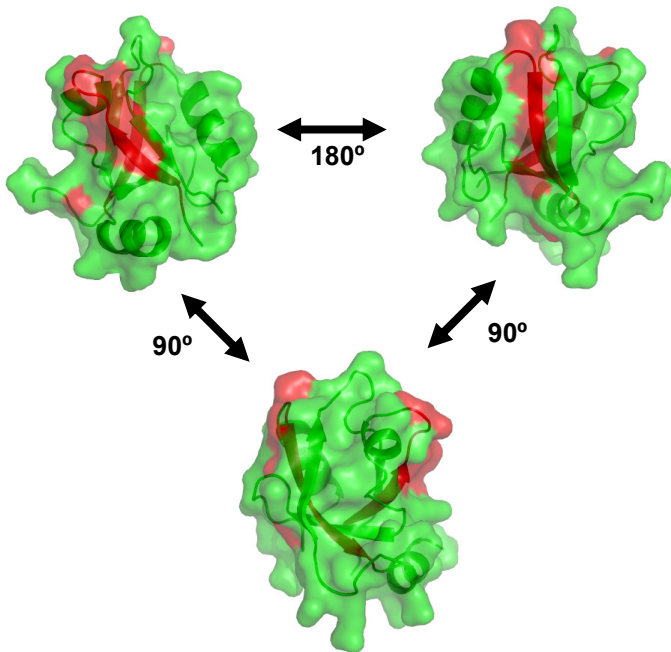


Figura 5.12. Representación de la estructura del dominio PDZ3-PSD95 en color verde con la región de los residuos con propensión a la agregación en color rojo. Las tres imágenes son vistas de la estructura del dominio giradas 90°.

Por otro lado, no se debe descartar la idea de que el proceso de oligomerización observado no corresponda a una única especie, sino que sea una mezcla de dímeros, trímeros, tetrámeros u otros oligómeros, aunque no lo muestren los experimentos de DLS. En ese caso, la fracción de aquellas especies de menor tamaño disminuye cuando baja la concentración de proteína, lo que además sería más probable desde un punto de vista del mecanismo. De hecho, ya ha sido descrita la coexistencia de diferentes estados de asociación para este tipo de especies de nucleación de fibras amiloides (Smith et al., 2006). La existencia de esos equilibrios adicionales podría ser otra posible explicación de las pequeñas desviaciones en los ajustes del DSC, y podría contribuir en cierta manera a la forma un poco aguda y con cierta asimetría que tiene la primera endoterma (Figura 5.4).

Para entender mejor la energética del estado intermedio, se deconvolucionaron las dos endotermas del DSC en sus distintas contribuciones al calor total (Figura 5.13). La evolución del calor durante el desplegamiento se da únicamente en la primera endoterma (Panel B de la Figura 5.13). Esta primera endoterma, además, es el resultado de dos procesos de signo contrario. Por un lado, la contribución positiva del proceso de desplegamiento de la proteína, y por otro, el proceso exotérmico debido al fenómeno de asociación del intermedio. En cuanto a la segunda endoterma, la contribución al calor es únicamente la debida a la disociación del oligómero, y no incluye ninguna otra contribución al calor, como por ejemplo, el completo desplegamiento del intermedio. De esta forma, parece que el estado I_n recuerda más bien a un estado parcialmente plegado en el que se han perdido la mayoría de las interacciones terciarias, y no a una conformación altamente estructurada y cooperativa.

El carácter no nativo de este estado parcialmente plegado se revela también con su incapacidad de reconocer el péptido KKETAV, ligando de alta afinidad del dominio PDZ3-PSD95. Tal y como se ve en la Figura 5.14, la variación en las trazas de DSC a una misma concentración de proteína y de tampón con distintas relaciones de péptido se debe únicamente al aumento de estabilidad que se debe a la formación del complejo con el estado N (Privalov & Dragan, 2007). No se ve afectada la segunda transición, que sería la propia de la disociación del estado I_n , revelando así la incapacidad que tiene el intermedio de reconocer el péptido. Además la presencia de éste tampoco influencia en nada la formación de las estructuras circulares o las fibras, lo que hemos comprobado haciendo incubaciones de los mismos tiempos con el péptido y no se ha alterado la secuencia de micrografías observada en la Figura 5.8.

En cualquier caso, los ajustes tan buenos de los datos al modelo (Figura 5.9) no son definitivos en cuanto al carácter no cooperativo del intermedio, por lo que no se puede descartar la existencia de regiones con una cierta organización espacial bien definida. Los oligómeros pueden posiblemente mantener ciertas estructuras locales o quizás contener nuevos elementos que no están presentes en el estado nativo. De hecho, los experimentos de SEC indican que a 60°C (Figura 5.10) se pueden resolver especies monoméricas dentro de la mezcla, lo cual apoya la idea de que posee cierta estabilidad termodinámica por sí mismo. Sin embargo, la posible estructura residual que poseen contribuiría relativamente poco al calor de desnaturalización del dominio PDZ3-PSD95.

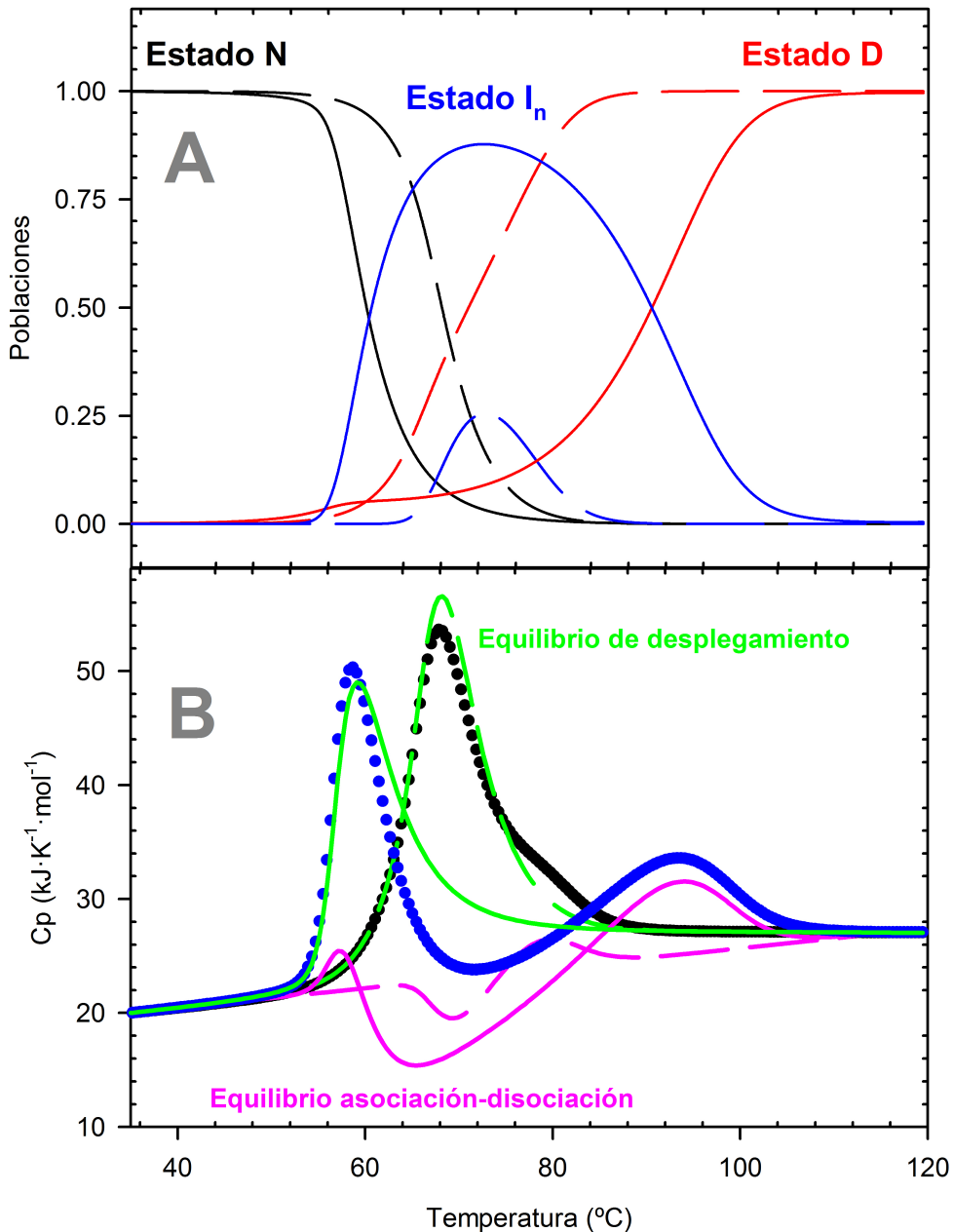


Figura 5.13. (A) Distribución de las poblaciones de los diferentes estados para las dos concentraciones de proteína usadas, en línea discontinua la de $0.4 \text{ mg}\cdot\text{mL}^{-1}$ y en línea continua la de $7.9 \text{ mg}\cdot\text{mL}^{-1}$. (B) Análisis de deconvolución de los diferentes equilibrios en el desplegamiento térmico del dominio PDZ3-PSD95. En color verde se representan las evoluciones térmicas del equilibrio de desplegamiento ($N \rightleftharpoons D$), mientras que en color rosa, el equilibrio asociación-disociación ($D \rightleftharpoons I_n$). Los círculos representan los datos experimentales de DSC. Tal y como se puede observar, el equilibrio de desplegamiento tiene forma simétrica mientras que el intermedio se asocia primero (mitad exotérmica) y se disocia después (mitad endotérmica) para completar el desplegamiento.

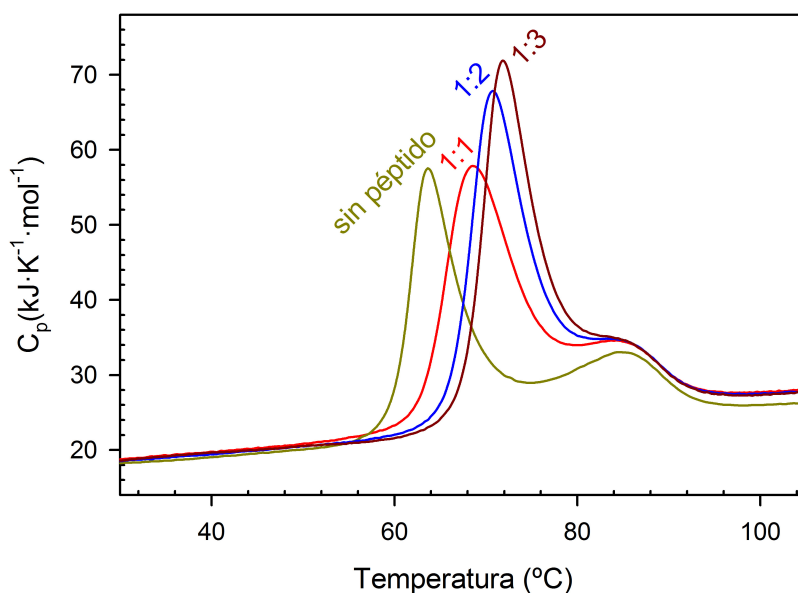


Figura 5.14. Curvas de desplegamiento térmico del PDZ3-PSD95 seguidas por DSC con distintas relaciones estequiométricas con un péptido de alta afinidad (KKETAV) a una misma concentración de proteína ($1.6 \text{ mg}\cdot\text{mL}^{-1}$) y a pH neutro. El péptido no es reconocido por el intermedio, ya que la segunda transición no varía.

5.1.4 Análisis espectroscópico y cinético de los estados conformacionales iniciales del mecanismo de fibrilación del dominio PDZ3-PSD95

Para poder conocer mejor la organización estructural del estado intermedio, se llevaron a cabo distintos experimentos con diferentes técnicas espectroscópicas. Por un lado, se registraron espectros de CD del dominio PDZ3-PSD95 a una concentración de $8 \text{ mg}\cdot\text{mL}^{-1}$ dializado en fosfato 50 mM pH 7.5 a dos temperaturas. Las temperaturas elegidas fueron 25°C y 70°C , aquellas en que hay una población máxima del estado nativo (N) y del intermedio (I_n), respectivamente (panel A de la Figura 5.13). El espectro correspondiente al intermedio, a 70°C , muestra un alto contenido en estructura secundaria en lámina β de acuerdo con el mínimo observado a 217 nm (Figura 5.15). Los espectros de CD no cambiaron después de varias horas de incubación a esa temperatura, por lo que esa conformación se mantiene a lo largo del tiempo. La ausencia de triptófanos, junto con el bajo contenido en residuos aromáticos de esta proteína, hizo que no se pudiese hacer un estudio de la estructura terciaria mediante CD en el UV-cercano.

Por otro lado tal y como se señala en la Sección 3.7, para detectar fibras amiloides es muy frecuente el uso del Congo Red (CR). Los ensayos de las fibras maduras obtenidas por la incubación durante un mes a 60°C del dominio PDZ3-PSD95 a $8 \text{ mg}\cdot\text{mL}^{-1}$ en fosfato 50 mM pH 7.5 con CR a una concentración de $20 \text{ }\mu\text{M}$ fueron positivos, dando lugar a un desplazamiento hacia longitudes de onda más cortas en el espectro del CR (Figura 5.16).

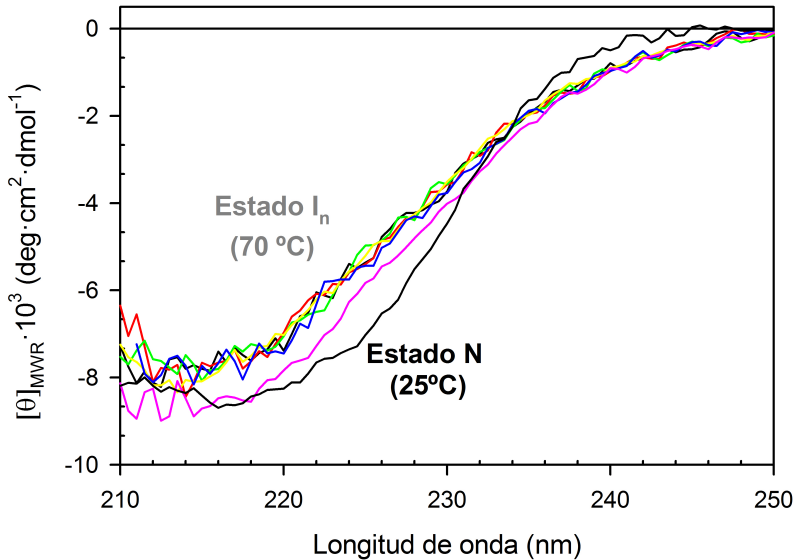


Figura 5.15. Espectros de CD del dominio PDZ3-PSD95 en tampón fosfato 50 mM pH 7.5 a 70°C en función del tiempo transcurrido después de haber alcanzado dicha temperatura. Se observa el mismo espectro prácticamente desde los 9 minutos hasta las 4 horas (líneas coloreadas) y hay un pequeño aumento de elipticidad después de las 4 horas, que es el espectro en color rosa. La concentración de proteína fue de $8 \text{ mg} \cdot \text{mL}^{-1}$, lo suficientemente alta como para asegurar que a 70°C el intermedio es la especie principal en la disolución (panel A de la Figura 5.13). El espectro en las mismas condiciones experimentales, pero a 25°C, está representado en color negro (Estado N).

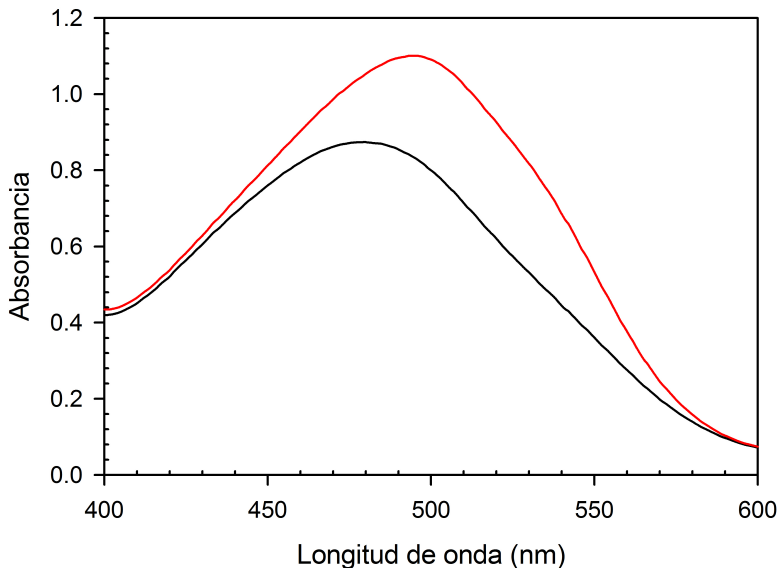


Figura 5.16. Espectros de absorbancia del Congo Red a una concentración de $20 \mu\text{M}$ en tampón fosfato 50 mM pH 7.5 en ausencia (espectro en color negro) y presencia de una muestra de PDZ3-PSD95 a una concentración de $8 \text{ mg} \cdot \text{mL}^{-1}$ incubada a 60°C durante una semana (espectro en color rojo). Se observa un aumento de absorbancia con un desplazamiento del máximo hacia longitudes de onda mayores, indicando la presencia de fibras amiloides (ver Sección 3.7).

Para conocer mejor la secuencia de eventos que se dan en la formación de las fibras, se llevaron a cabo experimentos en los que se siguió de forma continua la fluorescencia de ANS y ThT en función del tiempo. Estas sondas son bien conocidas por su capacidad de unirse a zonas hidrofóbicas de proteínas, así como a estructuras fibrilares de tipo β , respectivamente (ver Sección 3.8 para más detalle).

El dominio PDZ3-PSD95, a $8 \text{ mg}\cdot\text{mL}^{-1}$ en tampón fosfato 50 mM pH 7.5, junto con ThT a una concentración de $12.5 \text{ }\mu\text{M}$ o bien ANS a $20 \text{ }\mu\text{M}$, se calentó en el fluorímetro hasta 60°C , y una vez alcanzada esta temperatura, se registró la señal de fluorescencia en función del tiempo durante varias horas (Figuras 5.17 y 5.18). A continuación se registraron varios espectros después de enfriar la muestra hasta temperatura ambiente. De esta forma, se hicieron los experimentos paralelamente a los de DLS descritos en la Sección 5.1.2.

Cuando la temperatura alcanzó 50°C , hay un incremento de fluorescencia por parte de ambas sondas; sin embargo, se observó que nada más alcanzar 60°C y pasar al modo de medida en función del tiempo, las señales de fluorescencia de ambas sondas crecen muy rápidamente, alcanzando un máximo a los 100 minutos, en el caso de ThT y de 400 minutos en el de ANS. A continuación, la señal comienza a decaer. Esto último podría ser debido al aumento en la dispersión de luz como consecuencia del aumento de viscosidad de la disolución debido a la menor solubilidad de las estructuras supramacromoleculares que se están formando en estos tiempos (Hawe et al., 2008; Kad et al., 2003; Kardos et al., 2004). A continuación, cuando se enfrió hasta temperatura ambiente, se observa un efecto hipercrómico, posiblemente como consecuencia de la descomposición de esas estructuras más grandes en especies más solubles y de menor peso molecular (Figura 5.19 y 5.20).

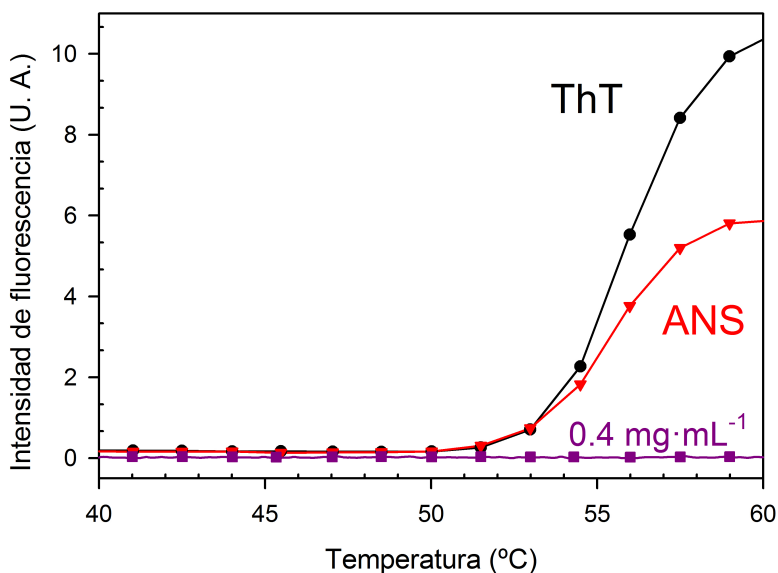


Figura 5.17. Barrido de temperatura de una muestra de PDZ3-PSD95 a $8 \text{ mg}\cdot\text{mL}^{-1}$ en presencia de $12.5 \text{ }\mu\text{M}$ de ThT (negro) o $20 \text{ }\mu\text{M}$ de ANS (rojo) en fosfato 50 mM pH 7.5 seguido por la emisión de fluorescencia de ambas sondas. En morado se muestra la muestra de PDZ3-PSD95 a $0.4 \text{ mg}\cdot\text{mL}^{-1}$ en la presencia de $12.5 \text{ }\mu\text{M}$ de ThT o $20 \text{ }\mu\text{M}$ de ANS a pH 7.5.

Asimismo, se realizaron los mismos experimentos a una concentración de proteína de $0.4 \text{ mg}\cdot\text{mL}^{-1}$ en vez de a $8 \text{ mg}\cdot\text{mL}^{-1}$ (Figuras 5.17 y 5.18). Se observó que no hay incremento de la fluorescencia de los sondas, confirmando así que no se puebla el intermedio lo suficientemente como para dar las estructuras macromoleculares (ver Sección 5.1.3, en concreto Panel A de la Figura 5.13).

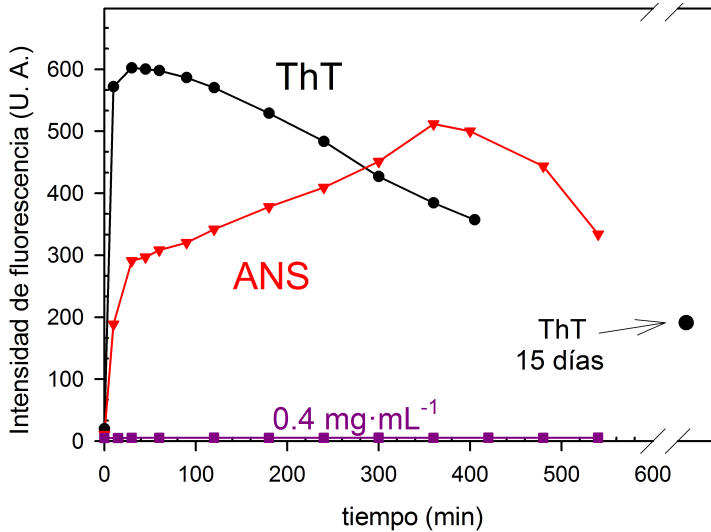


Figura 5.18. Emisión de fluorescencia en función del tiempo de incubación a 60°C de una muestra de PDZ3-PSD95 a $8 \text{ mg}\cdot\text{mL}^{-1}$ en presencia de $12.5 \mu\text{M}$ de ThT (negro) o $20 \mu\text{M}$ de ANS (rojo) en fosfato 50 mM pH 7.5. En morado se representa la muestra de PDZ3-PSD95 a $0.4 \text{ mg}\cdot\text{mL}^{-1}$ en la presencia de $12.5 \mu\text{M}$ de ThT o $20 \mu\text{M}$ de ANS a pH 7.5.

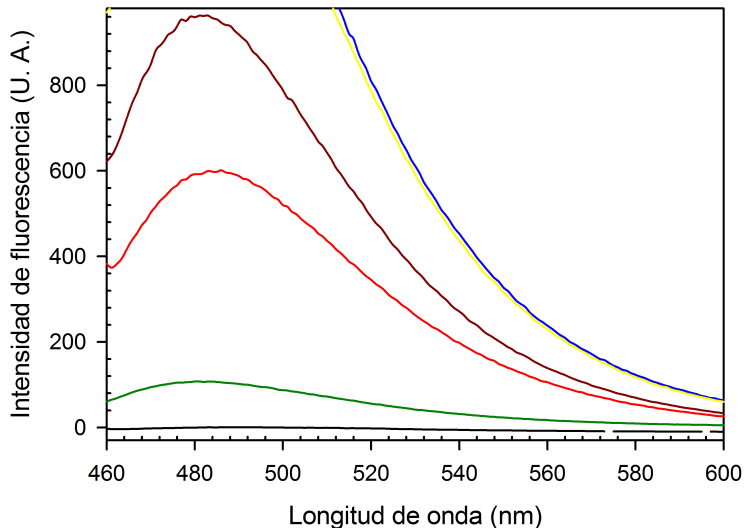


Figura 5.19. Espectros de emisión de ThT obtenidos a partir de una muestra del dominio PDZ3-PSD95 a $8 \text{ mg}\cdot\text{mL}^{-1}$ en fosfato 50 mM pH 7.5 en presencia de $12.5 \mu\text{M}$ de ThT en distintas condiciones de incubación. Se representa en color negro el espectro antes de calentar, mientras que en rojo después de haber incubado la muestra a 60°C durante una hora. A continuación la muestra se enfrió a 25°C y se registraron espectros un minuto después de haber alcanzado dicha temperatura (color azul), una hora después (color amarillo), un día después (color marrón) y una semana después (color verde).

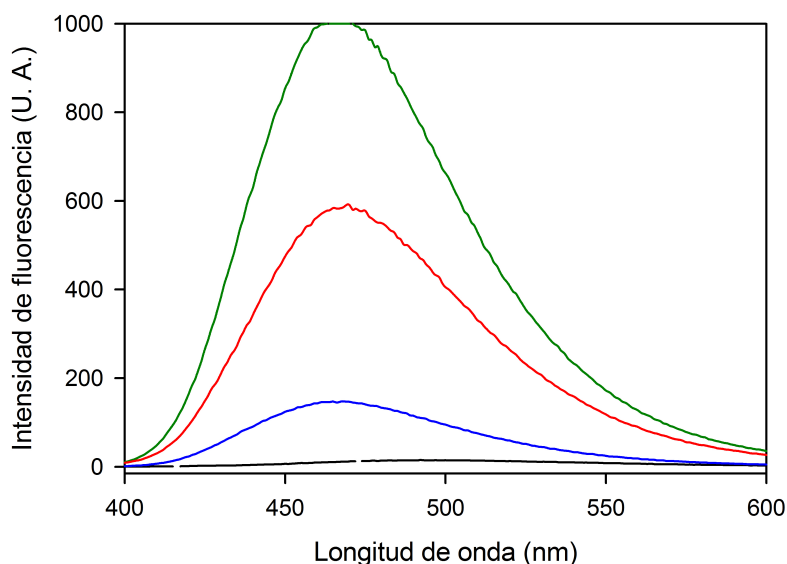


Figura 5.20. Espectros de emisión de ANS obtenidos a partir de una muestra del dominio PDZ3-PSD95 a $8 \text{ mg}\cdot\text{mL}^{-1}$ en fosfato 50 mM pH 7.5 en presencia de $20 \mu\text{M}$ de ANS en distintas condiciones de incubación. Se representa en color negro el espectro antes de calentar, mientras que en rojo después de haber incubado la muestra a 60°C durante 7 horas. A continuación la muestra se enfrió a 25°C y se registraron espectros 45 minutos después de haber alcanzado dicha temperatura (color verde) y un día después (color negro).

5.2 Discusión

Los parámetros termodinámicos que se muestran en la Tabla 5.3 indican que el dominio PDZ3-PSD95 es un dominio estable que tiene una energía libre de Gibbs de desplegamiento de unos $40 \text{ kJ}\cdot\text{mol}^{-1}$, valor que es comparable al de otras proteínas de tamaño similar, como por ejemplo la barnasa (Martinez et al., 1994). Sin embargo, este valor no se corresponde bien con los datos obtenidos para el dominio PDZ2-PTPBL o para el mismo PDZ3-PSD95, cuyos valores oscilan desde 14 hasta $30 \text{ kJ}\cdot\text{mol}^{-1}$ (Jemth & Gianni, 2007; Feng et al., 2005). Hay que tener en cuenta que estos resultados previos se obtuvieron a partir de un análisis con el modelo de dos estados de datos únicamente espectroscópicos.

El intermedio oligomérico de equilibrio descrito en esta Memoria se puede considerar como una nueva especie no detectada previamente en otros estudios de plegamiento de dominios PDZ. Así, se ha estudiado el mecanismo de plegamiento de varios dominios PDZ – incluyendo el del dominio PDZ3-PSD95 – mediante técnicas cinéticas. Los resultados obtenidos permitían describir dos estados de transición separados por un intermedio de alta energía a través del cual el mecanismo debe ir obligatoriamente (intermedio *on-pathway*). Estos dos estados de transición comunes a todos los dominios PDZ estudiados hasta la fecha, están situados aproximadamente en las mismas posiciones a lo largo de la coordenada de reacción, indicando así un mecanismo conservado de

plegamiento para estos dominios. Sin embargo, a pesar del carácter obligatorio del intermedio, no se acumula nunca en condiciones de equilibrio (Jemth & Gianni, 2007; Chi et al., 2007; Calosci et al., 2008).

En un estudio de resonancia magnética nuclear de intercambio hidrógeno/deuterio de una construcción más larga del dominio PDZ3-PSD95 se describe también un intermedio de equilibrio después de la etapa limitante del plegamiento (Feng et al., 2005). La estructura de ese intermedio se asemeja fundamentalmente a la forma nativa del PDZ3, con la excepción de los primeros residuos del extremo N-terminal, la tercera hélice α y una horquilla β que se organiza en el extremo C-terminal. Sin embargo, esta construcción del dominio PDZ3-PSD95 con la que se ha trabajado en esta Memoria contiene los residuos 302 al 402 de la proteína PSD95 y no contiene los residuos que forman la horquilla β , que, por otra parte, provienen del vector de expresión usado en los trabajos de Feng y colaboradores y otros análisis estructurales previos (Feng et al., 2005; Doyle et al., 1996). De esta forma, la evidencia adicional de la baja cooperatividad del intermedio, su contenido esencialmente en lámina β y la naturaleza oligomérica del intermedio detectado por DSC usando esta construcción sin la horquilla β , hace por tanto que ambos intermedios de equilibrio sean sustancialmente diferentes.

Solamente se ha encontrado un ejemplo previo en la bibliografía referente a DSC de una proteína similar en tamaño al dominio PDZ3-PSD95, la proteína de 14 kDa CheY, que también muestra una topología α - β y que fue analizada según el mismo modelo de equilibrio (Filimonov et al., 1993). Los intermedios de CheY y PDZ3-PSD95, ambos detectados por DSC, comparten algunas características, como la tendencia a asociarse y su baja cooperatividad en el desplegamiento térmico. Sin embargo, la formación de estructuras supramacromoleculares de la proteína CheY no ha sido descrita hasta ahora. Estos dos casos corroboran la idea de que proteínas monoméricas de tamaño medio (con pesos moleculares mayores a unos 10 kDa), se despliegan siguiendo un mecanismo más complejo que aquellas que son más pequeñas, presentando al menos un intermedio de equilibrio.

5.2.1 Aspectos termodinámicos y conformacionales de la formación de estructuras supramacromoleculares durante el desplegamiento del dominio PDZ3-PSD95

Hasta ahora sólo se ha descrito la formación de fibras en dominios PDZ para el caso del dominio PDZ2-PTPBL como resultado de la combinación de agitación mecánica y disolventes con una alta fuerza iónica a pH neutro (Sicorello et al., 2009). Las fibras, aunque se forman en condiciones nativas, poseen los espectros típicos de CD e infrarrojo de estructuras β puras, tal y como se da en el caso del intermedio del dominio PDZ3-PSD95.

Los estudios descritos en esta Memoria con este dominio muestran algunas características interesantes que se refieren a la naturaleza físico-química de los procesos que se dan en la formación de fibras. En primer lugar, la fibrilación puede ocurrir a través

de un oligómero que posee estructura β , lo que sugiere un modelo con varias etapas para la formación de fibras amiloides, tal y como ha sido descrito en otros ejemplos (Kad et al., 2003; Kodali & Wetzel, 2007; Jiang et al., 2004). Esta fase de nucleación, similar a la nucleación que se da en el crecimiento cristalino y que generalmente provoca la existencia de una fase de retardo en el crecimiento (Figura 5.21 tomada de (Nilsson, 2004)), se completa antes de alcanzar la temperatura donde se forman los amiloides. Esto hace que haya un comportamiento no cooperativo (no sigmoidal, por tanto) en el proceso de fibrilación del dominio PDZ3-PSD95 (Figura 5.18), que ocurre mediante la adición de los oligómeros a las estructuras supramacromoleculares en un proceso sin barreras energéticas, conocido como *downhill*, en inglés (Juárez et al., 2009; Carrotta et al., 2005; Kodali & Wetzel, 2007).

Por otro lado, las fibras se disuelven bajando a temperatura ambiente después de una dilución 1:1 de la muestra con el tampón y se vuelven a formar cuando se calienta de nuevo a 60°C. Esta evidencia implica la naturaleza reversible de los procesos involucrados en la formación de fibras, a pesar de los tiempos de relajación tan largos que tienen, ya que se pueden deshacer y volverse a formar de forma espontánea. Hay descritos en bibliografía únicamente un par de ejemplos en los que ocurre lo mismo, la acilfosfatasa (Calamai et al., 2005) y la apolipoproteína C-II (Binger et al., 2008). De esta manera, las fibras que se forman a partir del dominio PDZ3-PSD95 ofrecen una oportunidad interesante y poco común para estudiar cómo se disgregan estas estructuras, lo que podría ayudar en la detección de nuevas especies claves para explicar la naturaleza molecular de la formación de fibras y el posterior desarrollo de tratamientos para enfermedades asociadas a la agregación de proteínas.

Desde el punto de vista energético, este estudio muestra un ejemplo inusual de la termodinámica de formación de fibras. Como ya se ha señalado, el crecimiento de los oligómeros no parece estar acompañado por ningún cambio conformacional en el estado intermedio, porque la segunda transición que se observa en las trazas de DSC no incluye más efectos caloríficos que los debidos a la disociación de los oligómeros (Panel B de la Figura 5.13). De hecho, los espectros de CD registrados en el UV-lejano no muestran cambios aparentes en la forma (Figura 5.15). Esto ya se ha visto en las distintas caracterizaciones calorimétricas que se han hecho de formación de fibras amiloides, con la β_2 -microglobulina (Kardos et al., 2004) y con dominios SH3 (Morel et al., 2006), que además comparten otras características parecidas con el caso del dominio PDZ3-PSD95. Así, el cambio de entalpía de la disociación de los oligómeros de este dominio ($-130 \pm 20 \text{ kJ}\cdot\text{mol}^{-1}$) es muy similar con aquellos de las fibras de la β_2 -microglobulina ($-125 \text{ kJ}\cdot\text{mol}^{-1}$) y del SH3 ($-100 \text{ kJ}\cdot\text{mol}^{-1}$). Estos valores están también en el rango del debido a la disociación del intermedio dimérico de equilibrio presente en la proteína CheY, que posee un comportamiento experimental en DSC similar al dominio PDZ3-PSD95, tal y como ya se ha explicado. El pequeño valor de la entalpía específica estimado para la asociación del intermedio del dominio PDZ3-PSD95 ($11.8 \text{ J}\cdot\text{g}^{-1}$) junto con los valores similares de los otros tres ejemplos indica la relativamente baja densidad de las interacciones entre las cadenas laterales dentro de las fibras. De esta forma, parece que la fuerza conductora de la

formación de fibras son los factores entrópicos y muestran un mecanismo energético similar para el apilamiento de las láminas β , tal y como ha sido ya sugerido por experimentos cinéticos (Kardos et al., 2004; Morel et al., 2006).

Además, la energía libre de Gibbs a 70°C, que es la temperatura en la que las especies oligoméricas se pueblan según se puede ver en la Tabla 5.3, es de $25 \pm 5 \text{ kJ}\cdot\text{mol}^{-1}$, mientras que es prácticamente nula a 25°C, donde la estabilidad del estado nativo ha sido estimada en $40 \pm 6 \text{ kJ}\cdot\text{mol}^{-1}$. Esta estabilidad del estado nativo, mayor que la de las fibras, puede explicar la reversibilidad observada, ya que el estado nativo es más estable que las estructuras asociadas a temperatura ambiente. Esto está en acuerdo con el efecto contrario que se ve en el caso de la irreversibilidad de las fibras de β_2 -microglobulina, en las que se observa que hay un valor de estabilidad mayor para las fibras que para el estado nativo a cualquier temperatura (Kardos et al., 2004). La estabilidad de las fibras del dominio PDZ3-PSD95 es considerablemente menor que las formadas a partir de β_2 -microglobulina ($44 \text{ kJ}\cdot\text{mol}^{-1}$), lo cual podría deberse a diferencias estructurales de origen entrópico y podría justificar también su menor grosor y longitud.

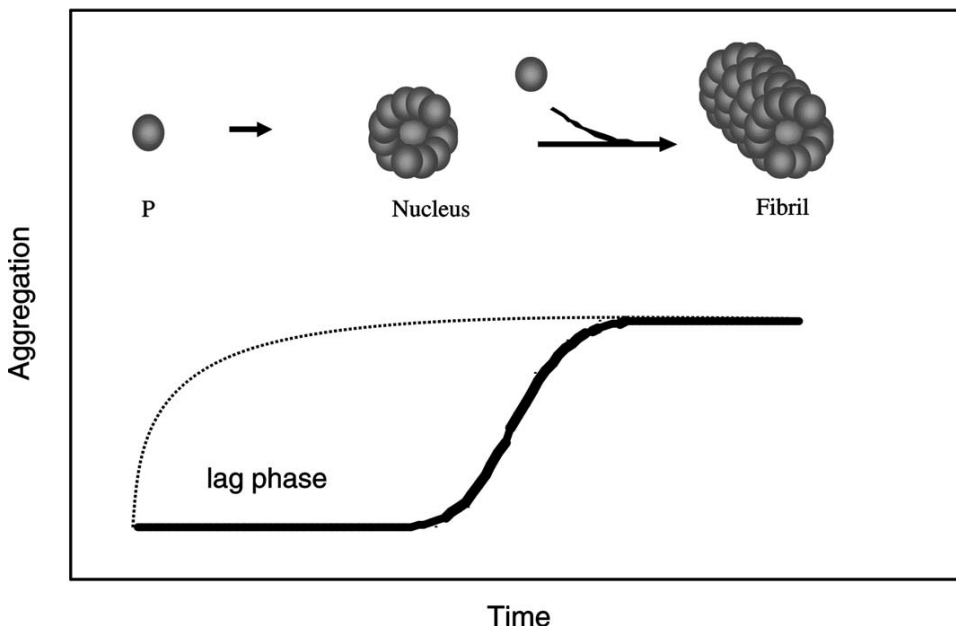


Figura 5.21. La formación de fibras a partir de una proteína (P) puede empezar con una fase de retardo (lag phase) en la que no se observa agregación. Durante este tiempo, se da el proceso entrópicamente desfavorable de asociación inicial. Una vez que el proceso de agregación comienza y que un núcleo con el tamaño crítico se forma, la agregación se conduce rápidamente hacia la formación de las fibras (línea continua). También puede suceder que no haya esta fase inicial de retardo y el proceso de agregación se dé desde el comienzo (línea de puntos) puesto que ya hay presente un núcleo, tal y como ocurre en el caso del dominio PDZ3-PSD95 (Figura tomada de (Nilsson, 2004)).

5.2.2 El esquema de desplegamiento del dominio PDZ3-PSD95 como posible paradigma para el estudio biofísico de interconversión entre estructuras macromoleculares

La capacidad de formar fibras se considera hoy en día como una propiedad general de cualquier cadena polipeptídica. Hay menos evidencia de que cualquier proteína pueda formar estructuras anulares y esféricas, aunque en un trabajo de Quist y colaboradores (Quist et al., 2005) se ha visto que una serie de proteínas y péptidos amiloidogénicos que no están relacionados entre sí forman estructuras anulares similares cuando están en presencia de membranas lipídicas y que además son capaces de actuar como transportadores de iones, es decir, como canales iónicos. Estas evidencias indican que es bastante común la formación de estas estructuras. El diámetro de los canales es de unos 8 a 12 nm, un tamaño que es parecido a las estructuras de 12 nm que han sido descritas en el apartado 5.1.2. Según se puede apreciar en las micrografías donde aparecen dichas estructuras (aquellas que están señaladas con “3 días”, “8 días” y “15 días (A)” de la Figura 5.8), parece que hay una cierta homogeneidad en tamaño, lo que sugiere que la organización de estas estructuras está bien definida (Quist et al., 2005). A su vez, también se han descrito la formación de agregados esféricos de tamaños que van desde los 5 nm hasta los 200 nm, estableciéndose como una propiedad general de las cadenas polipeptídicas (Krebs et al., 2007) y formando poros en su interacción con membranas lipídicas (Walsh et al., 2009).

La relación entre especies anulares y esféricas con las fibras amiloides no ha sido establecida hasta ahora. Debido a limitaciones experimentales no ha podido establecerse si hay un mecanismo común para la organización de estas estructuras supramacromoleculares o si son productos de mecanismos alternativos de plegamiento anómalo de las proteínas (Jang et al., 2008). En la mayoría de los ejemplos descritos en literatura se muestra una clara tendencia a la formación de especies oligoméricas pequeñas (generalmente desde dímeros hasta hexámeros) en la disolución, aunque la formación de las estructuras anulares ocurre únicamente cuando interactúan con membranas (Quist et al., 2005). Las estructuras esféricas se suelen formar en condiciones cercanas al punto isoeléctrico (Krebs et al., 2007) y/o al añadir disolventes orgánicos (Walsh et al., 2009).

Debido a esto es bastante interesante que el dominio PDZ3-PSD95 forme esas estructuras globulares sin la presencia de esas membranas, además de que coexisten con las fibras, en una disolución a pH fisiológico y sin presencia de ningún disolvente orgánico. Este fenómeno, junto con que estas estructuras pueden permanecer en disolución hasta dos semanas, apoya la idea de que la formación de estas estructuras sea un paso obligatorio, dentro del mecanismo de formación de las fibras, tal y como ha sido propuesto para otros casos, como la albúmina de suero humano (Juárez et al., 2009), o las proteínas Sup35 (Serio et al., 2000) y Ure2p (Jiang et al., 2004). De hecho, como ya se ha comentado, las fibras crecen a partir de las estructuras globulares (micrografía señalada con “3 días” de la Figura 5.8), que desaparecen de la disolución (micrografía señalada con

“1 mes” de la Figura 5.8). Sin embargo, la aparición de protofibras también en las etapas iniciales (micrografía señalada con “5 días” de la Figura 5.8) así como la única aparición de protofibras y algunas fibras finas cuando se incubó la muestra a 37°C (micrografía señalada con “1 mes (37°C)” de la Figura 5.8), no nos permite establecer una conclusión definitiva sobre el mecanismo de formación de estas estructuras.

Los experimentos de DLS y ANS/ThT dan alguna información adicional sobre este tema. La ausencia de fase de retardo (*lag phase*) y la coincidencia de los tiempos registrados con las distintas pruebas espectroscópicas, implica que una cierta cantidad de estructuras oligoméricas que muestran al disolvente regiones hidrofóbicas y estructuras de lámina β se forman después de un evento de nucleación común. Esto ha sido también observado con la proteína sinucleína- α (Hoyer et al., 2002), por lo que se podría concluir que si existe otro mecanismo, el punto de ramificación estaría después de la nucleación, es decir, después de la formación de las especies I_n . De esta forma, se podría resumir el comportamiento experimental de la formación de fibras del dominio PDZ3-PSD95 en tampón fosfato 50 mM a pH 7.5 según el esquema reflejado en la Figura 5.22.

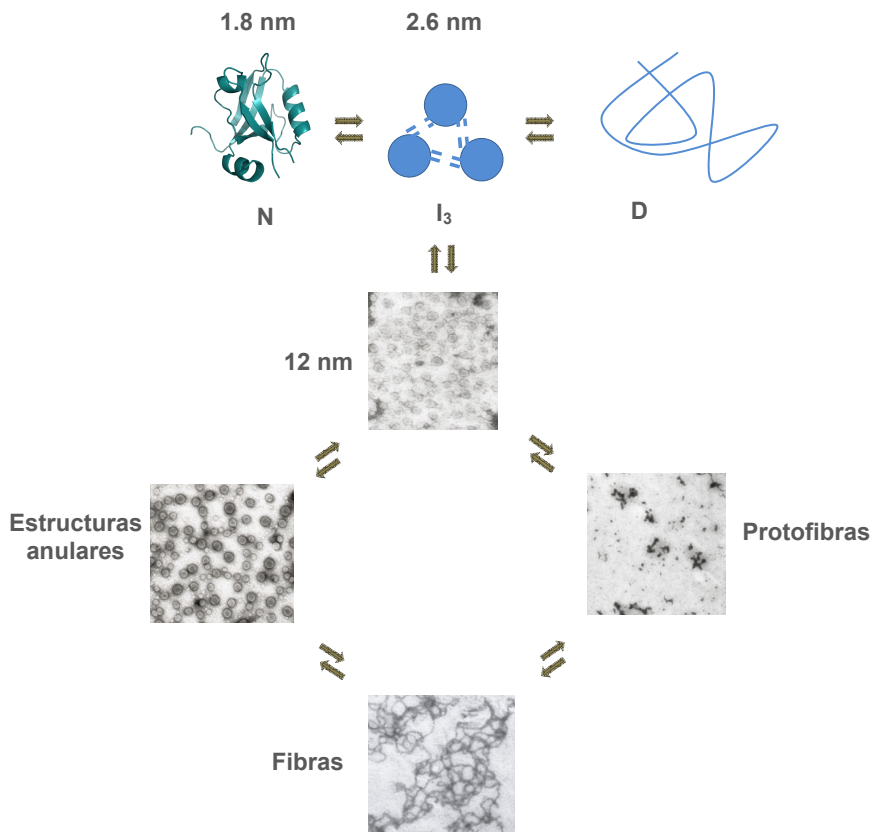


Figura 5.22. Esquema en que se representan las distintas etapas por las que pasa el dominio PDZ3-PSD95 cuando forma estructuras supramacromoleculares en fosfato 50 mM pH 7.5.

En el que las dobles flechas indican no tanto una posible situación de equilibrio, sino más bien la naturaleza reversible de los procesos que se están dando (al menos aquellos que se refieren a los procesos que ocurren en la fibrilación). Este comportamiento del dominio PDZ3-PSD95 abre la posibilidad de estudiar más en profundidad las distintas estructuras supramacromoleculares, teniendo en cuenta además que exponen al disolvente una superficie hidrofóbica relativamente alta, según indican los experimentos con ThT y ANS, así como que el estado I_n no es propiamente nativo, tal y como se ha visto por DSC (Figura 5.14) y que no es capaz de reconocer ligandos de alta afinidad (como el KKETAV, que no altera la secuencia de eventos descritos en la Figura 5.8).

Los estudios más recientes indican que las fibras amiloides podrían ser simplemente un mecanismo de almacenamiento y que las estructuras globulares serían suficientes para inducir la degradación de la célula mediante la disrupción de la homeostasis debida al intercambio iónico a través del canal formado por estas estructuras en la membrana. Así, el bloqueo selectivo de estos canales iónicos parece una estrategia prometedora para el tratamiento de enfermedades derivadas de la formación de amiloides. De esta forma, el mecanismo de desplegamiento del dominio PDZ3-PSD95 es muy informativo para obtener datos experimentales sobre el mecanismo de la formación de amiloides y su toxicidad. Además, el carácter reversible de los procesos descritos, que son en condiciones prácticamente fisiológicas, junto con la formación de estructuras anulares o esféricas sin el contacto con membranas, hace especialmente interesante este sistema para estudios posteriores de procesos de agregación acoplados al plegamiento.

CAPÍTULO 6

ESTUDIO DE LOS DETERMINANTES
ELECTROSTÁTICOS DEL EQUILIBRIO
CONFORMACIONAL DEL TERCER DOMINIO
PDZ DE LA PROTEÍNA PSD95

6.1 Resultados

El desplegamiento térmico del dominio PDZ3-PSD95 muestra un intermedio de equilibrio que presenta oligomerización a pH neutro y que actúa como precursor de estructuras supramacromoleculares tales como anillos, protofibras o fibras amiloides. Como además los procesos que se dan tienen un alto carácter reversible, hacen de esta proteína un modelo de estudio de la formación de estas estructuras. Para conocer mejor los aspectos moleculares de estos procesos y ver cómo se ven afectados por las distintas condiciones de la disolución, se realizaron más estudios de plegamiento a valores ácidos de pH.

6.1.1 Estudio del desplegamiento térmico del dominio PDZ3-PSD95 a valores de pH ácido

Inicialmente se realizó un barrido de temperatura seguido por CD en tampón glicina 50 mM pH 3 con la misma concentración de proteína con la que se realizó el experimento a pH neutro ($0.3 \text{ mg}\cdot\text{mL}^{-1}$). El resultado fue observar de nuevo una curva sigmoide que se analizó con el modelo de dos estados (Figura 6.1, Tabla 6.1). Se vio que a pH 7.5 el uso de esta técnica espectroscópica por sí sola no dio resultados concluyentes (ver Sección 5.1.1), por lo que se realizó de nuevo un estudio de DSC para comprobar ese comportamiento. Se dializó el dominio PDZ3-PSD95 en glicina 50 mM pH 3 teniendo una concentración final de proteína de $1.5 \text{ mg}\cdot\text{mL}^{-1}$ y se realizó un barrido de temperatura desde 5°C hasta 110°C en el calorímetro a una velocidad de barrido de $1.5 \text{ K}\cdot\text{min}^{-1}$. En este caso se observó que la traza experimental se corresponde con el modelo de dos estados, puesto que la endoterma se ajusta perfectamente al modelo matemático correspondiente (Figura 6.2), el cual está descrito con detalle en la Sección 3.4.2.1 Además, el grado de reversibilidad que presenta es de más del 90%. Sin embargo, cuando se realizó el mismo experimento pero aumentando la concentración de proteína hasta $5.1 \text{ mg}\cdot\text{mL}^{-1}$ la curva no es ya tan simétrica, por lo que no se puede afirmar que el comportamiento experimental se explique mediante un modelo de dos estados, puesto que es bien conocido que este modelo es independiente de la concentración (Viguera et al., 1994; Martínez et al., 1994). El análisis matemático de esta traza realizada con una mayor concentración de proteína confirma este hecho por la menor calidad del ajuste (Figura 6.2).

Al ver que a pH 3 cambia el comportamiento, el siguiente paso fue hacer un estudio de desplegamiento térmico seguido por DSC a más valores de pH ácido. Se hicieron los experimentos a una misma concentración de tampón (50 mM), se usó glicina para pH 2, 2.5 y 3.5, mientras que para pH 4 se usó acetato. La concentración de PDZ3-PSD95 fue la misma para todas las trazas, unos $1.5 \text{ mg}\cdot\text{mL}^{-1}$. El resultado es que hay una variación clara en la forma de las trazas calorimétricas en función del pH (Figura 6.3). Se puede observar que a pH igual o mayor que 3.5 aparecen dos transiciones. De hecho, a pH 4 aparecen dos transiciones claramente diferenciadas, como ocurría a pH neutro (ver Sección 5.1.1), aunque a pH 3.5 ya la segunda transición aparece muy solapada a la primera, recordando

a la forma de la traza de $0.4 \text{ mg}\cdot\text{mL}^{-1}$ a pH neutro (Figura 5.4). Tal y como ya se ha explicado en esa Sección, ese comportamiento se debe a que se puebla muy poco el estado intermedio.

Por tanto, la variación que se observa en el comportamiento a distintos pH ácidos indica que el dominio PDZ3-PSD95 sigue un desplegamiento térmico en el que hay un intermedio que se desestabiliza al bajar el pH de la disolución. Ese estado intermedio se puebla en cantidades apreciables a valores de pH mayores de 3.5. A valores de pH menores de este valor no se detecta ese estado intermedio; sin embargo, el hecho de que haya una dependencia de la concentración de proteína indica que en el mecanismo de desplegamiento térmico sigue estando presente dicho intermedio, aunque no se puebla a valores de pH menores o iguales a 3 y baja concentración de proteína, debido a una desestabilización energética.

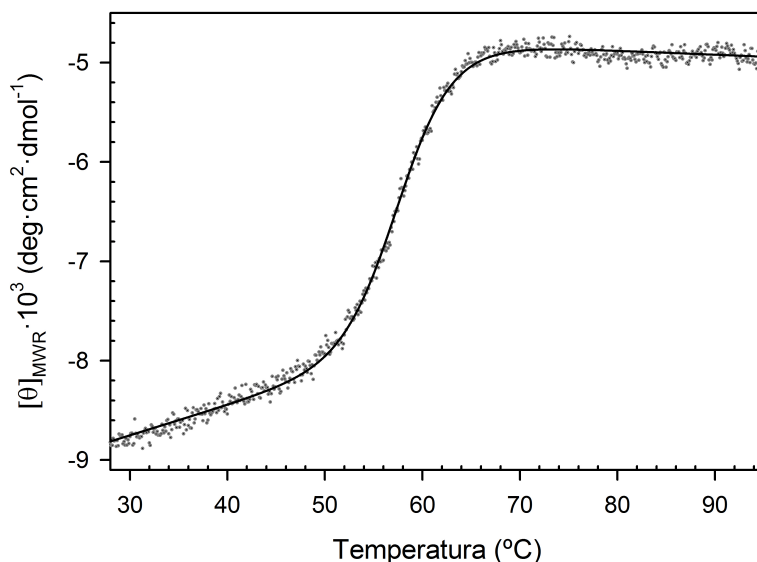


Figura 6.1. Curva de desplegamiento térmico seguido por CD a 218 nm del dominio PDZ3-PSD95 a una concentración de $0.3 \text{ mg}\cdot\text{mL}^{-1}$ en glicina 50 mM pH 3. Los puntos representan los datos experimentales, mientras que la línea es el ajuste al modelo de dos estados.

Tabla 6.1. Parámetros termodinámicos del desplegamiento térmico del dominio PDZ3-PSD95 en tampón glicina 50 mM pH 3 obtenidos del análisis de experimentos de CD con el modelo de dos estados.

T_{N-D} (°C)	$\Delta H_{N-D}(T_{N-D})$ (kJ·mol ⁻¹)	$\Delta G_{N-D}(298)$ (kJ·mol ⁻¹)	R^2
57.8 ± 0.1	323 ± 12	32 ± 10	0.998

Los intervalos de error se han considerado como tres veces los errores del ajuste (correspondiente a un intervalo de confianza del 99%).

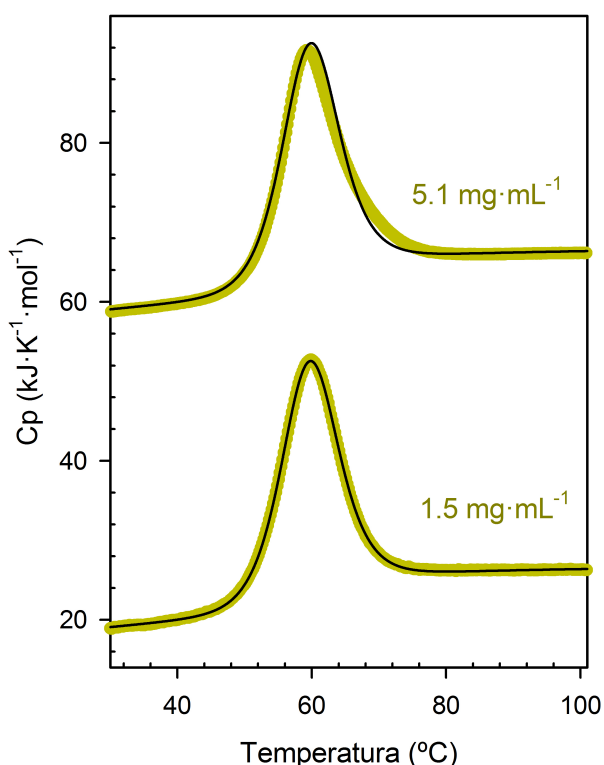


Figura 6.2. Curvas de desplegamiento térmico del PDZ3-PSD95 seguidas por DSC en función de la concentración de proteína a pH 3 en glicina 50 mM. Los puntos experimentales son los círculos amarillos, mientras que el ajuste al modelo $N=D$ se representa mediante líneas negras.

6.1.2 Evaluación de las contribuciones electrostáticas de los residuos ácidos a los parámetros termodinámicos de desplegamiento

El comportamiento experimental observado en las trazas calorimétricas ha de deberse al equilibrio de protonación de los aminoácidos glutamato o aspartato, puesto que sus valores de pK_a son los únicos que están en el intervalo de pH 3-4 (Weast, 1979). Para evaluar la contribución energética del balance electrostático de estos residuos, se han representado las entalpías de desplegamiento como función de sus respectivas temperaturas de transición obtenidas a diferentes condiciones de pH (Figura 6.4). Los valores obtenidos a partir del análisis con el modelo de dos estados ($pH \leq 3$) se correlacionan bien según una dependencia lineal. De hecho, se puede observar una buena coincidencia con el hipotético valor de $54 \text{ J}\cdot\text{g}^{-1}$ a 110°C , que representa un valor convergente de las funciones de entalpía de proteínas globulares pequeñas, tal y como está postulado por varios autores (Martinez et al., 1994; Fu & Freire, 1992; Baldwin & Muller, 1992; Ragone & Colonna, 1994). La pendiente de la regresión da un valor para el cambio de capacidad calorífica de desplegamiento de $5.6 \text{ kJ}\cdot\text{K}^{-1}\cdot\text{mol}^{-1}$, que es comparable a los valores obtenidos por otras proteínas de tamaños similares (Martinez et al., 1994).

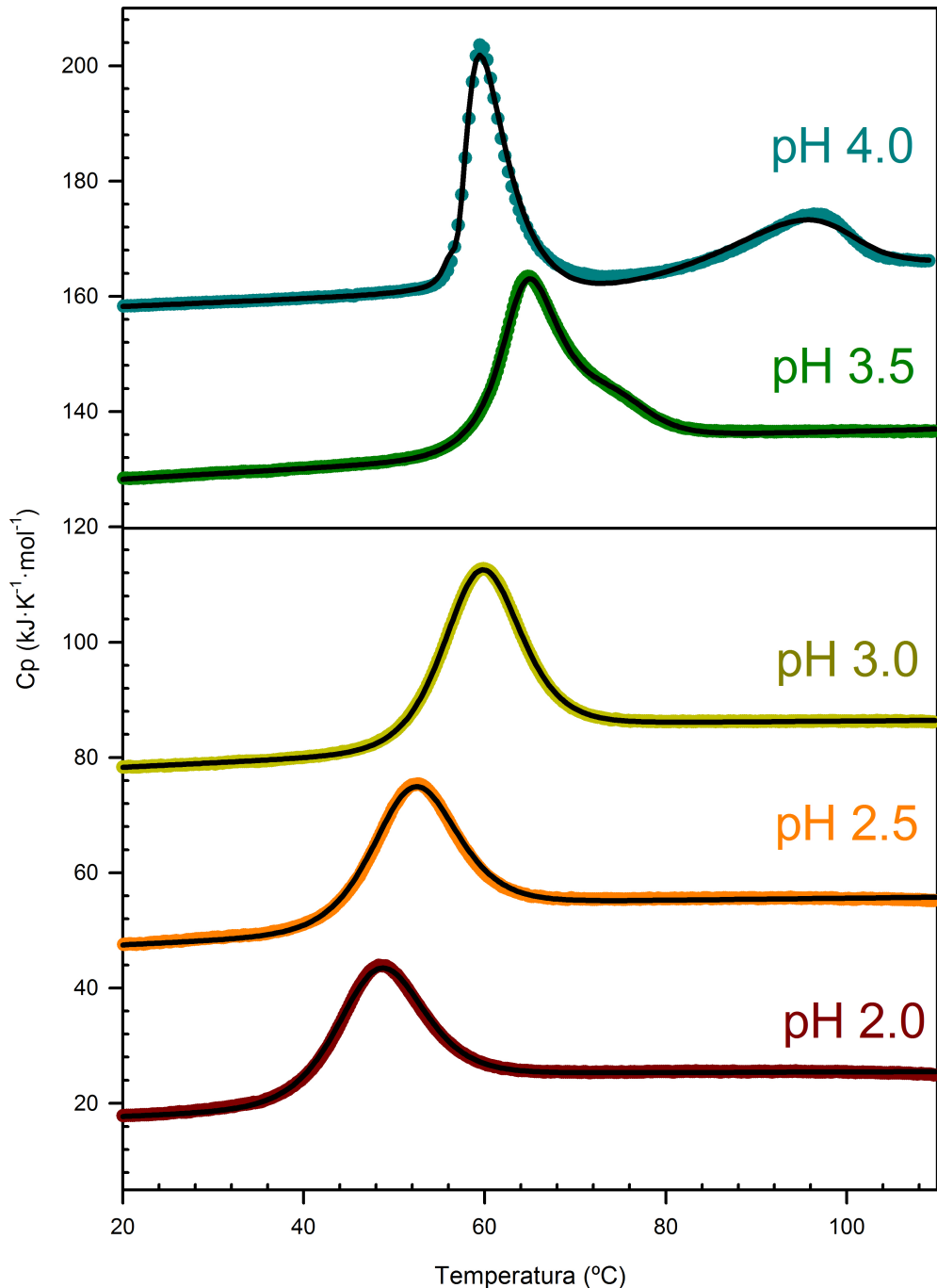


Figura 6.3. Curvas de desplegamiento térmico del dominio PDZ3-PSD95 seguidas por DSC en función del pH a una concentración de proteína de $1.5 \text{ mg}\cdot\text{mL}^{-1}$ y con una concentración de tampón de 50 mM . En puntos coloreados están representados los datos experimentales, mientras que las líneas de color negro son los análisis matemáticos, realizados con el modelo $3N \rightleftharpoons I_3 \rightleftharpoons 3D$ para valores de pH iguales o mayores de 3.5 y con el modelo $N \rightleftharpoons D$ para valores menores de este pH.

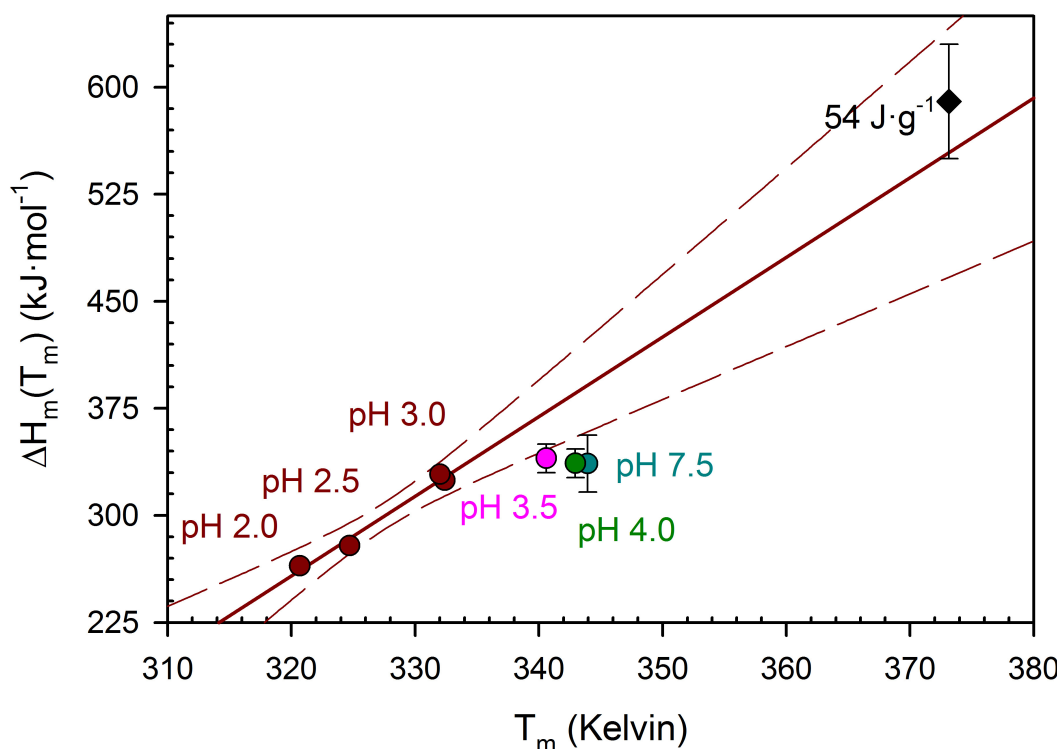


Figura 6.4. Análisis de la contribución electrostática de los residuos glutamatos y aspartatos a la entalpía de desplegamiento del dominio PDZ3-PSD95. La correlación lineal entre las entalpías obtenidas a partir del análisis con el modelo de dos estados (círculos en color marrón) es la línea continua en color marrón. Las líneas discontinuas representan los intervalos de confianza al 95% de dicha regresión lineal. Los círculos con sus barras de error en colores rosa (pH 3.5), verde (pH 4) y azul (pH 7.5) son los valores de entalpía para los análisis realizados con el modelo de tres estados. El diamante en color negro representa el valor de convergencia de $54 \text{ J}\cdot\text{g}^{-1}$ a 110°C para proteínas globulares pequeñas.

Los valores de entalpía obtenidos mediante el análisis con el modelo de tres estados descrito en la Sección 3.4.2.2 ($\text{pH} > 3$) y que están recogidos en la Tabla 6.2, no entran dentro de esa dependencia lineal (Figura 6.4). Por lo tanto, considerando que a esos valores de pH el hecho más relevante energéticamente está relacionado con la protonación de restos glutamato y/o aspartato, se puede concluir que los casi $40 \text{ kJ}\cdot\text{mol}^{-1}$ de diferencia deben estar relacionados con la contribución al calor de las interacciones electrostáticas desarrolladas después de la pérdida del protón de tales residuos por encima de pH 3. Esta estimación sería debida únicamente a la contribución neta de este balance electrostático, ya que los espectros de CD no muestran diferencias apreciables entre pH 3 y pH 7.5 (Figura 6.5), lo que indica una disposición estructural similar para el dominio PDZ3-PSD95 en las condiciones estudiadas de disolvente.

Tabla 6.2. Parámetros termodinámicos del desplegamiento térmico del dominio PDZ3-PSD95 en tampón acetato, para pH 4, y glicina para el resto de valores de pH, a una concentración de 50 mM obtenidos del análisis de experimentos de DSC según el modelo $3N \rightleftharpoons I_3 \rightleftharpoons 3D$ para pH mayores o iguales a 3.5 y según el modelo $N \rightleftharpoons D$ para el resto de valores de pH.

pH	T_{N-D} (°C)	$\Delta H_{N-D}(T_{N-D})$ (kJ·mol ⁻¹)	T_{D-In} (°C)	$\Delta H_{D-In}(T_{D-In})$ (kJ·mol ⁻¹)	$\Delta G_{N-D}(298K)$ (kJ·mol ⁻¹)	R ²
3N ⇌ I₃ ⇌ 3D						
7.5	70.4 ± 0.5	335 ± 20	79.2 ± 1.2	-130 ± 10	39 ± 6	0.996
4	70.0 ± 0.3	335 ± 10	89.3 ± 1.4	-145 ± 30	33 ± 8	0.983
3.5	67.4 ± 0.2	340 ± 10	77.0 ± 1.7	-110 ± 20	37 ± 7	0.999
N ⇌ D						
3	59.0 ± 0.1	325 ± 5	—	—	26 ± 2	0.999
2.5	51.6 ± 0.1	280 ± 5	—	—	17 ± 2	0.998
2	47.5 ± 0.1	265 ± 5	—	—	14 ± 2	0.998

Los intervalos de error se han considerado como tres veces los errores del ajuste (correspondiente a un intervalo de confianza del 99%). La concentración de proteína fue de 1.5 mg·mL⁻¹ en todos los experimentos, excepto los datos de pH 3 marcados con un asterisco, cuya concentración fue de 5.1 mg·mL⁻¹ (ver texto). Se han incluido los datos de pH 7.5 para la comparación.

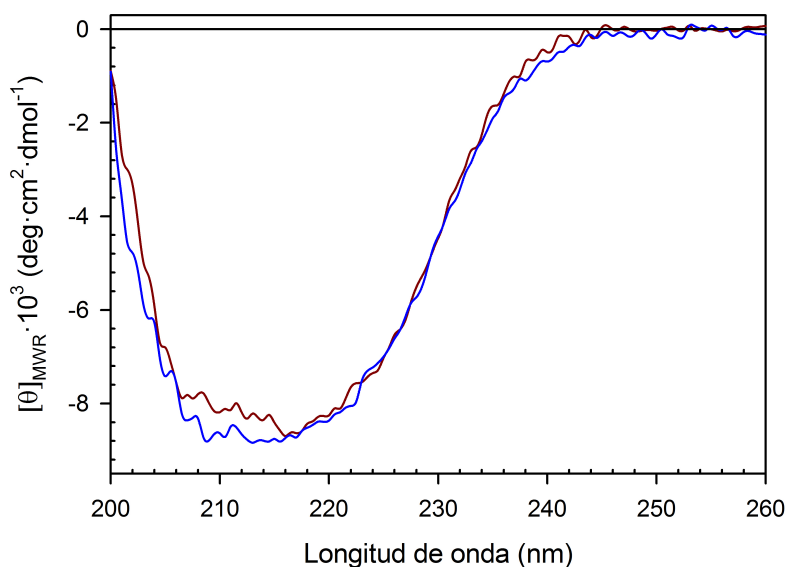


Figura 6.5. Comparación de los espectros de CD del dominio PDZ3-PSD95 en diferentes condiciones de pH, a 8 mg·mL⁻¹. En color azul está representado el espectro a pH 7.5 y en verde a pH 3.

6.1.3 Mecanismo de formación de fibras del dominio PDZ3-PSD95 a pH ácido

Para comprobar la existencia de fenómenos de asociación a pH 3, se realizó un barrido de temperatura seguido por DLS. Se tomó una muestra del dominio PDZ3-PSD95 a $8 \text{ mg}\cdot\text{mL}^{-1}$ en tampón glicina 50 mM pH 3 y se hizo un barrido de temperatura desde 20°C hasta 60°C . El resultado se muestra en la Figura 6.6 y permite confirmar que se están dando fenómenos de asociación, que comienzan a 56°C . A esa temperatura coexisten en la disolución el monómero y una especie de 4.7 nm (que contiene aproximadamente 10 monómeros). La especie que presenta oligomerización se da en una proporción que no llega al 10%.

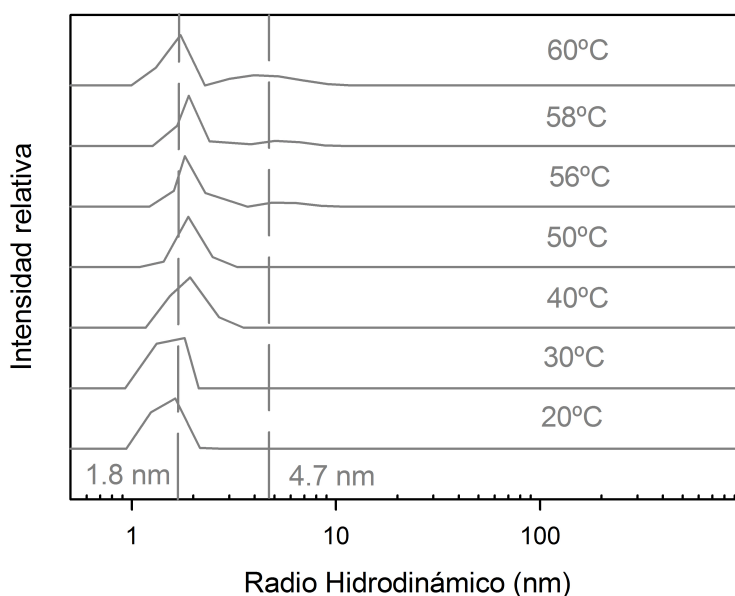


Figura 6.6. Barrido de temperatura desde 20°C hasta 60°C del dominio PDZ3-PSD95 a una concentración de $8 \text{ mg}\cdot\text{mL}^{-1}$ y pH 3 seguido por DLS. La velocidad de barrido fue de $60^\circ\text{C}\cdot\text{h}^{-1}$.

A continuación se dejó incubando varias horas a 60°C con el fin de estudiar la evolución de las especies presentes en la disolución. Básicamente, ambas especies permanecen en la disolución hasta las cuatro horas, pero con un incremento de masa de la especie de 4.7 nm de hasta un 40% (Tabla 6.3). Después de este tiempo, los monómeros desaparecen por completo y las partículas de 4.7 nm son las únicas presentes en la disolución. A continuación, estas partículas crecen dando tamaños mayores. Todos estos resultados, que se muestran en la Figura 6.7, indican que la asociación ocurre de forma distinta al proceso observado a pH neutro. A pH neutro los monómeros desaparecen y trimerizan antes de alcanzar 60°C , no existiendo tales trímeros a pH ácido. También a diferencia con pH neutro, el proceso no es reversible a pH ácido, puesto que al bajar la temperatura de nuevo a 20°C no se recupera el estado monomérico.

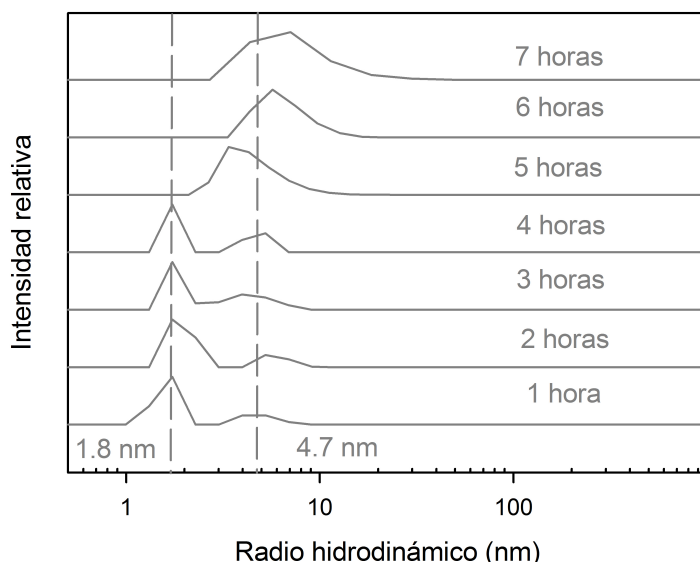


Figura 6.7. Evolución del radio hidrodinámico en función del tiempo de las especies presentes en una disolución del dominio PDZ3-PSD95 a una concentración de $8 \text{ mg}\cdot\text{mL}^{-1}$ en glicina 50 mM pH 3 seguida por DLS.

Tabla 6.3 Poblaciones de las distintas especies detectadas por DLS que presenta el dominio PDZ3-PSD95 cuando se incubó a 60°C durante siete horas en tampón glicina 50 mM pH 3.

Tiempo (horas)	Radio (nm)	% Masa	Peso Molecular (kDa)
1 h	1.6	75.9	10
	4.9	24.1	138
2 h	1.9	79.1	16
	6.0	20.9	220
3 h	1.8	57.7	13
	4.5	42.3	115
4 h	1.7	60.3	12
	4.7	39.7	128
5 h	4.7	100.0	126
6 h	6.6	100.0	280
7 h	7.8	100.0	414

Puesto que el intermedio oligomérico que se forma a pH neutro da lugar a la formación de estructuras anulares y fibrilares (Sección 5.1.2), se quiso establecer si la asociación que se da pH ácido también puede formar dichas estructuras. Para ello se incubaron varias muestras de PDZ3-PSD95 a $8 \text{ mg}\cdot\text{mL}^{-1}$ en 50 mM glicina pH 3 a distintos tiempos para verlas con el microscopio electrónico. La secuencia temporal de los procesos se resume en la Figura 6.8. En dicha Figura se observa la formación de fibras amiloides

con notables diferencias con respecto a la fibrilación que se da en condiciones de pH neutro. Se pueden observar algunas partículas de varios nm después de cuatro horas de incubación, aquellas de 4.7 nm vistas por DLS (micrografía señalada con “4 horas” de la Figura 6.8) y sólo hay estructuras fibrilares después de un día (micrografía señalada con “1 día” de la Figura 6.8). Con más tiempo de incubación estas fibras se hacen más largas y gruesas (micrografía señalada con “7 días” de la Figura 6.8).

No se observan estructuras anulares en ninguna de las muestras preparadas, estando de acuerdo con los resultados de DLS, en que no se detectan partículas de 12 nm. Además, las fibras a pH ácido crecen más rápidamente y maduran antes que las que se forman a pH neutro. La posibilidad de la formación de las estructuras anulares a pH neutro parece evitar la maduración de las fibras, mientras que a pH ácido las fibras crecen exclusivamente a expensas de los oligómeros precursores. Finalmente, a diferencia de las fibras a pH neutro, que desaparecen confirmando la reversibilidad del proceso, a pH ácido se dan de forma totalmente irreversible. Al diluirlas y ponerlas a temperatura ambiente siguen estando presentes en la disolución.

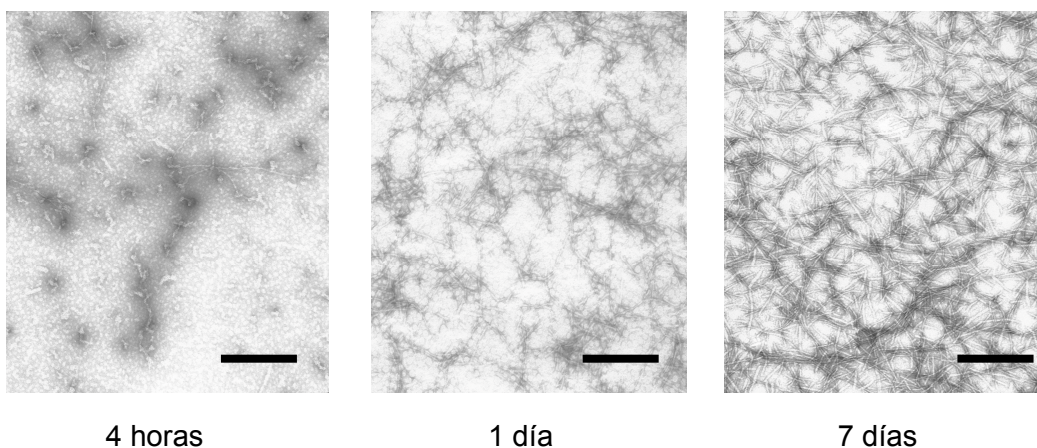


Figura 6.8. Análisis de la evolución de las fibras formadas a partir de una disolución del PDZ3-PSD95 a $8 \text{ mg}\cdot\text{mL}^{-1}$ y pH 3 seguidas por microscopía de transmisión electrónica. En el pie de cada micrografía está indicado el tiempo de incubación a 60°C de la disolución de proteína. La barra de escala corresponde a un tamaño de 200 nm.

6.1.4 Análisis cinético del crecimiento de fibras del dominio PDZ3-PSD95 a pH ácido

Puesto que se observan diferencias en el mecanismo de formación de fibras según el pH se realizaron varias pruebas espectroscópicas para estudiar en más profundidad las diferencias entre los distintos mecanismos. Para ello se hizo un barrido de temperatura del dominio PDZ3-PSD95 a $8 \text{ mg}\cdot\text{mL}^{-1}$ a pH 3 desde 20°C hasta 60°C en presencia de $12.5 \mu\text{M}$ de ThT y por otro lado $20 \mu\text{M}$ de ANS. Se observó que no hay aumento de señal de fluorescencia en el caso del experimento con ThT, mientras que hay un ligero incremento a partir de 50°C en la de ANS (Figura 6.9).

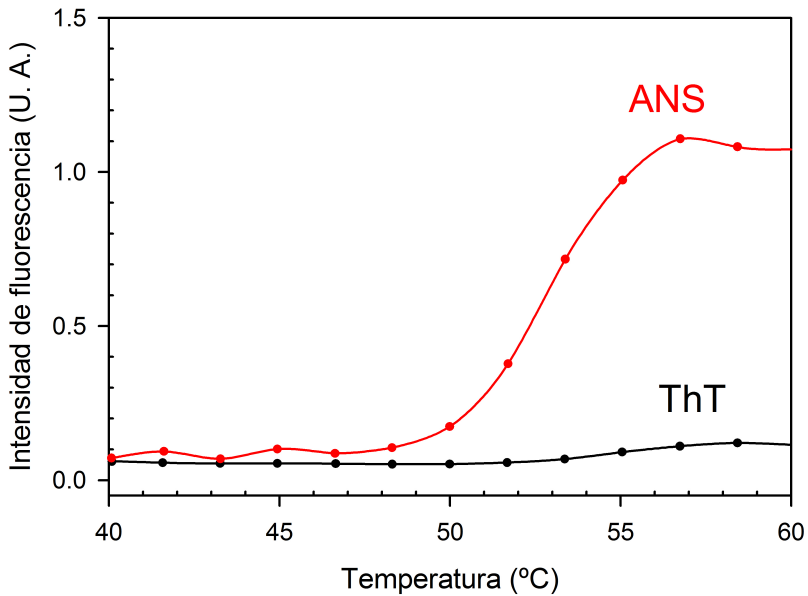


Figura 6.9. Barrido de temperatura seguido por la emisión de fluorescencia de ThT (negro) y ANS (rojo). Se incubó una muestra de PDZ3-PSD95 a $8 \text{ mg}\cdot\text{mL}^{-1}$ en presencia de $12.5 \text{ }\mu\text{M}$ de ThT o $20 \text{ }\mu\text{M}$ de ANS en glicina 50 mM pH 3.

A continuación se mantuvo la temperatura constante a 60°C y se fue registrando la señal durante un cierto tiempo. La emisión de fluorescencia presenta una curva cinética sigmoide (Paneles A y B de la Figura 6.10) con una fase de retardo de alrededor de 120 minutos, en los que la fluorescencia no aumenta prácticamente. Después de este tiempo, hay un incremento abrupto de la señal, que tiende a un máximo o *plateau*. La forma sigmoide de estas curvas no se observa a pH neutro, donde tenemos un crecimiento esencialmente no cooperativo, tal y como se explica en la Sección 5.2.1. Está aceptado que la fase de retardo de estas curvas se debe a una fase de nucleación en la que unas especies oligoméricas no fluorescentes se organizan previamente a las estructuras fibrilares, que crecerán a continuación según un régimen cooperativo y rápido. Este comportamiento es el más común entre la multitud de proteínas amiloidogénicas estudiadas hasta la fecha (Harper & Lansbury, 1997; Kelly, 2000; Kad et al., 2003; Knowles et al., 2009; Smith et al., 2006).

La diferencia sustancial en la forma de las curvas de ThT y ANS indica de nuevo el mecanismo diferente de fibrilación que presenta el dominio PDZ3-PSD95 a pH neutro y ácido. El mínimo incremento de fluorescencia de ANS y ThT después de unas dos horas indica que las especies no fluorescentes que se forman en el período de retardo (monómeros y partículas de 4.7 nm según los experimentos de DLS) no exponen superficies hidrofóbicas y no muestran la agregación β típica que sí hacen las especies triméricas y oligómeros de 12 nm a pH neutro. Por tanto, la naturaleza molecular de los núcleos formados en condiciones de pH diferentes es sustancialmente distinta.

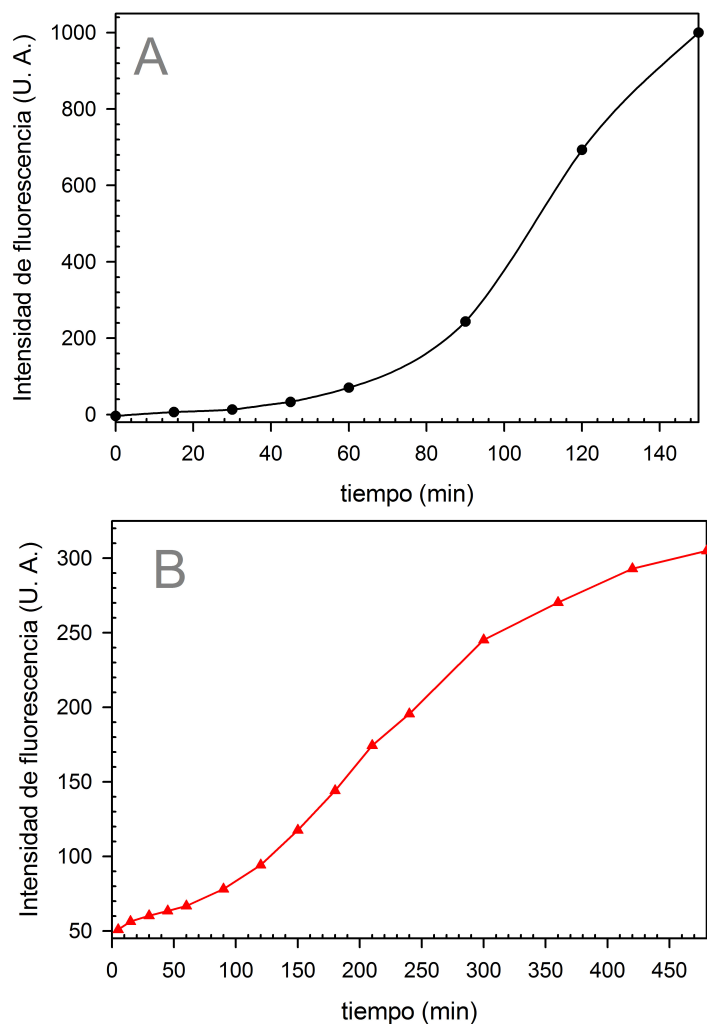


Figura 6.10. Cinéticas de crecimiento de la emisión de fluorescencia a 60°C de dos disoluciones del dominio PDZ3-PSD95 a $8 \text{ mg} \cdot \text{mL}^{-1}$ en glicina 50 mM pH 3, una en presencia de $12.5 \text{ } \mu\text{M}$ de ThT (Panel A) y otra en presencia de $20 \text{ } \mu\text{M}$ de ANS (Panel B).

Es interesante notar que a pH 3 los experimentos de DLS no indican la población de otras partículas distintas a monómeros y partículas de 4.7 nm hasta las cuatro horas, lo cual no coincide con las dos horas de fase de retardo. La cuestión que surge es qué especie o especies son las responsables del incremento en fluorescencia después de la fase de retardo. Una posible explicación podría ser que el incremento de fluorescencia de las dos sondas que se da después de esas dos horas se deba a una reestructuración interna de los monómeros y/o oligómeros. Este cambio conformacional podría permitir la unión de ambas sondas fluorescentes al dominio. Este posible cambio no se detecta por CD, siendo todos los espectros recogidos a las diferentes horas prácticamente coincidentes en forma, mostrando un mínimo a 217 nm que indica la presencia de estructura β y que va aumentando en intensidad conforme aumenta el tiempo (Figura 6.11).

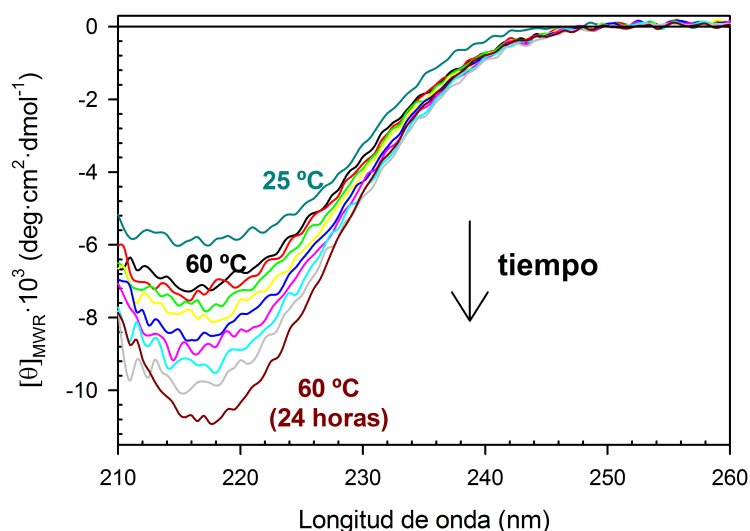


Figura 6.11. Espectros de CD del dominio PDZ3-PSD95 a $8 \text{ mg}\cdot\text{mL}^{-1}$ en tampón glicina 50 mM pH 3 a 25°C (espectro en color verde) y a distintos tiempos de incubación a 60°C , 2 minutos (negro), 1 hora (rojo), 2 horas (verde), 3 horas (amarillo), 4 horas (azul), 5 horas (rosa), 6 horas (cian), 7 horas (gris) y 24 horas (marrón).

Finalmente, después de la incubación a 60°C , se enfrió la disolución hasta 20°C de nuevo para ver si alcanzaba el estado inicial (Figuras 6.12 y 6.13). De nuevo con ambas sondas se observa que el proceso es irreversible, ya que en ninguno de los dos casos la señal disminuye con el tiempo. Esto concuerda con las evidencias obtenidas por DLS y TEM, confirmando la diferencia con el comportamiento a pH neutro, que sí es reversible (ver Sección 5.1.1).

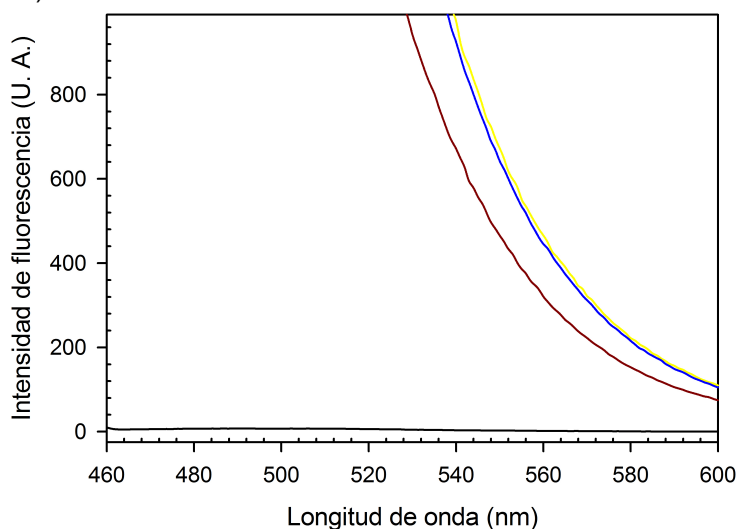


Figura 6.12. Espectros de emisión de fluorescencia de ThT una vez enfriada la disolución del dominio PDZ3-PSD95 a $8 \text{ mg}\cdot\text{mL}^{-1}$ en glicina 50 mM pH 3 a temperatura ambiente después de haberla calentado a 60°C durante unas horas. Los espectros han sido recogidos a distintos tiempos después del enfriamiento de la muestra a 25°C : 30 minutos (amarillo), 2 horas y media (azul) y un día (marrón). Se ha incluido también el espectro a 25°C antes de calentar, en color negro.

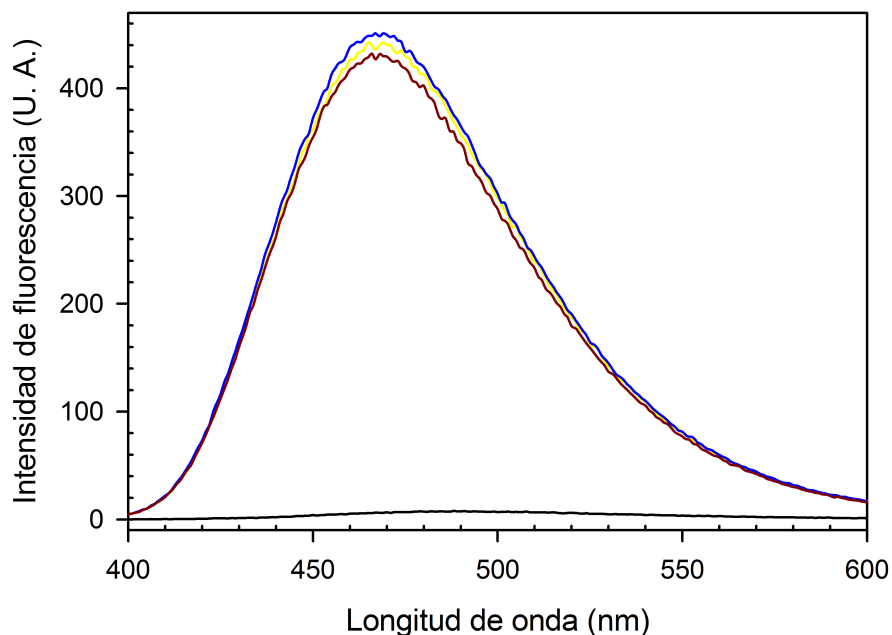


Figura 6.13. Espectros de emisión de fluorescencia de ANS una vez enfriada la disolución del dominio PDZ3-PSD95 a $8 \text{ mg}\cdot\text{mL}^{-1}$ en glicina 50 mM pH 3 a temperatura ambiente después de haberla calentado a 60°C durante unas horas. Los espectros han sido recogidos a distintos tiempos después del enfriamiento de la muestra a 25°C : 30 minutos (amarillo), 2 horas y media (azul) y un día (marrón). Se ha incluido también el espectro a 25°C antes de calentar, en color negro.

6.2 Discusión

Como se describe en el Capítulo 5, el desplegamiento en el equilibrio del dominio PDZ3-PSD95 a pH 7.5 muestra un intermedio parcialmente plegado que presenta fenómenos de asociación, cuya estequiometría es un trímero, y que conduce a la formación reversible de estructuras anulares y fibrilares (panel señalado con “pH 7.5” de la Figura 6.14). Sin embargo, a valores de pH menores de 3, este dominio se desnaturaliza según el modelo de dos estados. Sin embargo, los resultados de los experimentos realizados mediante DLS y fluorescencia de ThT y ANS muestran que dicho intermedio sigue formándose, aunque el hecho de que no se pueble indica que se desestabiliza energéticamente. Esta desestabilización, sin embargo, no provoca la desaparición de la formación de fibras amiloides que se observa mediante TEM, aunque hace que sea más rápida e irreversible, y que no se formen estructuras anulares (panel señalado con “pH 3” de la Figura 6.14). Este cambio de comportamiento que se produce a valores de pH 3 y menores, conduce a pensar que los residuos responsables de la desestabilización del intermedio han de ser de tipo glutamato y/o aspartato, por los valores de pK_a de su cadena lateral.

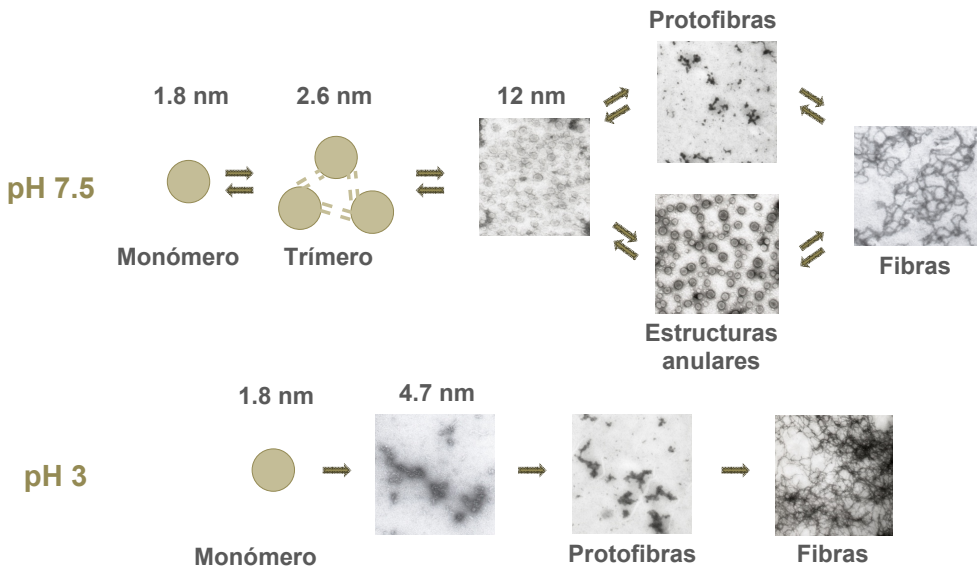


Figura 6.14. Esquema en que se representan las distintas etapas por las que pasa el dominio PDZ3-PSD95 cuando forma estructuras supramacromoleculares en condiciones de pH neutro y ácido.

6.2.1 Análisis termodinámico de la irreversibilidad de los procesos de fibrilación del dominio PDZ3-PSD95 a pH ácido

Tal y como se ha mostrado previamente en algunos ejemplos (Kardos et al., 2004) el carácter energético de los procesos de fibrilación se puede explicar mediante criterios termodinámicos. La estabilidad del dominio PDZ3-PSD95 a $\text{pH} \leq 3$ está en el intervalo de $15\text{-}25 \text{ kJ}\cdot\text{mol}^{-1}$ (ver Tabla 6.2), mientras que la estabilidad de las estructuras macromoleculares está estimada en unos $25 \text{ kJ}\cdot\text{mol}^{-1}$, según se vio en la Sección 5.2.1. De hecho, ya se explicó en dicha Sección que la mayor estabilidad del estado nativo a pH neutro (que es de unos $40 \text{ kJ}\cdot\text{mol}^{-1}$) explica la reversibilidad del proceso. Ahora en cambio, a valores ácidos de pH la estabilidad de las fibras es ligeramente mayor, prefiriéndose por eso este estado. Por otro lado, la mayor velocidad en la formación de las fibras a pH 3 que a pH 7.5 podría ser debida a que no se da la competencia en la formación de estructuras anulares y fibrilares que se da a pH 7.5. De esta forma, las fibras a pH 3 crecen exclusivamente a partir de los oligómeros precursores.

6.2.2 Influencia de las fuerzas electrostáticas como responsables del comportamiento diferencial en el plegamiento del dominio PDZ3-PSD95 a pH neutro y ácido

El papel relevante de las interacciones electrostáticas en la organización de las fibras amiloides ha sido ya descrito en otros ejemplos (Gerber et al., 2008; Juárez et al., 2009). Se sabe que las variaciones en el pH pueden modular el balance electrostático de

las fuerzas intra e intermoleculares que se dan en las moléculas de proteína. El comportamiento distinto del crecimiento de las fibras del dominio PDZ3-PSD95 observado a pH neutro (no sigmoidal) y pH ácido (sigmoidal) debe ser una consecuencia directa de las distintas estabilidades del intermedio precursor. Así, la presencia de una fase de retardo a valores ácidos de pH apunta a que la formación de estructuras oligoméricas necesita alcanzar un número o tamaño crítico previo a inducir una posterior agregación. A pH neutro este proceso está energéticamente favorecido por el estado intermedio precursor que ya oligomeriza durante el calentamiento desde temperatura ambiente hasta 60°C (Sección 5.1.2). Este comportamiento dependiente del pH ya ha sido observado en los procesos de fibrilación de la albúmina de suero humano (Juárez et al., 2009).

Los experimentos cinéticos también muestran que la duración de la fase de retardo es de unas dos horas aproximadamente, pero no se observa ninguna partícula nueva durante este tiempo que pueda corresponderse con el aumento de fluorescencia. Así, el monómero y las partículas de 4.7 nm son las únicas especies presentes hasta unas cuatro horas (ver Figura 6.7 y Tabla 6.3). Además, la ausencia de fluorescencia debida a las sondas ANS y ThT de dichas partículas hace pensar que estas especies no exponen superficies hidrofóbicas ni muestran la agregación β características de las estructuras amiloides, como sí que ocurre a pH neutro con los trímeros y las partículas de 12 nm (ver Sección 5.1.4). Todo esto está indicando una diferencia estructural de las especies asociadas que se dan a los distintos pH. De esta forma, el aumento de fluorescencia después de la fase de retardo a pH 3 sería la consecuencia de un cambio conformacional de las moléculas de proteína previo a la fibrilación, pero que no ha podido ser detectado por CD.

6.2.3 Análisis de las posiciones de los aminoácidos aspartato y glutamato en la estructura del dominio PDZ3-PSD95 que pueden ser responsables de la desestabilización energética del intermedio a pH ácido

La desestabilización del estado intermedio a $\text{pH} \leq 3$ se ha de deber a un equilibrio de protonación de restos glutamato o aspartato, puesto que son los únicos cuyos valores de pK_a están en el rango de pH 3-4. Las estructuras del dominio PDZ3-PSD95 (Capítulo 4) muestran que estos residuos están muy expuestos al disolvente, con algunas excepciones. Así, algunos glutamatos parecen ser los principales responsables de la unión de la tercera hélice α al resto del dominio completo (Figura 6.15). Estos residuos, principalmente el Glu334 de la lámina β_3 y Glu396 y Glu401 de la hélice α_3 , pueden interactuar por medio de sus cadenas alifáticas entre sí y con los anillos aromáticos de los residuos Phe337 de la lámina β_3 y Tyr397 y Phe400 de la hélice α_3 . Sin embargo, cuando estos glutamatos se desprotonan por encima de pH 3, puede haber cierta repulsión electrostática entre sus respectivas cargas negativas. Además, puede haber una contribución a esa repulsión por parte de la nube de electrones π de los anillos aromáticos (Ma & Dougherty, 1997; Lund & Jönsson, 2005). Este fenómeno podría generar la apertura de la hélice α_3 y de la región β_2 - β_3 , tal y como se ha observado por resonancia magnética nuclear (Feng et al., 2005).

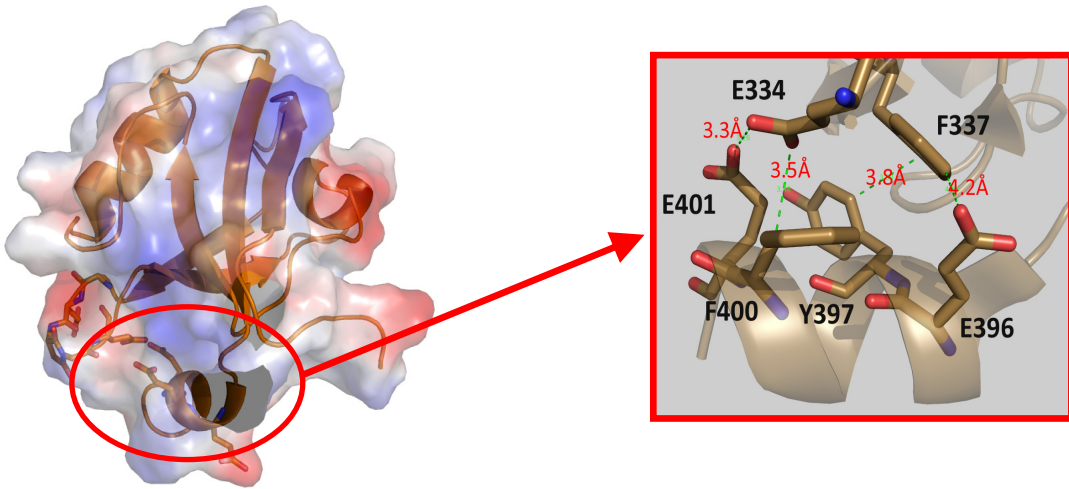


Figura 6.15. Superficie electrostática del dominio PDZ3-PSD95 junto con el detalle de la interacción entre la hélice $\alpha 3$ y el giro entre las cadenas $\beta 2$ y $\beta 3$ en el que se muestra la disposición espacial de la Phe400, Tyr397 y Glu396. Las líneas discontinuas de color verde representan las distancias relativas entre residuos.

Como ya se comentó en el Capítulo 2, esta hélice $\alpha 3$ constituye un elemento estructural adicional comparado con otros dominios PDZ típicos, que se caracterizan por tener seis láminas β y sólo dos hélices α . Además, esta hélice ha sido descrita como un elemento alostérico regulador que puede modular la afinidad de unión del dominio PDZ3-PSD95 (Petit et al., 2009).

Por otro lado, si se observa el giro que conecta las cadenas $\beta 2$ y $\beta 3$, formado por los residuos Gly329-Gly335 (GGEDGEG), las estructuras resueltas en esta Memoria (ver Sección 4.2.2) indican la presencia de una succinimida en la posición del Asp332. Desde un punto de vista conformacional, esta modificación disminuye la flexibilidad del giro a la vez que ocurre una pérdida de una carga negativa de la cadena lateral. Aparte de estas observaciones, la única característica relevante del resto de residuos ácidos es un puente salino presente entre los residuos Arg312-Asp357, que probablemente no tiene un papel importante al cambiar las condiciones de pH, puesto que no está expuesto al disolvente.

6.2.4 Racionalización del mecanismo de desplegamiento térmico del dominio PDZ3-PSD95

El análisis de la tendencia a agregación β del dominio PDZ3-PSD95 mediante el algoritmo *Tango* (Fernandez-Escamilla et al., 2004) no revela diferencias sustanciales entre pH 7.5 y pH 3, siendo los residuos 340-350 (pertenecientes a las estructuras secundarias $\beta 3$ y $\alpha 1$) y 392-397 (de $\beta 5$ y $\alpha 3$) los más propensos a agregación β . La posible repulsión existente entre los residuos Glu334 de $\beta 3$ y Glu396, Glu401 de $\alpha 3$ cuando el pH es neutro puede hacer salir hacia fuera del núcleo hidrofóbico de la proteína los residuos vecinos con tendencia a agregación β de las regiones $\beta 3$ - $\alpha 1$ y $\beta 5$ - $\alpha 3$ (Figura 6.16). Estos intermedios parcialmente plegados se estabilizan entre ellos a través de la formación de especies

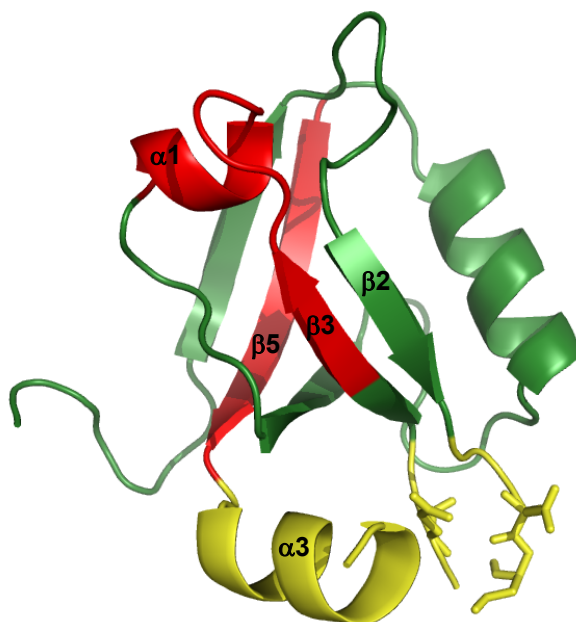


Figura 6.16. Estructura del dominio PDZ3-PSD95 (código PDB 3K82) con las regiones con más propensión a agregación β en color rojo y la hélice $\alpha 3$ y el giro entre las cadenas $\beta 2$ y $\beta 3$ en color amarillo. Las cadenas laterales de los residuos de este giro, entre los que se encuentra el anillo de succinimida, se han representado con barras (sticks).

triméricas vistas a pH neutro. Sin embargo, a $\text{pH} \leq 3$, los glutámicos protonados no presentarían esas fuerzas electrostáticas repulsivas y así no se expondrían al exterior del dominio PDZ3-PSD95, permaneciendo como un monómero en disolución. Las partículas de 4.7 nm podrían deberse a una agregación inespecífica de monómeros tal y como ya se ha visto en otros ejemplos (Smith et al., 2006). Estas características podrían explicar la ausencia de fluorescencia de ANS y ThT de las partículas ácidas y el carácter altamente fluorescente de dichas sondas de las observadas a pH neutro.

La apertura de la hélice $\alpha 3$ descrita a partir de experimentos de intercambio de hidrógeno con resonancia magnética nuclear apoya esta hipótesis (Feng et al., 2005). Además, ya se ha hablado en la Sección 4.2.2 de la interconexión estructural entre la hélice $\alpha 3$ y el giro que conecta las láminas $\beta 2$ y $\beta 3$. Por otro lado, teniendo en cuenta lo que ya se ha comentado de que esta tercera hélice α es un elemento regulador de las propiedades de unión del dominio PDZ3-PSD95 (Petit et al., 2009) y que es un elemento estructural que no se encuentra en la topología típica de los dominios PDZ, se podría concluir que la hélice $\alpha 3$ parece ser también un elemento regulador del plegamiento del dominio PDZ3-PSD95. Para apoyar esta idea, se puede señalar la diferente naturaleza físico-química de la interacción entre la hélice $\alpha 3$ y el resto del dominio, que es principalmente entre algunos restos glutamato y residuos aromáticos, y el resto de los elementos estructurales secundarios, en los que el tipo de interacciones entre sí son fundamentalmente entre cadenas laterales alifáticas (Gianni et al., 2006).

CAPÍTULO 7

ANÁLISIS TERMODINÁMICO Y CINÉTICO
DEL EQUILIBRIO CONFORMACIONAL DEL
TERCER DOMINIO PDZ DE LA PROTEÍNA
PSD95 SIN SU TERCERA HÉLICE α

7.1 Resultados

El comportamiento diferente del dominio PDZ3-PSD95 en su desplegamiento en el intervalo de pH entre los valores 3 y 4, en el que a valores de pH por debajo de 3 muestra un comportamiento de dos estados, mientras que a valores de pH por encima de 3.5 sigue un mecanismo de tres estados, indica que los residuos responsables de dicho comportamiento han de ser de tipo glutamato y/o aspartato, debido que son los únicos que presentan un pK_a que está en torno a ese valor. Además, han de ser aquellos que están en la hélice α_3 , así como los presentes en el giro entre las láminas β_2 y β_3 (Figura 7.1), puesto que el resto de estos aminoácidos tiene las cadenas laterales muy expuestas al disolvente y no presentan interacciones relevantes para el plegamiento del dominio. En definitiva, los resultados previos sugieren un papel destacado de la hélice α_3 en el plegamiento de este dominio.

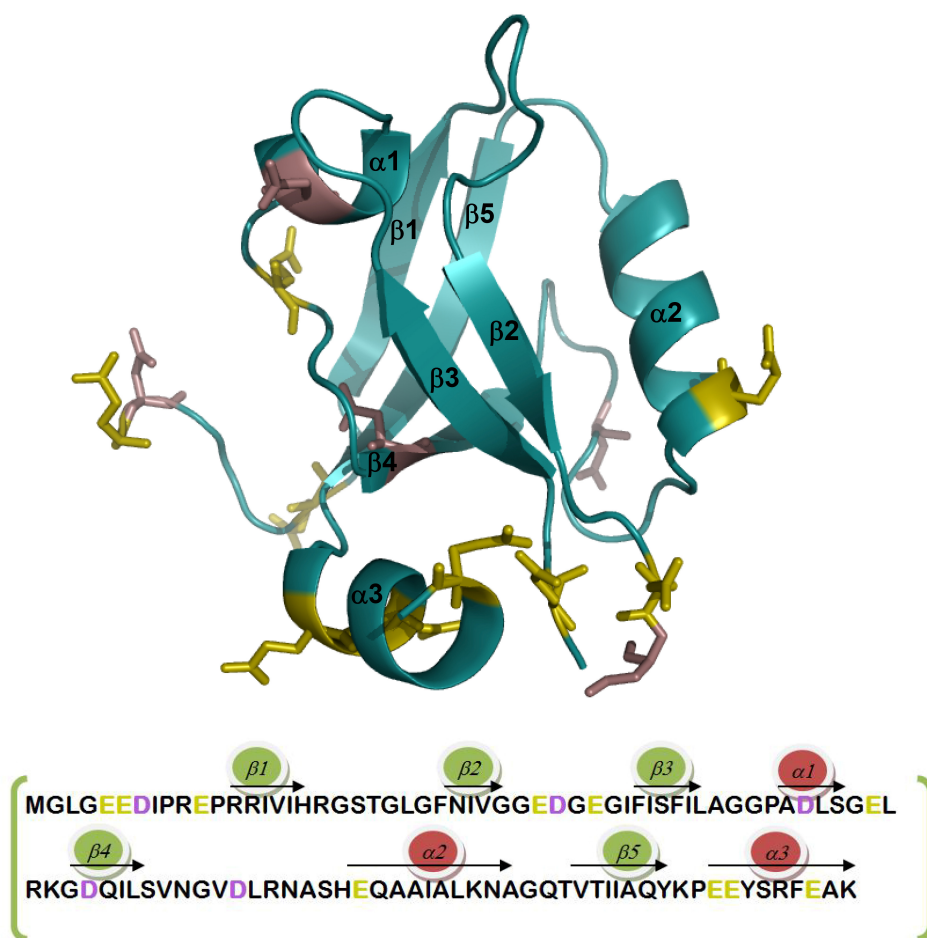


Figura 7.1. Estructura del dominio PDZ3-PSD95 (código PDB 3K82) y su secuencia de aminoácidos con los respectivos elementos estructurales secundarios indicados. En violeta están señalados los aspartatos (D), mientras que en amarillo los glutamatos (E).

Por otro lado, la estructura canónica de un dominio PDZ contiene cinco o seis láminas β y únicamente dos hélices α (Fanning & Anderson, 1996), siendo la tercera hélice α del dominio PDZ3-PSD95 un elemento estructural adicional que está presente en la proteína PSD95 (Doyle et al., 1996). Hay que tener en cuenta que esta hélice se conserva en la familia de proteínas MAGUK (Morais Cabral et al., 1996) sirviendo de elemento de unión entre el tercer dominio PDZ y el dominio SH3 de estas proteínas (Kim & Sheng, 2004), por lo que ha de tener un papel biológico concreto, aunque este papel aún no se conoce.

La importancia de la hélice α_3 ha sido corroborada con el papel regulador que ésta tiene en las propiedades de unión del dominio PDZ3-PSD95. Este dominio presenta una regulación alostérica en su unión a ligandos debida a este elemento estructural cuyos residuos no interactúan directamente con el ligando. Así, cuando se estudiaron las propiedades de unión de este dominio sin la hélice α_3 , se observó que la afinidad varía notablemente, en un orden 20 veces menor (Petit et al., 2009). Esta variación parece no estar acompañada de un cambio estructural, y se debería a fenómenos dinámicos de carácter entrópico.

7.1.1 Estudio de la interacción del dominio PDZ3-PSD95 y de su construcción sin la hélice α_3 con el péptido de secuencia KKETAV mediante ITC

A la vista de la importancia evidente de esta hélice dentro de este dominio se decidió estudiar las consecuencias de su eliminación en el proceso de plegamiento, para lo cual se clonó y purificó una nueva construcción del tercer dominio PDZ de la proteína PSD95, en la que se eliminó esta hélice α_3 . Esta construcción, a la que se le va a denominar PDZ3-h-PSD95 a lo largo de la Memoria, contiene los residuos 302 a 393 de la proteína PSD95 (numeración proveniente de (Doyle et al., 1996)). Se han truncado, por tanto, los nueve residuos presentes en el extremo C-terminal del dominio PDZ3-PSD95, cuya secuencia aminoacídica es EEYSRFEAK.

Se comprobó inicialmente que efectivamente la hélice es un elemento regulador de las propiedades de unión del dominio. Para ello se realizaron experimentos de ITC en las condiciones de tampón con que se han hecho gran parte de los experimentos de plegamiento, fosfato 50 mM pH 7.5. Dichos experimentos de unión se realizaron entre cada una de las dos construcciones, tanto del dominio PDZ3-PSD95 como del dominio PDZ3-h-PSD95 y un péptido cuya secuencia aminoacídica es KKETAV, descrito en la literatura como uno de los ligandos que presentan mayor afinidad para el dominio PDZ3-PSD95 (Saro et al., 2007). En el caso del dominio PDZ3-PSD95, la concentración de proteína usada fue de $0.2 \text{ mg}\cdot\text{mL}^{-1}$, mientras que la de péptido fue de 0.2 mM. El ligando se inyectó a la célula calorimétrica con volúmenes constantes, en intervalos de 10 μL . Estas mismas condiciones se probaron inicialmente en la construcción PDZ3-h-PSD95, pero los calores de inyección no generaron una señal lo suficientemente adecuada como para poder analizarla, por lo que se usó un perfil de volúmenes. De esta forma, se empezó inyectando pequeños volúmenes de ligando al principio para a continuación inyectar volúmenes

mayores. En este caso, la concentración de proteína fue de $1 \text{ mg}\cdot\text{mL}^{-1}$ y la de ligando de 2.4 mM .

Los calores detectados por ITC fueron analizados con el modelo de modelo de “n” sitios de unión idénticos e independientes, puesto que es bien conocido que estos dominios unen únicamente a una molécula diana. Los resultados de los análisis de ambos experimentos confirman la diferencia en afinidad entre sendas construcciones (Figura 7.2), que es de unas cuatro veces, tal y como reflejan sus constantes de disociación con valores de $3.6 \text{ }\mu\text{M}$ y $1 \text{ }\mu\text{M}$ respectivamente (Tabla 7.1). Esta diferencia no se debe a factores entálpicos, puesto que los valores de entalpías para ambas construcciones son muy parecidos (Tabla 7.1), tal y como describen Petit y colaboradores (Petit et al., 2009). Sin embargo, esta diferencia en afinidad no es tan acusada como la descrita en dicho trabajo, en que la construcción con la hélice $\alpha 3$ resulta hasta veinte veces más afín. Posiblemente, en el caso del trabajo de Petit, al usar un péptido de secuencia distinta (TKNYKQTSV) la influencia de la hélice $\alpha 3$ tiene más relevancia.

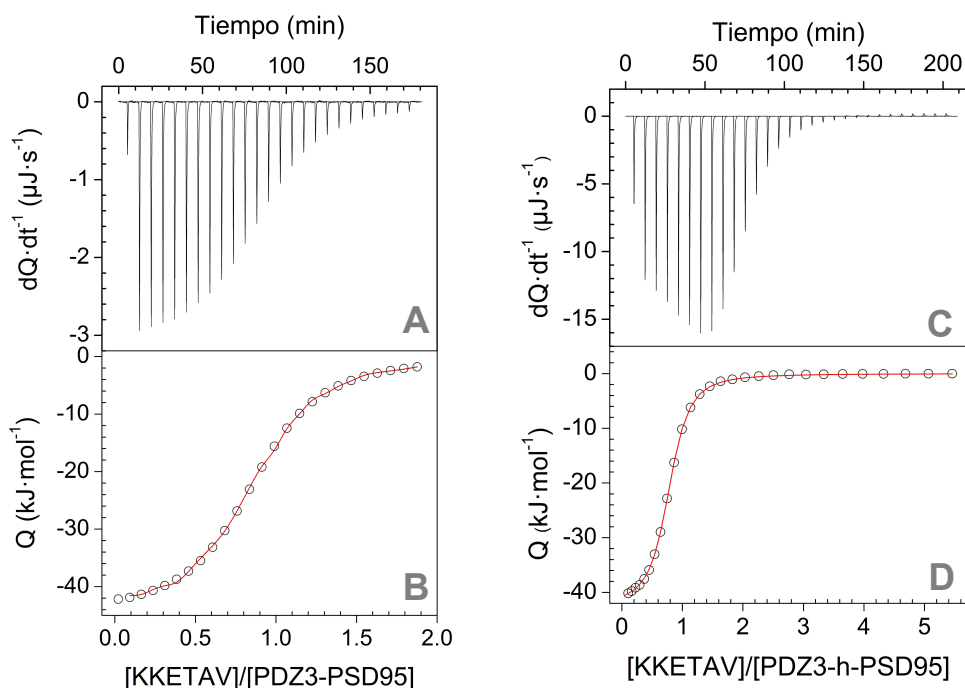


Figura 7.2. Experimentos de titulación de los dominios PDZ3-PSD95 (paneles A y B) y PDZ3-h-PSD95 (paneles C y D) con el péptido de secuencia KKETAV mediante ITC en fosfato 50 mM pH 7.5 . (A) Termograma de titulación de los calores por unidad de tiempo liberados tras la inyección del péptido (25 inyecciones de $10 \text{ }\mu\text{L}$ de una disolución de péptido a 0.2 mM) sobre la proteína PDZ3-PSD95 ($20 \text{ }\mu\text{M}$ en la célula calorimétrica). (B) Isotherma de unión correspondiente al experimento en (A), obtenida después de normalizar los calores de titulación con la concentración de péptido y proteína. (C) Termograma de titulación de los calores por unidad de tiempo liberados tras la inyección del péptido (inyecciones de volúmenes variables de una disolución de péptido a 2.4 mM sobre la proteína PDZ3-h-PSD95 ($100 \text{ }\mu\text{M}$ en la célula calorimétrica). (D) Isotherma de unión correspondiente al experimento en (C), obtenida después de normalizar los calores de titulación. Los datos experimentales se representan como círculos, mientras que la línea continua corresponde al mejor ajuste para el modelo de “n” sitios de unión idénticos e independientes.

Tabla 7.1. *Parámetros termodinámicos obtenidos mediante el ajuste de los datos experimentales de la unión del ligando con secuencia KKETAV con las construcciones PDZ3-PSD95 y PDZ3-h-PSD95 en tampón fosfato 50 mM pH 7.5 mediante ITC al modelo de “n” sitios de unión idénticos e independientes.*

	ΔH_d (kJ·mol ⁻¹)	K_d (μ M ⁻¹)	n
PDZ3-PSD95	-44.7 ± 0.6	1.0 ± 0.4	0.90 ± 0.04
PDZ3-h-PSD95	-42.5 ± 0.2	3.6 ± 0.1	0.81 ± 0.01

7.1.2 Mecanismo de desplegamiento térmico del dominio PDZ3-h-PSD95 a pH neutro

El efecto de la eliminación de este elemento estructural en el plegamiento del dominio se ha estudiado mediante DSC. Se llevaron a cabo experimentos en las mismas condiciones de tampón que se usaron para el dominio PDZ3-PSD95, fosfato 50 mM pH 7.5, usando varias concentraciones de proteína y una velocidad de barrido de 1.5 K·min⁻¹. Las trazas calorimétricas obtenidas con el dominio PDZ3-h-PSD95 presentan dos transiciones (Figura 7.3), al igual que las obtenidas con el dominio PDZ3-PSD95. Además, el cambio en la posición de las trazas en función de la concentración de proteína es equivalente en ambas construcciones. De forma análoga al dominio PDZ3-PSD95, al aumentar la concentración del dominio PDZ3-h-PSD95 la primera endoterma se desplaza hacia temperaturas menores, mientras que la segunda lo hace hacia temperaturas mayores. Tal y como se explica en la Sección 5.1.1, el desplazamiento de las trazas en la escala de temperatura al cambiar la concentración se relaciona con fenómenos de asociación. Por lo tanto, de manera similar, en este caso, la primera endoterma refleja un proceso de asociación del estado final, mientras que la segunda uno de disociación del estado inicial. Sin embargo, se observa una diferencia cualitativa entre la forma de las trazas de las dos construcciones, siendo más agudas las transiciones del dominio PDZ3-h-PSD95 (Figura 7.3) con respecto a las del dominio PDZ3-PSD95. Además también presentan una diferencia en el grado de reversibilidad. En el caso del dominio PDZ3-PSD95 es de un 60%, mientras que en el dominio PDZ3-h-PSD95 no llega al 10% (Figura 7.4).

Para determinar que los experimentos realizados con el dominio PDZ3-h-PSD95 están en condiciones de equilibrio y poder aplicar así los modelos termodinámicos para su análisis, se realizaron nuevos experimentos de DSC en las mismas condiciones de tampón, fosfato 50 mM pH 7.5, y una misma concentración de proteína, 1.7 mg·mL⁻¹, pero a diferentes velocidades de barrido: 2 K·min⁻¹, 1 K·min⁻¹, 0.5 K·min⁻¹ y 0.25 K·min⁻¹. Se observa que los máximos de las trazas se desplazan ligeramente hacia temperaturas mayores conforme aumenta la velocidad de barrido (Figura 7.4), aunque esta tendencia, sin embargo, no afecta la forma de las trazas, por lo que la información obtenida mediante el uso de la termodinámica del equilibrio resultaría válida.

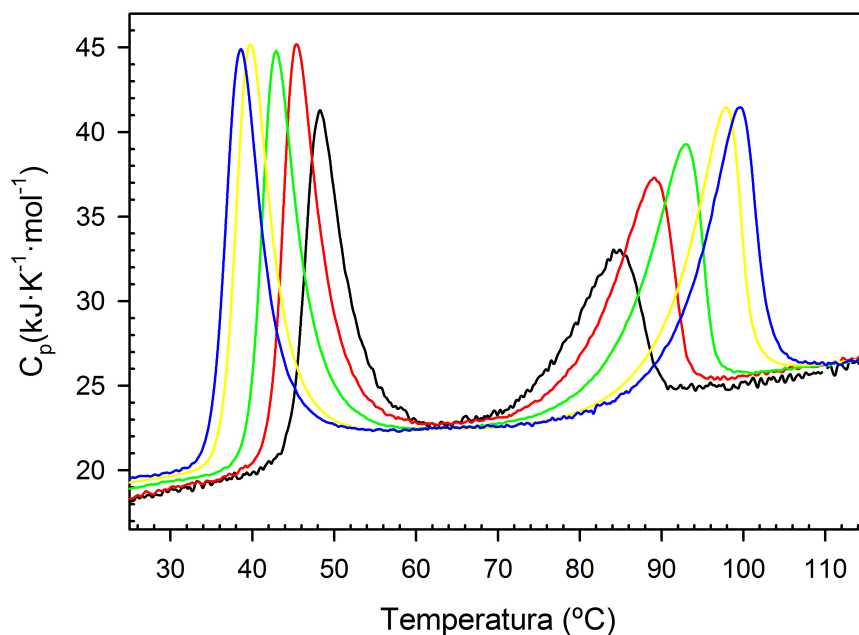


Figura 7.3. Curvas de desplegamiento térmico del dominio PDZ3-h-PSD95 seguidas por DSC en función de la concentración de proteína en fosfato 50 mM pH 7.5. Las concentraciones de proteína son $0.4 \text{ mg}\cdot\text{mL}^{-1}$ (negro), $0.8 \text{ mg}\cdot\text{mL}^{-1}$ (rojo), $1.7 \text{ mg}\cdot\text{mL}^{-1}$ (verde), $4.8 \text{ mg}\cdot\text{mL}^{-1}$ (amarillo) y $7.2 \text{ mg}\cdot\text{mL}^{-1}$ (azul).

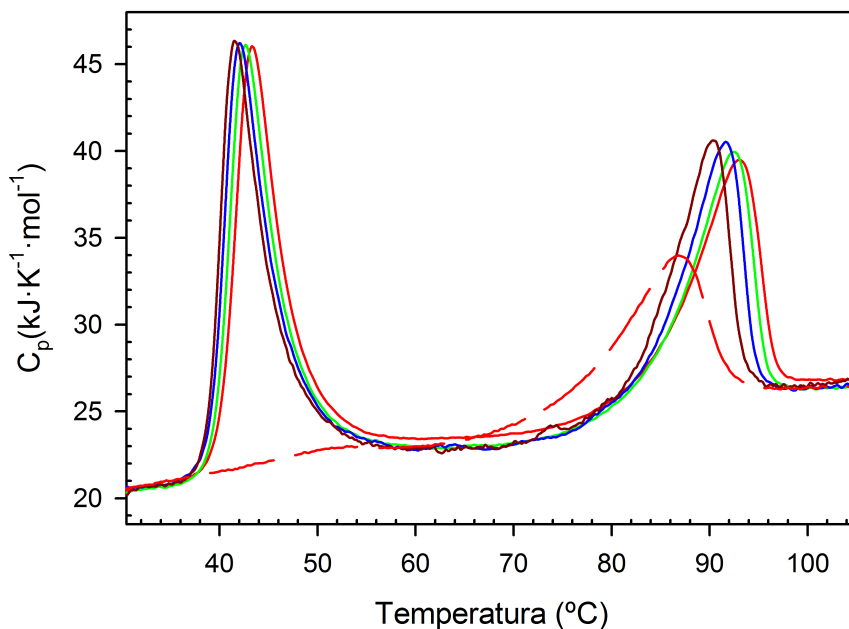


Figura 7.4. Curvas de DSC a distintas velocidades de barrido para el dominio PDZ3-h-PSD95 a una concentración de $1.7 \text{ mg}\cdot\text{mL}^{-1}$ y pH neutro. Las velocidades son $2 \text{ K}\cdot\text{min}^{-1}$ (rojo), $1 \text{ K}\cdot\text{min}^{-1}$ (verde), $0.5 \text{ K}\cdot\text{min}^{-1}$ (azul) y $0.25 \text{ K}\cdot\text{min}^{-1}$ (marrón). El segundo barrido de la traza de $2 \text{ K}\cdot\text{min}^{-1}$ está representado con una línea discontinua en color rojo. Para simplificar la gráfica sólo se ha mostrado el segundo barrido del experimento realizado a $2 \text{ K}\cdot\text{min}^{-1}$.

Adicionalmente, se comprobó la reversibilidad del proceso de desplegamiento térmico mediante el registro de espectros de CD en fosfato 50 mM pH 7.5 a una concentración de PDZ3-h-PSD95 de $0.3 \text{ mg}\cdot\text{mL}^{-1}$. Inicialmente la temperatura fue 2°C , en que la proteína está en el estado nativo. A continuación se aumentó hasta 98°C , en que está desnaturalizada. Luego, al enfriar la disolución hasta la temperatura inicial de 2°C se comprobó que la forma del espectro es la misma, aunque no se recupera toda la intensidad de la señal (Figura 7.5). Todos estos resultados indican que hay una menor reversibilidad en el mecanismo de plegamiento del tercer dominio PDZ de la proteína PSD95 cuando se elimina su hélice α_3 .

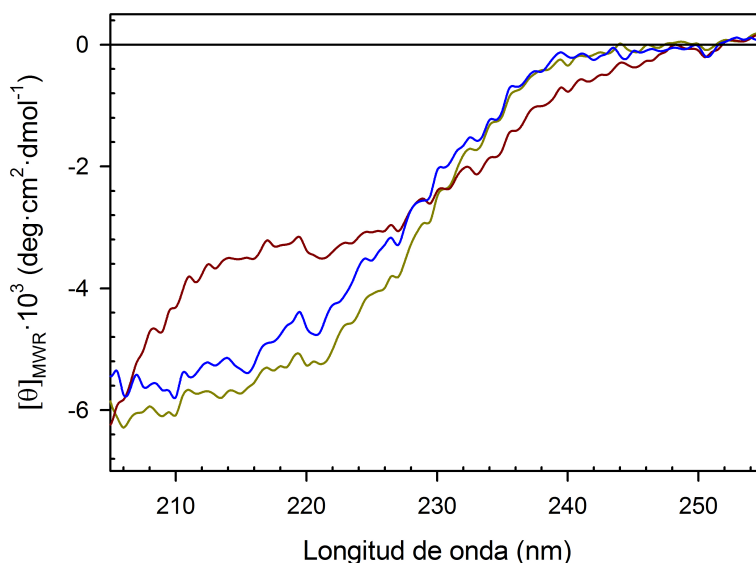


Figura 7.5. Espectros de CD del dominio PDZ3-h-PSD95 en fosfato 50 mM pH 7.5 a una concentración de $0.3 \text{ mg}\cdot\text{mL}^{-1}$. En verde se representa el espectro de la proteína nativa a 2°C y en rojo el espectro a 98°C . Para comprobar la reversibilidad se enfrió la muestra dentro del dicrógrafo hasta 2°C y se registró otro espectro a los 60 minutos después de haber alcanzado esta temperatura (espectro en color azul).

Puesto que las trazas calorimétricas obtenidas para los dos dominios de estudio tienen el mismo comportamiento experimental en función de la concentración de proteína, se trataron de analizar las trazas con el mismo modelo con el que se analizaron las correspondientes al dominio PDZ3-PSD95. Dicho modelo, cuyo esquema es $nN \rightleftharpoons I_n \rightleftharpoons nD$, es el más sencillo para explicar un proceso que evidencia la existencia de tres estados conformacionales y en el que se produce un fenómeno de asociación y uno de disociación. Ya se ha comentado que al haber dos transiciones deben existir al menos tres estados conformacionales y que el desplazamiento observado de las trazas a lo largo del eje de la temperatura en función de la concentración implica un fenómeno de asociación y otro de disociación. El ajuste de nuestros datos experimentales a dicho modelo proporcionó resultados de baja calidad, sugiriendo la invalidez del mismo. Así pues, teniendo en cuenta

los fenómenos de asociación-disociación debidos a los efectos de concentración y la presencia de al menos tres estados conformacionales macroscópicos, el esquema alternativo más sencillo que se consideró para analizar el proceso de desplegamiento del dominio PDZ3-h-PSD95 incluye la presencia de un estado intermedio monomérico, capaz además de presentar un equilibrio de asociación. Dicho esquema podría resumirse de la siguiente forma:



Según el esquema hay tres procesos de equilibrio. Por un lado, está el paso del estado nativo al intermedio, $N \rightleftharpoons I$, al que sigue el desplegamiento completo del dominio, $I \rightleftharpoons D$. Por otro lado, está el equilibrio de asociación, $I \rightleftharpoons (1/n) I_n$. Al considerar de esta forma los equilibrios, los valores termodinámicos obtenidos de los análisis son comparables con los del modelo más sencillo, puesto que se describe el proceso de desplegamiento del estado nativo al desplegado considerando la presencia de un estado intermedio, así como el proceso de asociación relacionado con dicho intermedio (para ver los detalles del modelo ver la Sección 3.4.2.3).

Los datos experimentales convergen muy bien con los resultados del análisis aplicando dicho modelo (Figura 7.6). Tal y como se realizó con las trazas obtenidas del dominio PDZ3-PSD95, para los análisis se consideraron comunes a todas las curvas los valores de las capacidades caloríficas correspondientes a cada estado conformacional así como los parámetros termodinámicos relacionados con el proceso de disociación. Sin embargo, los valores de las entalpías y temperaturas asociadas a las transiciones del desplegamiento de la proteína ($N \rightleftharpoons I \rightleftharpoons D$), se obtuvieron de forma individual. El intervalo de error considerado para estos parámetros fue entonces el asociado a la desviación estándar media de los parámetros obtenidos individualmente. Para conocer la estequiometría de la asociación, se dejó como valor flotante n , el índice de asociación del intermedio, y se obtuvo un valor de 2.8, que indica, como en el caso del dominio PDZ3-PSD95, que la asociación del intermedio correspondería globalmente a un trímero. De hecho, los análisis mostrados en la Figura 7.6, cuyos resultados están recogidos en la Tabla 7.2, se han obtenido para un valor de $n = 3$.

Según el análisis de las poblaciones se puede observar que el intermedio oligomérico se puebla en una alta proporción (panel A de la Figura 7.7). De hecho, a baja concentración de proteína también se puebla de forma mayoritaria, hasta un 90%, en el intervalo de temperaturas comprendido entre los 50°C y unos 85°C. Esto refleja una diferencia importante con respecto al dominio PDZ3-PSD95, en cuyo caso a baja concentración de proteína el intermedio se puebla menos, llegando al 50%, y en el que además el intervalo de temperaturas en el que está presente es bastante menor, sólo desde 65°C hasta 85°C (panel B de la Figura 7.7). Por lo tanto, al truncar la tercera hélice α del dominio, el intermedio se estabiliza, puesto que aparece a menores concentraciones de proteína y se puebla en mayor proporción.

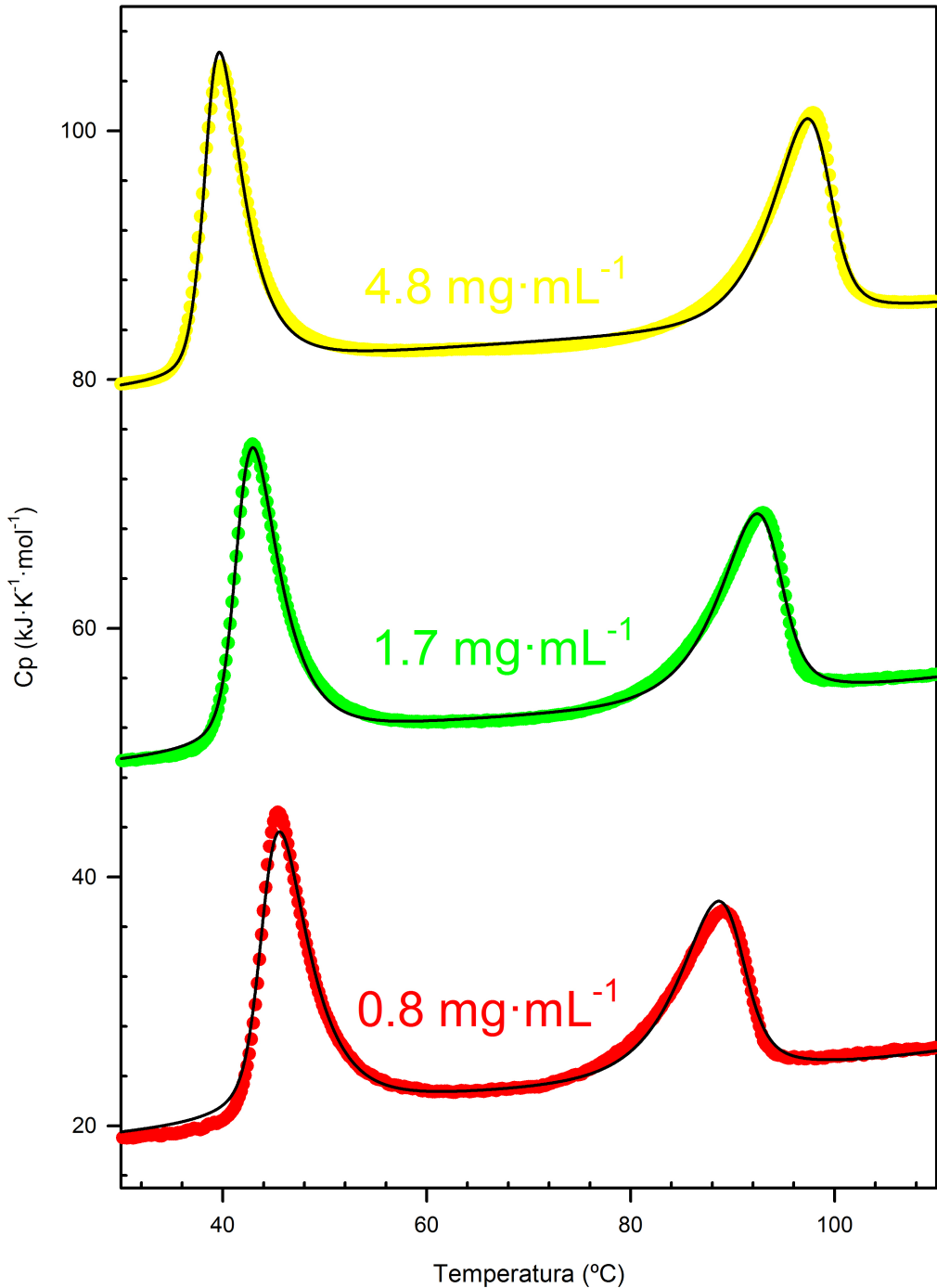


Figura 7.6. Curvas de desplegamiento térmico del dominio PDZ3-h-PSD95 en fosfato 50 mM pH 7.5 obtenidas mediante DSC a varias concentraciones de proteína, indicadas sobre cada una de las trazas. Los puntos experimentales se corresponden con los círculos, mientras que el análisis global con el modelo que se corresponde con el [Esquema 3.3](#), se representa mediante líneas de color negro.

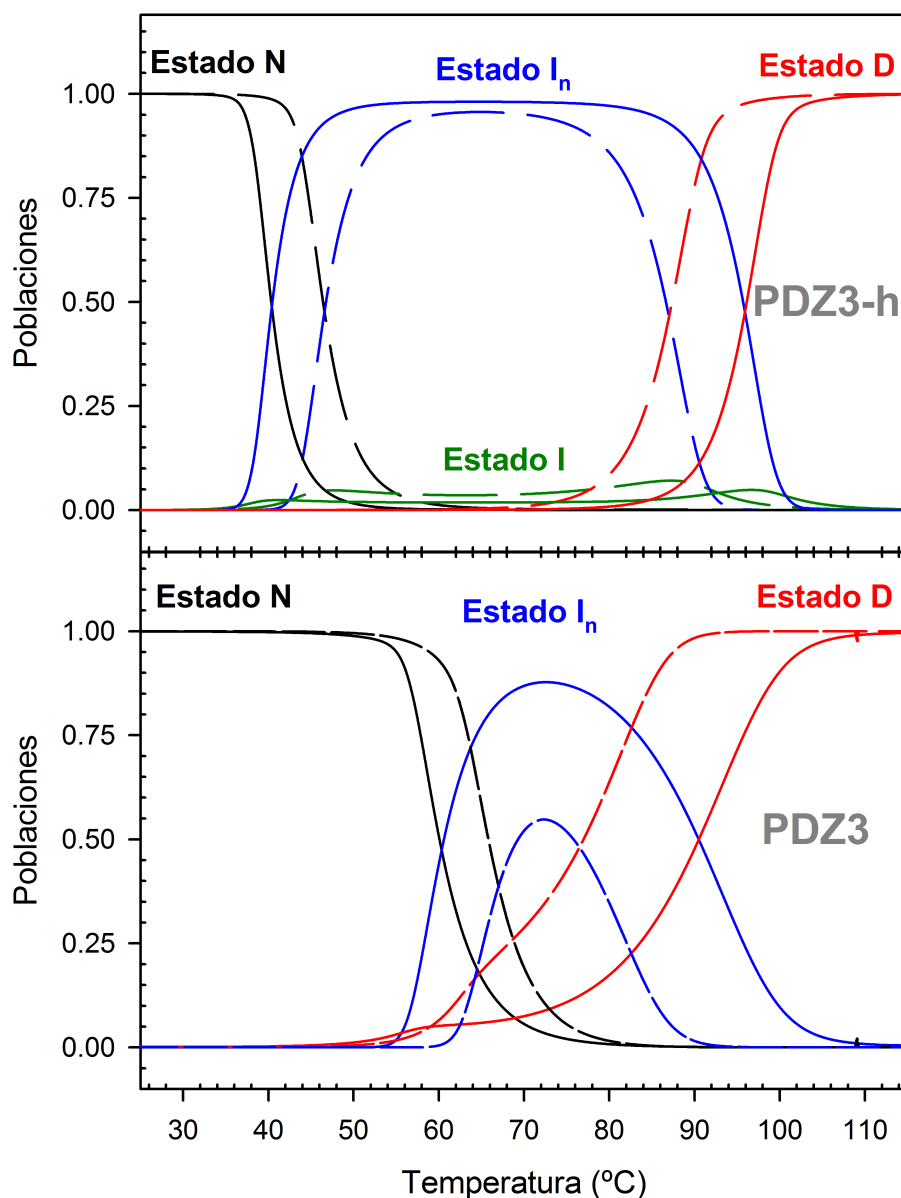


Figura 7.7. Distribución de las poblaciones de los diferentes estados en el desplegamiento térmico obtenidos mediante el análisis de las trazas calorimétricas. Panel indicado con PDZ3-h: Distribución de las poblaciones de los diferentes estados en el desplegamiento térmico del dominio PDZ3-h-PSD95 para las dos concentraciones de proteína usadas, en línea continua la de $0.8 \text{ mg}\cdot\text{mL}^{-1}$ y en línea discontinua la de $4.8 \text{ mg}\cdot\text{mL}^{-1}$. Panel indicado con PDZ3: Distribución de las poblaciones de los diferentes estados en el desplegamiento térmico del dominio PDZ3-PSD95 para las dos concentraciones de proteína iguales en molaridad que el dominio PDZ3-h-PSD95, en línea continua la de $1 \text{ mg}\cdot\text{mL}^{-1}$ (unos $85 \mu\text{M}$ para ambas proteínas) y en línea discontinua la de $5.3 \text{ mg}\cdot\text{mL}^{-1}$ (unos $480 \mu\text{M}$ para ambas proteínas). En color negro se representa la población del estado nativo (N), en rojo el desnaturalizado (D), en azul el estado intermedio asociado (I_n) y en verde el estado intermedio monomérico (I), representado únicamente para el caso del dominio PDZ3-h-PSD95.

Tabla 7.2. *Parámetros termodinámicos del desplegamiento térmico del dominio PDZ3-h-PSD95 en tampón fosfato 50 mM a pH 7.5 obtenidos del análisis de experimentos de DSC con el modelo de cuatro estados en que aparece un intermedio que en otro paso se asocia.*

T_{N-I} (°C)	$\Delta H_{N-I}(T_{N-I})$ (kJ·mol ⁻¹)	T_{I-D} (°C)	$\Delta H_{I-D}(T_{I-D})$ (kJ·mol ⁻¹)	T_{I-In} (°C)	$\Delta H_{I-In}(T_{I-In})$ (kJ·mol ⁻¹)	$\Delta G_{N-D}(298)$ (kJ·mol ⁻¹)	R^2
52 ± 3	235 ± 20	81 ± 3	204 ± 10	107 ± 2	-157 ± 12	35 ± 5	0.993

Los intervalos de error se han considerado como el valor asociado a la desviación estándar media de los parámetros obtenidos individualmente. Los valores de las magnitudes termodinámicas del estudio han sido estimados con una $P_{ref} = 100 \mu\text{M}$. Las funciones de capacidad calorífica obtenidas del ajuste fueron $C_{pN} = -9.72 + (0.096 \cdot T)$; $C_{pD} = -19.1 + (0.117 \cdot T)$; $C_{pIn} = -30.4 + (0.131 \cdot T)$; $C_{pI} = -31.9 + (0.146 \cdot T)$ en kJ·mol⁻¹.

Para estudiar con más detalle la distribución de especies observadas en los análisis de las trazas calorimétricas, se realizaron experimentos adicionales de DLS en los que se realizó un barrido de temperatura desde 20°C hasta 60°C a una velocidad de barrido de 1 K·min⁻¹ en fosfato 50 mM pH 7.5, las mismas condiciones que el experimento análogo realizado para el dominio PDZ3-PSD95. La concentración de PDZ3-h-PSD95 fue de 7.2 mg·mL⁻¹, la misma en molaridad, 720 μM , que la que se usó con el dominio PDZ3-PSD95, 8 mg·mL⁻¹. Durante el barrido se produce asociación tal y como indican los resultados de DSC. Así, a 20°C hay especies con un radio hidrodinámico de 1.8 nm, que corresponde a un monómero, según la calibración con proteínas de pesos moleculares conocidos que está incluida en el programa que controla el instrumento (Dynamics V6, ver Sección 3.6), con la que se puede hacer una estimación del peso molecular de la partícula en función de su radio hidrodinámico. Entre 34°C y 38°C hay especies de 1.8 nm y 6 nm, que según dicha calibración correspondería a un peso molecular de unos veinte monómeros, aunque si la disposición de la partícula no se asemeja a una proteína globular, cosa que desconocemos, no podemos decir con fiabilidad que la partícula contiene veinte monómeros de forma exacta. A partir de 40°C desaparece la especie de 1.8 nm y la especie de 6 nm co-existe junto con una nueva de 24 nm, de tamaño muy superior, que, al ser tan grande, es poco probable que su forma se asemeje a una proteína globular, por lo que no se puede estimar de forma fiable su peso molecular.

De esta forma, de acuerdo con los resultados de los análisis de DSC, se confirma la existencia de un estado monomérico junto con las especies asociadas hasta los 40°C (Figura 7.8). Sin embargo, en el caso de la construcción PDZ3-h-PSD95 no se observa la especie de 2.6 nm observada en el caso del dominio PDZ3-PSD95 y que corresponde al peso molecular de un trímero. Las especies oligoméricas de la construcción sin la hélice $\alpha 3$ son de 6 nm, tamaño considerablemente mayor que las observadas en el caso del dominio PDZ3-PSD95.

A partir de la presencia de estas especies de mayor índice de oligomerización observadas por DLS, se replanteó el análisis de las trazas calorimétricas variando la estequiometría de asociación hasta valores de $n = 20$. Los resultados del análisis no reproducen los datos experimentales, lo que nos hace pensar que el proceso de asociación podría ser aún más complejo que el esquema sugerido, en el que haya incluidos otros

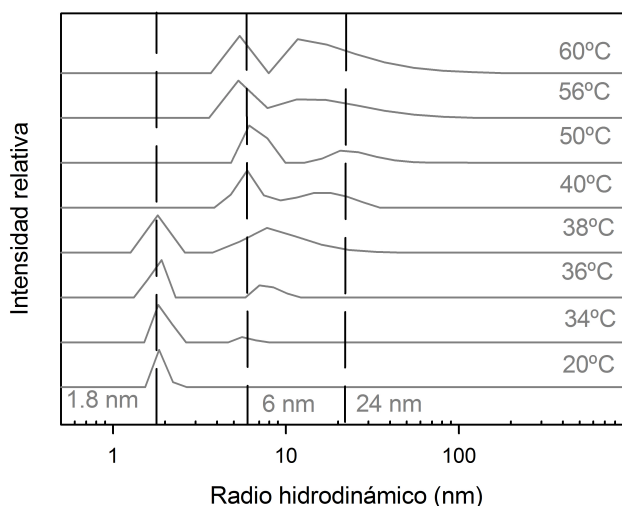


Figura 7.8. Barrido de temperatura desde 20°C hasta 60°C del dominio PDZ3-h-PSD95 a una concentración de $7.2 \text{ mg}\cdot\text{mL}^{-1}$ en fosfato 50 mM pH 7.5 seguido por DLS. La velocidad de barrido fue de $1 \text{ K}\cdot\text{min}^{-1}$.

equilibrios adicionales, como por ejemplo la inter-asociación de los trímeros. En todo caso, las evidencias experimentales sí muestran claramente que en el desplegamiento del dominio PDZ3-h-PSD95 hay un intermedio que se asocia, existiendo además una estabilización de este intermedio con respecto al de la construcción PDZ3-PSD95.

Esta estabilización del intermedio y su correspondiente asociación se confirma mediante otro dato experimental de DLS. Ya se explicó en la Sección 5.1.3 que cuando la concentración de PDZ3-PSD95 usada es mínima (tanto para DSC como para DLS, unos $0.4 \text{ mg}\cdot\text{mL}^{-1}$), no se detecta ningún proceso de asociación. En cambio, tal asociación sí se produce con el dominio PDZ3-h-PSD95. Así, cuando se realiza un barrido de temperatura seguido por DLS con el dominio PDZ3-h-PSD95 a unos $0.4 \text{ mg}\cdot\text{mL}^{-1}$ desde 20°C hasta 50°C únicamente existe una especie monomérica, y a partir de 50°C se observa de nuevo una especie a 6 nm. A 54°C desaparece la especie monomérica y no se observan partículas de tamaños mayores (Figura 7.9). Esto está de acuerdo con los análisis de DSC, según los cuales también se puebla el intermedio a baja concentración. De hecho, los resultados de DLS muestran que el proceso de asociación comienza hacia los 50 °C en el caso de una concentración baja de proteína, mientras que a alta concentración se da a temperaturas menores, unos 38 °C, lo cual coincide de nuevo con las poblaciones obtenidas del análisis de los datos calorimétricos (Figura 7.7).

Para intentar conseguir información conformacional adicional del mecanismo de desplegamiento, se realizaron experimentos de CD. Así, en primer lugar, se realizó un barrido de temperatura desde 2°C hasta 98°C a una velocidad de barrido de $1.5 \text{ K}\cdot\text{min}^{-1}$, seguido por CD a una longitud de onda de 218 nm en las mismas condiciones que el previo realizado con el dominio PDZ3-PSD95, en fosfato 50 mM pH 7.5. La concentración de PDZ3-h-PSD95 usada fue de $0.3 \text{ mg}\cdot\text{mL}^{-1}$, el mismo valor en molaridad, $30 \text{ }\mu\text{M}$, que la usada para el dominio PDZ3-PSD95. Estos experimentos muestran que el comportamiento

es significativamente diferente cuando se trunca la hélice (Figura 7.10). Tal y como se describe en la Sección 5.1.1, el dominio PDZ3-PSD95 sigue un comportamiento sigmoide, en cambio, la nueva construcción ya no tiene este comportamiento. Además, no es fácil analizar la curva para saber qué fenómenos están ocurriendo en una primera aproximación, sin más datos.

Se registraron espectros de CD desde 10°C hasta 98°C en intervalos de 10°C en fosfato 50 mM pH 7.5 y una concentración de PDZ3-h-PSD95 de 0.3 mg·mL⁻¹. En este caso, se distinguen claramente tres tipos de espectros (Figura 7.11). Por un lado aquellos obtenidos entre 10 y 40°C, que corresponden a la especie nativa. Por otro, los espectros registrados entre 50 y 70°C, en los que presenta un mínimo hacia 217 nm, lo que indica claramente un aumento de conformación en lámina β. Finalmente, están los espectros registrados a más de 90°C, correspondientes al estado desnaturalizado. Todos estos resultados explican el barrido de temperatura seguido por CD mostrado en la Figura 7.10, registrado a una longitud de onda de 218 nm. Se observa que hacia los 50°C hay un aumento de la señal que corresponde al aumento de conformación de lámina β, para luego hacia los 80°C dar otra transición en la que se va perdiendo dicha conformación. Esto concuerda con los resultados obtenidos de la calorimetría, en los que la temperatura de la transición entre el estado nativo y el estado intermedio es de unos 50°C y la correspondiente entre el estado intermedio y el estado desplegado, de unos 80°C (ver Tabla 7.2).

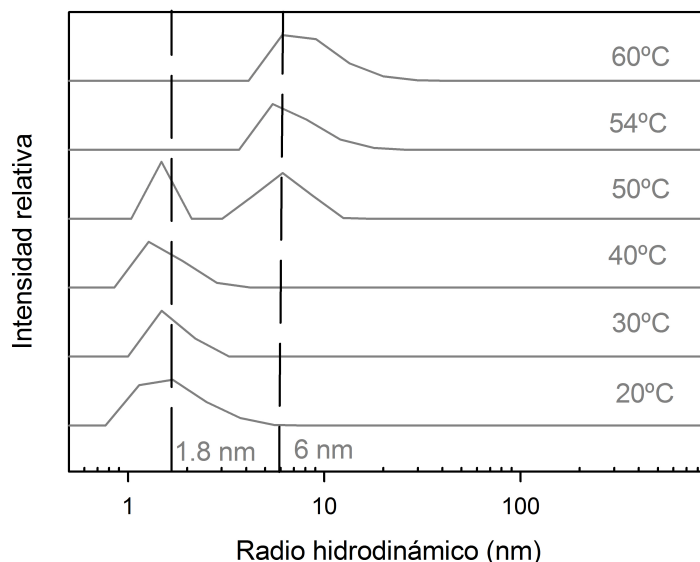


Figura 7.9. Barrido de temperatura desde 20°C hasta 60°C del dominio PDZ3-h-PSD95 a una concentración de 0.4 mg·mL⁻¹ en fosfato 50 mM pH 7.5 seguido por DLS. La velocidad de barrido fue de 1 K·min⁻¹.

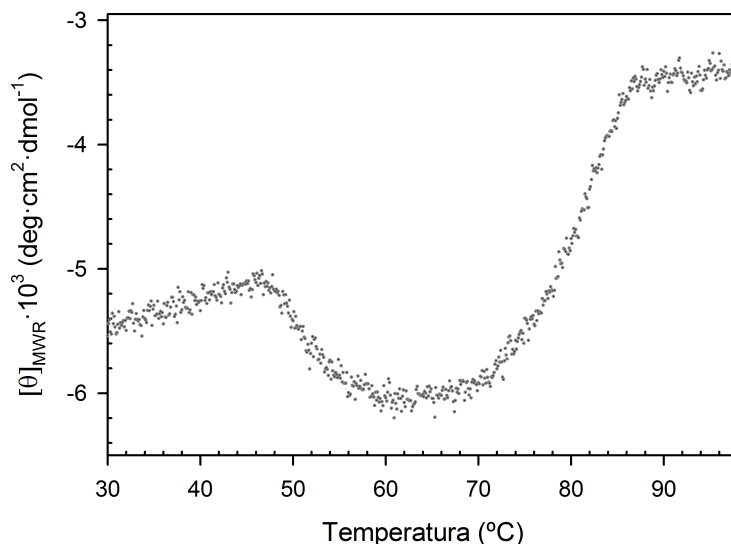


Figura 7.10. Curva de desplegamiento térmico seguida por CD a 218 nm del dominio PDZ3-h-PSD95 a una concentración de $0.3 \text{ mg} \cdot \text{mL}^{-1}$ en fosfato 50 mM pH 7.5.

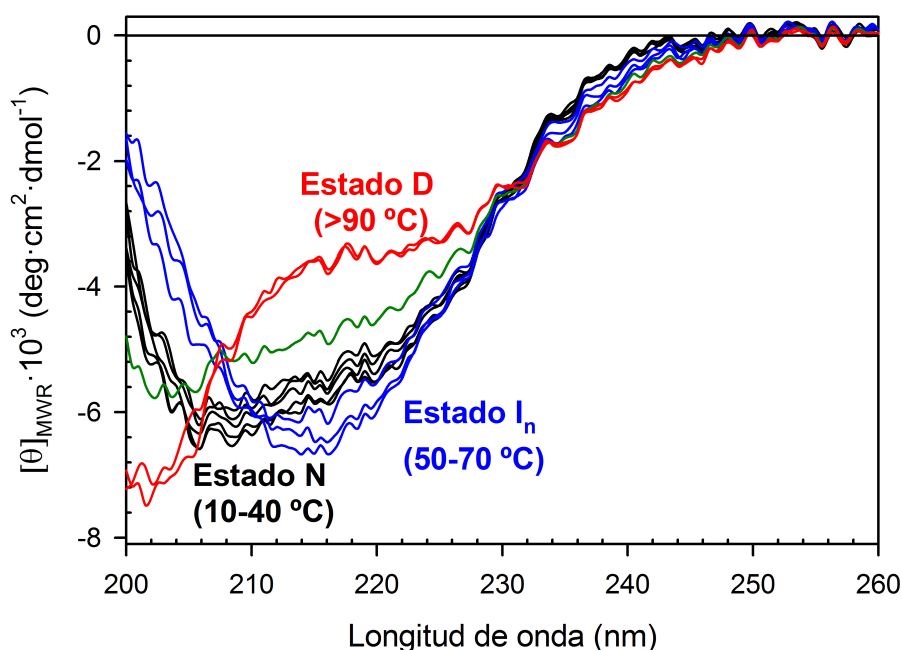


Figura 7.11. Espectros de CD del dominio PDZ3-h-PSD95 a $0.3 \text{ mg} \cdot \text{mL}^{-1}$ en tampón fosfato 50 mM pH 7.5 registrados desde 10°C a 98°C en un intervalo de 10°C . En el intervalo de temperaturas de 10°C a 40°C los espectros están representados en color negro (Estado N), de 50°C a 70°C en color azul (Estado I_n) y a más de 90°C en color rojo (Estado D). En color verde se representa el espectro a 80°C , temperatura que se corresponde con la transición del estado intermedio al desplegado.

7.1.3 Mecanismo de fibrilación del dominio PDZ3-h-PSD95

Para averiguar si el estado intermedio evoluciona a estructuras fibrilares como ocurre en el caso del dominio PDZ3-PSD95, se dejó incubando la muestra a 60°C en las mismas condiciones de tampón, fosfato 50 mM pH 7.5, y a alta concentración de proteína (7.2 mg·mL⁻¹) para estudiar las especies que se pueblan a lo largo del tiempo mediante DLS. En consonancia con los resultados del barrido de temperatura realizados con esta misma técnica, hay una población de partículas con tamaños de 6 nm y otro tipo de unos 35-40 nm (Figura 7.12). Estas especies van aumentando de tamaño a lo largo del tiempo (Tabla 7.3).

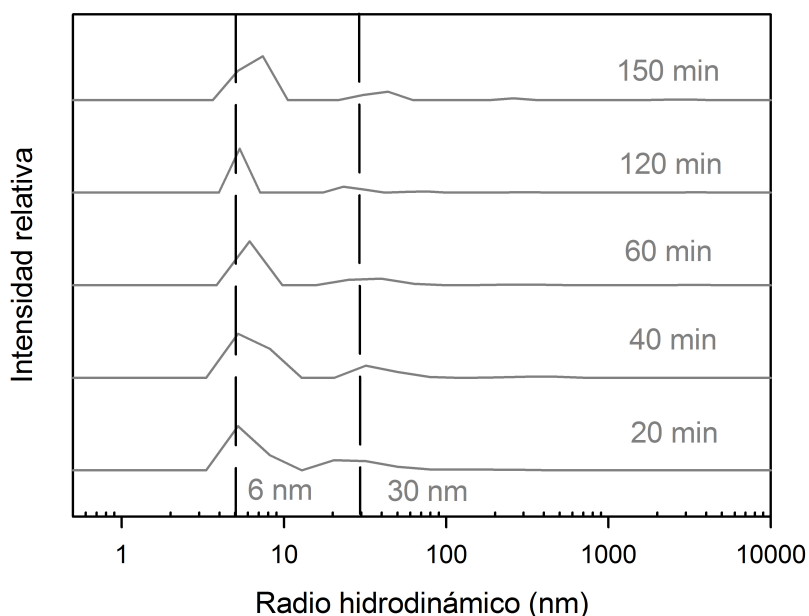


Figura 7.12. Evolución del radio hidrodinámico en función del tiempo seguida por DLS para las especies presentes en una disolución del dominio PDZ3-h-PSD95 a una concentración de 7.2 mg·mL⁻¹ en fosfato 50 mM pH 7.5.

Asimismo, se incubaron a 60°C durante distintos intervalos de tiempo, desde un día hasta cinco, varias muestras de PDZ3-h-PSD95 dializadas en fosfato 50 mM pH 7.5 con una concentración final de 7.2 mg·mL⁻¹, que es la misma que se usó para seguir la evolución de las especies a 60°C en disolución mediante DLS, para su posterior análisis mediante TEM. Las micrografías muestran en este caso únicamente fibras, que aparecen además desde un día de incubación (Figura 7.13). Asimismo, la preparación de una muestra en las mismas condiciones, pero con una incubación de un mes a 37°C, resultó de nuevo en la formación de fibras (micrografía señalada con “1 mes 37°C” de la Figura 7.13).

Tabla 7.3 Poblaciones de las distintas especies detectadas por DLS que presenta el dominio PDZ3-h-PSD95 cuando se incubaba a 60°C durante dos horas y media en tampón fosfato 50 mM pH 7.5.

Tiempo (horas)	Radio (nm)	% Masa	Peso Molecular (kDa)
20 min	6	70.2	220
	42.7	29.8	22000
40 min	6.4	77.8	259
	39.7	20	18485
	364	2.2	3308160
60 min	6.1	74.4	234
	36.2	22.5	14928
	297.3	2.7	2060120
	3708.5	0.4	$7.56 \cdot 10^8$
120 min	5.3	76.8	169
	26.1	16.2	6941
	70	3.9	69962
	302.1	1.4	2139820
	3556.9	1.7	$6.85 \cdot 10^8$
150 min	6.5	80.5	270
	38.8	15.1	17601
	259.2	2.1	1495310
	2721.9	1.2	$3.67 \cdot 10^8$
	53728.3	1	$3.93 \cdot 10^{11}$

Todos estos resultados muestran claras diferencias con los obtenidos con el dominio PDZ3-PSD95. Tal y como se ha descrito en el Capítulo 5, en el caso de dicho dominio se observan especies de 2.6 nm, así como partículas de 12 nm y a continuación especies de tamaños mayores (Figura 5.8). En el microscopio se ven en varios días unas estructuras anulares y en otras ocasiones protofibras, para dar lugar finalmente a fibras amiloides en tiempos de incubación más largos, de al menos una semana. Además, al incubar a 37°C sólo se observaban protofibras, mientras que en el caso del dominio PDZ3-h-PSD95 se observan fibras maduras. Por lo tanto, se ve claramente que al truncar la tercera hélice α la fibrilación se produce mucho más rápidamente, además de que se impide la formación de las estructuras anulares. Por otro lado, las fibras obtenidas a partir del dominio PDZ3-h-PSD95 siguen presentes al poner la muestra a temperatura ambiente y diluir con tampón, en contraposición con la reversibilidad que ante este hecho presenta la construcción que tiene la hélice. Estos resultados están de acuerdo con la mayor estabilidad del intermedio de plegamiento, precursor de estas estructuras supramacromoleculares.

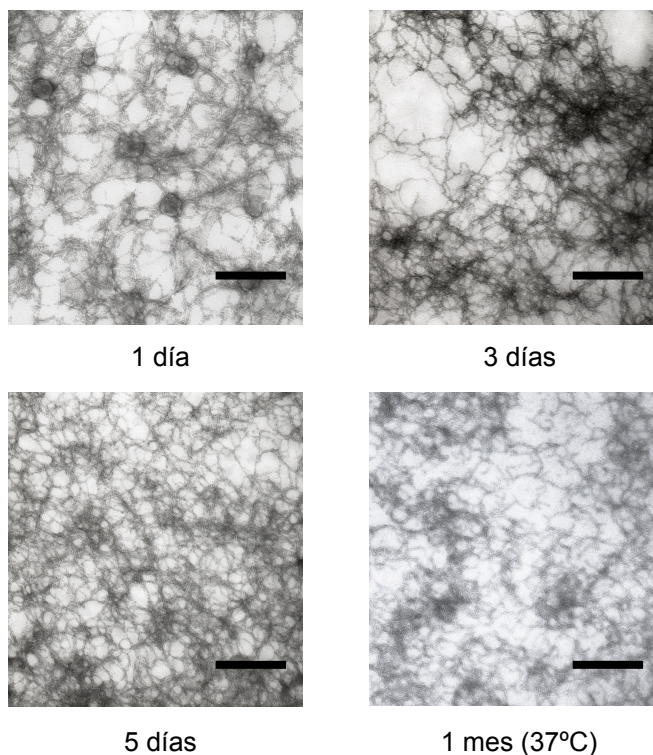


Figura 7.13. Análisis de las estructuras supramacromoleculares del dominio PDZ3-h-PSD95 seguidas por TEM. En el pie de cada micrografía está indicado el tiempo de incubación de la disolución de proteína a $7.2 \text{ mg}\cdot\text{mL}^{-1}$ en fosfato 50 mM pH 7.5 a 60°C , excepto en la micrografía que está indicada con “1 mes (37°C)”, que fue tomada después de 1 mes de incubación a 37°C . La barra de escala corresponde a un tamaño de 200 nm.

De hecho, este carácter más irreversible y la mayor rapidez en la formación de especies de tamaños mayores que presenta el dominio PDZ3-h-PSD95 se observó también mediante experimentos seguidos por SEC. Así, se dializó el dominio PDZ3-h-PSD95 en fosfato 50 mM pH 7.5 y, en una concentración final de $7.2 \text{ mg}\cdot\text{mL}^{-1}$, y se inyectó en la columna de exclusión molecular. Al realizar la cromatografía a 25°C se observó únicamente un máximo (Figura 7.14), cuyo volumen de elución corresponde a un peso molecular del monómero, según la calibración de la columna que se realizó con patrones de proteína de pesos moleculares conocidos (ver Sección 3.10). En cambio, únicamente con el tiempo de incubación de la muestra a 60°C dentro de la columna de exclusión, es decir, sin incubación previa a 60°C , prácticamente la única especie que se observa eluye en el volumen de exclusión de la columna, correspondiente a tamaños muy grandes de partículas. Ya se ha visto que en el caso del dominio PDZ3-PSD95 hace falta más tiempo de incubación para ver especies distintas al monómero, así como que se observan más especies cuyos tamaños corresponden a la posible presencia de un equilibrio entre monómeros y oligómeros de distinto índice de asociación.

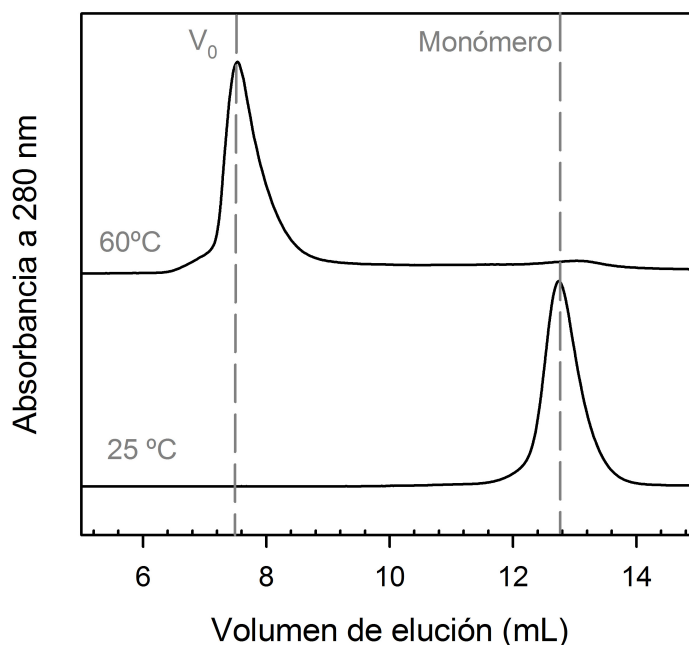


Figura 7.14. Perfil de eluciones del dominio PDZ3-h-PSD95 a una concentración de $7.2 \text{ mg}\cdot\text{mL}^{-1}$ en fosfato 50 mM pH 7.5 obtenido después de la inyección de la muestra en la columna Superdex 75 HR 10/30 termostatizada a 25°C y a 60°C . Las líneas verticales corresponden al volumen de elución de la especie monomérica del dominio PDZ3-h-PSD95, y al volumen de exclusión de la columna (V_0).

7.1.4 Cinética de la formación de fibras en el dominio PDZ3-h-PSD95

Para estudiar en mayor profundidad el mecanismo de formación de fibras en el dominio PDZ3-h-PSD95, se llevó a cabo un barrido de temperatura hasta 60°C para seguir la fluorescencia de ThT y ANS. Las condiciones fueron las mismas que las que se usaron para la construcción que incluye la hélice α_3 , fosfato 50 mM pH 7.5 y a concentraciones de ThT de $12.5 \text{ }\mu\text{M}$ y de ANS de $20 \text{ }\mu\text{M}$. Se realizaron los experimentos a las dos concentraciones de proteína extremas estudiadas (0.4 y $7.2 \text{ mg}\cdot\text{mL}^{-1}$). La fluorescencia de ambas sondas aumenta ligeramente hasta unos 40°C , temperatura a partir de la cual la señal disminuye en ambas sondas (Figura 7.15). Esto está de acuerdo con la temperatura a la que se da el aumento de conformación en lámina β observado por CD (Figura 7.11) puesto que este cambio estructural implica una mayor unión de ThT por su capacidad de unión a agregados tipo β . El cambio en la fluorescencia de ANS implica que en ese cambio estructural se expone una mayor superficie hidrofóbica al disolvente. El comportamiento a $0.4 \text{ mg}\cdot\text{mL}^{-1}$ es similar al que se da a alta concentración de proteína, con dos diferencias relevantes. Por un lado, la señal que dan ambas sondas fluorescentes es significativamente menor. Por otro, la temperatura hasta la cual aumenta la intensidad de fluorescencia de ambas sondas es mayor que a alta concentración, unos 50°C , temperatura que coincide de nuevo con los resultados del análisis de poblaciones obtenidos a partir de las trazas de DSC, así como los resultados de DLS.

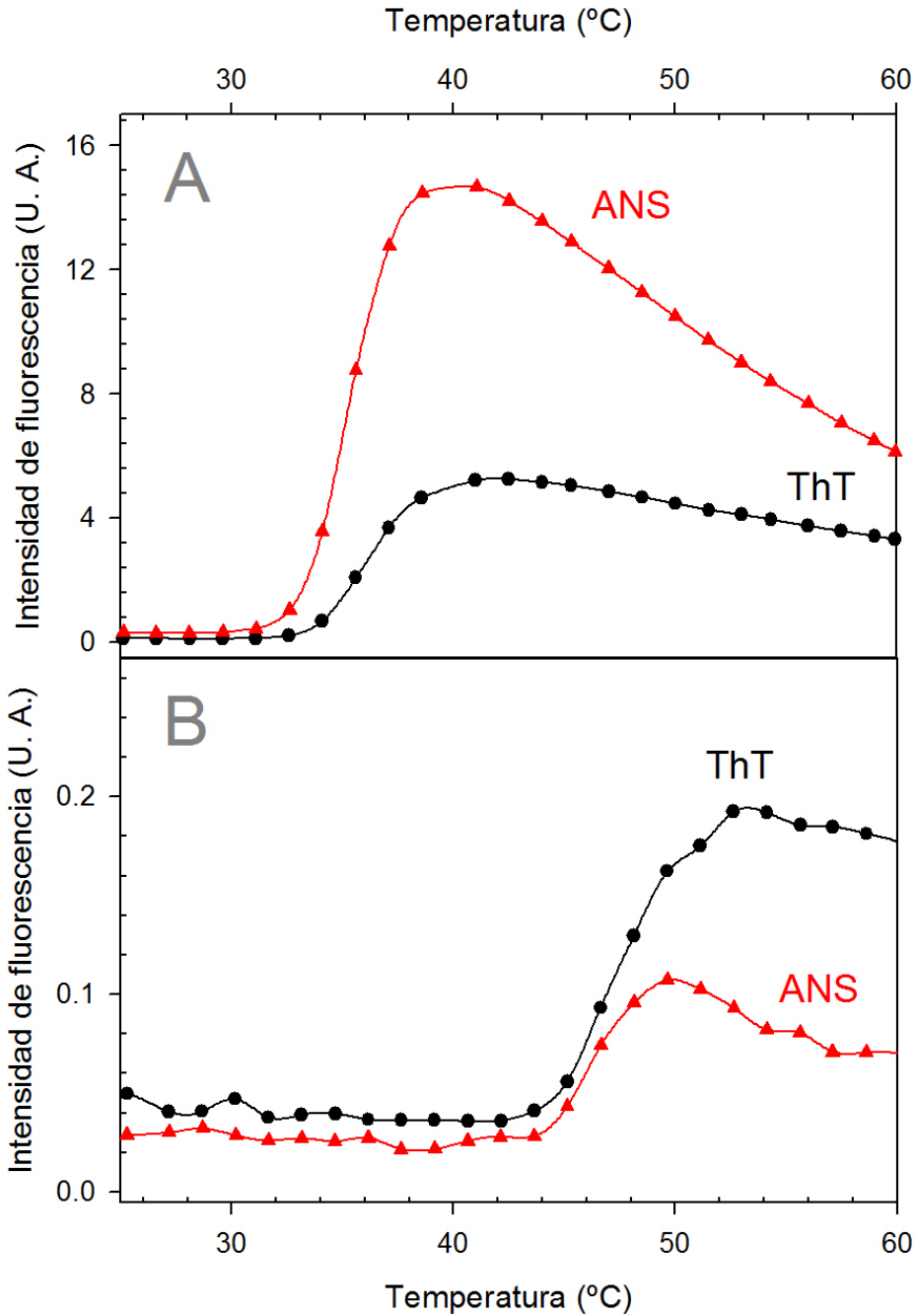


Figura 7.15. Barrido de temperatura del dominio PDZ3-h-PSD95 hasta 60°C seguido por la emisión de fluorescencia de ThT (negro) y ANS (rojo). (A) a una concentración de 7.2 mg·mL⁻¹ en fosfato 50 mM pH 7.5 en la presencia de 12.5 μM de ThT o 20 μM de ANS y (B) en las mismas condiciones experimentales pero a una concentración de proteína de 0.4 mg·mL⁻¹.

Cuando se dejaron incubando a 60°C esas mismas muestras, se observó un aumento de fluorescencia muy acusado en función del tiempo (Figura 7.16), tal y como sucede en el caso del experimento paralelo del dominio PDZ3-PSD95. Sin embargo, hay una diferencia en los tiempos a los que la fluorescencia alcanza un máximo, a unos 30 minutos en el caso de ThT y a unos 350 minutos en el de ANS para el dominio PDZ3-h-PSD95, mientras que en el caso del dominio PDZ3-PSD95 son 100 minutos y 400 minutos, respectivamente (Sección 5.1.4). Después de estos tiempos, la señal comienza a decaer en el caso de los experimentos con ambos dominios. Ya se explicó que este fenómeno puede ocurrir cuando hay un aumento en la dispersión de luz que se debe al aumento de la viscosidad de la disolución puesto que las estructuras supramacromoleculares que se están formando tienen menor solubilidad (Hawe et al., 2008; Kad et al., 2003; Kardos et al., 2004).

Por otro lado, a diferencia de lo que ocurría con el dominio PDZ3-PSD95, todo este proceso ocurre paralelamente a las dos concentraciones estudiadas. Tal y como se observa en la Figura 7.17, también se produce un aumento abrupto de fluorescencia en ambas sondas fluorescentes cuando se incubó PDZ3-h-PSD95 en fosfato 50 mM pH 7.5 con ThT a 12.5 μM o ANS a 20 μM y a una concentración baja de proteína, de 0.4 $\text{mg}\cdot\text{mL}^{-1}$, aunque el tiempo al que ocurre es diferente, a unos diez minutos en ambas sondas fluorescentes. Que de nuevo se produzca un paralelismo en el comportamiento a concentraciones altas y bajas de proteína está de acuerdo una vez más con la mencionada estabilización del intermedio que se produce al truncar la tercera hélice α . Finalmente, se sigue reafirmando el carácter irreversible que tienen los procesos en el dominio PDZ3-h-PSD95, puesto que al bajar a temperatura ambiente la fluorescencia de ambas sondas permanece casi inalterada (Figuras 7.18 y 7.19).

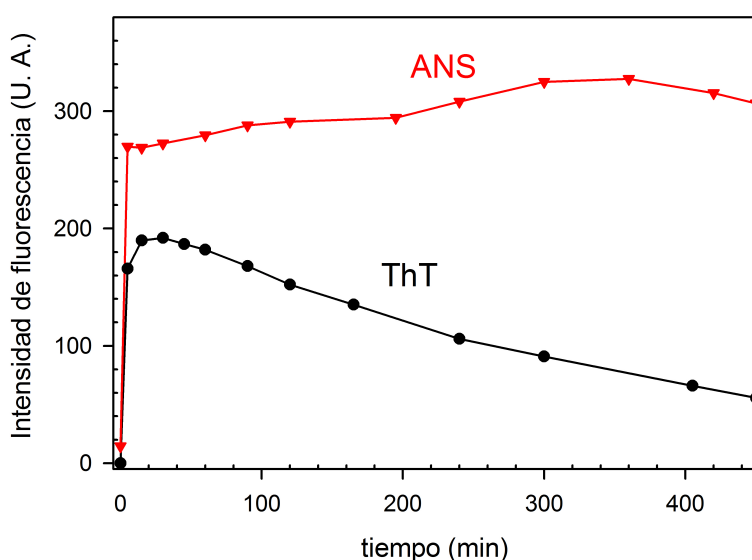


Figura 7.16. Cinéticas de crecimiento de la emisión de fluorescencia de ThT (negro) y ANS (rojo). Se incubó una muestra de PDZ3-h-PSD95 a 7.2 $\text{mg}\cdot\text{mL}^{-1}$ en la presencia de 12.5 μM de ThT o 20 μM de ANS en fosfato 50 mM pH 7.5 a 60°C.

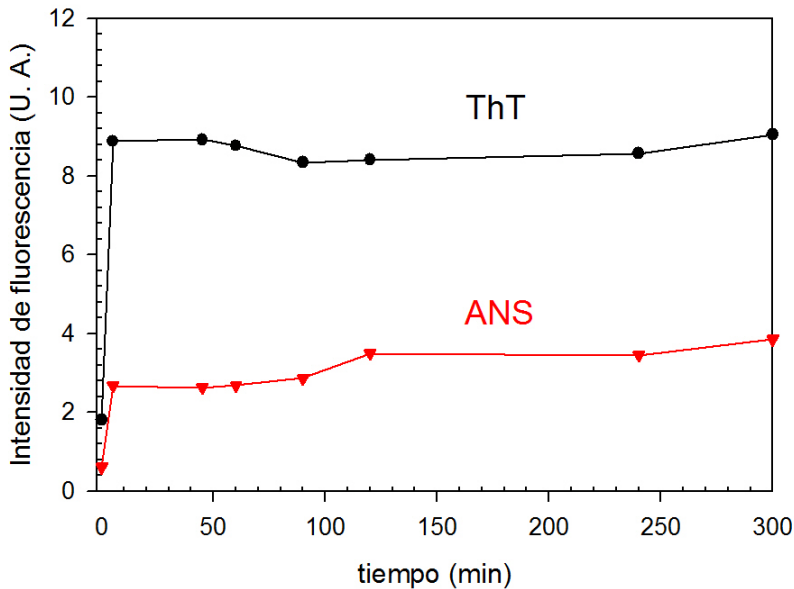


Figura 7.17. Cinéticas de crecimiento de la emisión de fluorescencia de ThT (negro) y ANS (rojo). Se incubó una muestra de PDZ3-h-PSD95 a $0.4 \text{ mg}\cdot\text{mL}^{-1}$ en la presencia de $12.5 \mu\text{M}$ de ThT o $20 \mu\text{M}$ de ANS en fosfato 50 mM pH 7.5 a 60°C .

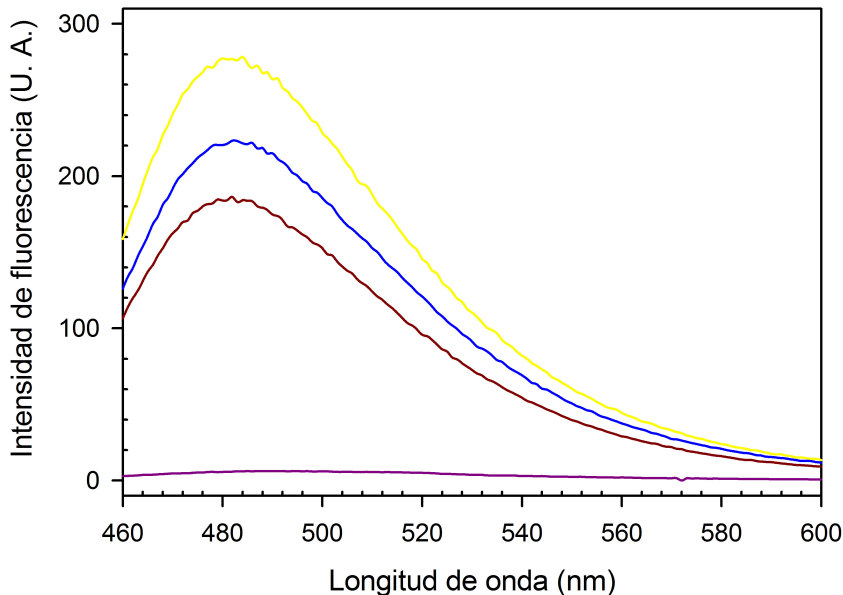


Figura 7.18. Espectros de emisión de fluorescencia de ThT una vez enfriada la disolución de PDZ3-h-PSD95 a $7.2 \text{ mg}\cdot\text{mL}^{-1}$ y pH 7.5 a 25°C después de haberla calentado a 60°C durante unas horas. Los espectros se registraron a distintos tiempos una vez que se alcanzan los 25°C : 1 minuto (amarillo), 1 hora (azul) y tres días (marrón). Se ha incluido también el espectro a 25°C antes de calentar, en color morado.

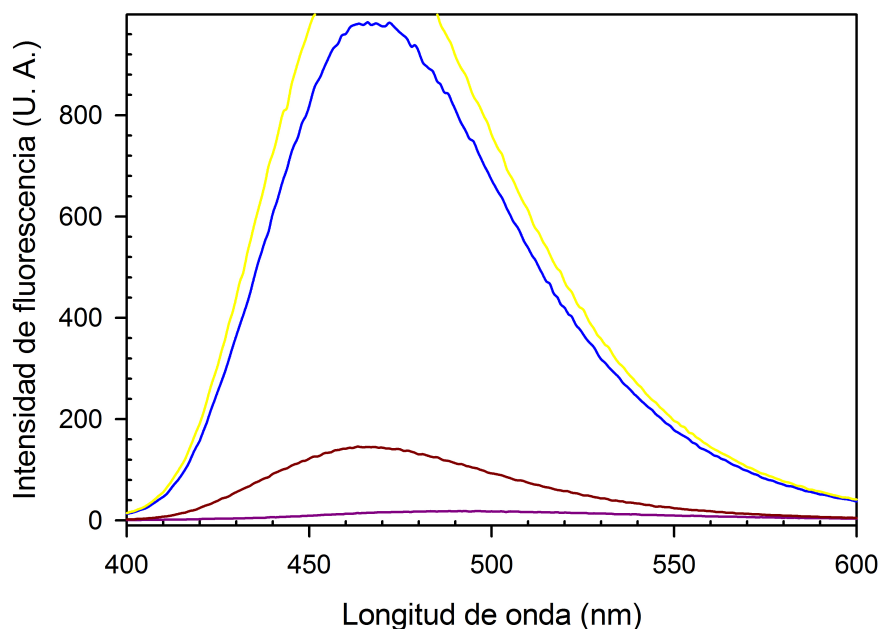


Figura 7.19. Espectros de emisión de fluorescencia de ANS una vez enfría la disolución de PDZ3-h-PSD95 a $7.2 \text{ mg}\cdot\text{mL}^{-1}$ y pH 7.5 a 25°C después de haberla calentado a 60°C durante unas horas. Los espectros se registraron a distintos tiempos una vez que se alcanzan los 25°C : 1 minuto (amarillo), 1 hora (azul) y tres días (marrón). Se ha incluido también el espectro a 25°C antes de calentar, en color morado.

7.1.5 Influencia del pH en el intermedio de plegamiento del dominio PDZ3-h-PSD95

Puesto que en el caso del dominio PDZ3-PSD95 cambia el equilibrio conformacional en función del pH, se estudió si se daba el mismo efecto en la construcción sin su hélice $\alpha 3$. Se realizaron experimentos de DSC con el dominio PDZ3-h-PSD95 a $1.3 \text{ mg}\cdot\text{mL}^{-1}$ en tampón 50 mM a distintos valores de pH. Para tamponar a pH 2.5, 3 y 3.5 se usó glicina y para pH 4, acetato. Se observó que de nuevo el comportamiento experimental varía en función del pH (Figura 7.20). A simple vista, la traza a pH 4 se parece a las trazas a pH 7.5 del dominio PDZ3-h-PSD95, puesto que las curvas son más agudas. En cambio, las trazas de pH 3 y 3.5 se parecen más a las de pH 7.5 del dominio PDZ3-PSD95, puesto que sus transiciones con más redondeadas. Por otro lado, la curva obtenida a pH 2.5 podría reflejar un comportamiento de dos estados ($N \rightleftharpoons D$), aunque el pico es un poco más ancho y menos simétrico que el que se da en estos casos. De hecho, esta curva obtenida a pH 2.5 recuerda a la traza a alta concentración del dominio PDZ3-PSD95 a pH 3, que no se ajusta bien al modelo de dos estados (Figura 6.2).

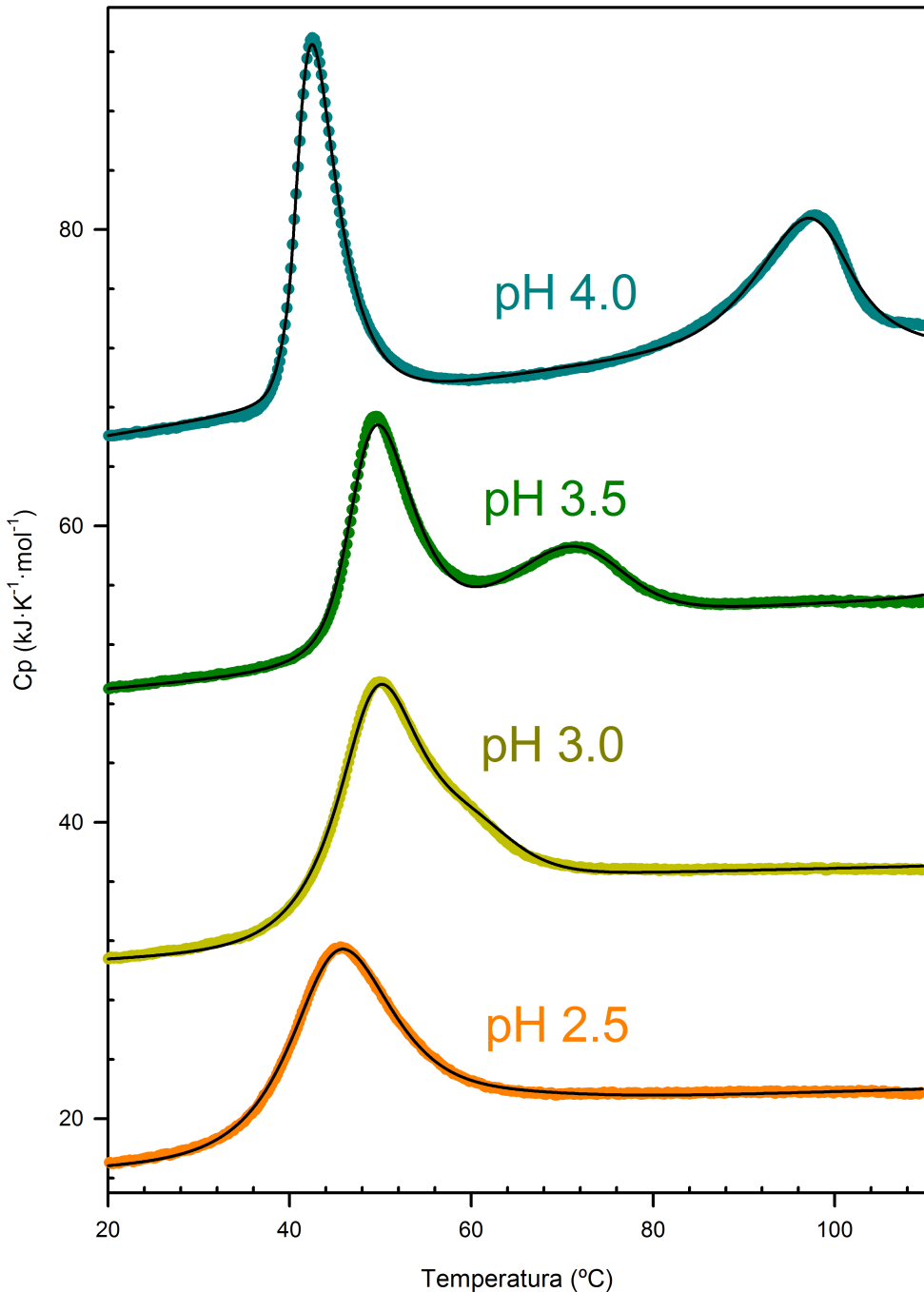


Figura 7.20. Curvas de desplegamiento térmico del dominio PDZ3-h-PSD95 seguidas por DSC en función del pH a una concentración de proteína de $1.3 \text{ mg}\cdot\text{mL}^{-1}$ y de tampón 50 mM. En puntos coloreados representan los datos experimentales, mientras que las líneas de color negro son los análisis. La curva de pH 4.0 se analizó con el modelo correspondiente al Esquema 3.3, mientras que para valores menores a este pH se usó el modelo $nN \rightleftharpoons I_n \rightleftharpoons nD$.

A la vista de estas diferencias, en lugar de realizar un análisis de todas las trazas con un modelo único, se han utilizado modelos distintos en función del valor del pH. Así, a pH 4 el modelo utilizado fue el que se usó para el dominio PDZ3-h-PSD95 a pH neutro (esquema 3.3). En cambio, las trazas obtenidas a valores de pH de 3.5 y 3.0, se utilizó el modelo usado para el dominio PDZ3-PSD95 a pH neutro (esquema 3.2), ya que el análisis previo de estas trazas que se realizó con el modelo correspondiente al esquema 3.3, que presenta un intermedio monomérico previo a la asociación, no reproduce correctamente el comportamiento experimental. Finalmente, en el caso de la curva a pH 2.5 ocurrió algo similar. El análisis de dicha curva realizado inicialmente con el modelo de dos estados no reproduce bien la traza experimental, tal y como se observa a modo de ejemplo en la Figura 7.21. Además, aunque reproduce correctamente la función de la capacidad calorífica del estado desnaturalizado, C_{pD} , no se reproduce correctamente la pendiente de la capacidad calorífica del estado nativo, C_{pN} (Figura 7.21). Esto condujo a analizar la traza a pH 2.5 con el modelo de tres estados con asociación del intermedio, correspondiente al esquema 3.2. Así, la calidad del ajuste con el modelo de dos estados corresponde a un $R^2 = 0.992$, mientras que en la de tres estados presenta un $R^2 = 0.998$ (Tabla 7.5), lo que termina por justificar el uso del modelo del esquema 3.2.

Todos los análisis se realizaron con una estequiometría $n = 3$ para el intermedio oligomérico, la misma estequiometría obtenida en los análisis a pH 7.5, que además coincide con la estequiometría del intermedio del dominio PDZ3-PSD95. Los resultados de los análisis muestran de nuevo que se produce una desestabilización del intermedio conforme disminuye el pH de la disolución, ya que se va produciendo una disminución de la población del mismo. De hecho, la proporción máxima que presenta el intermedio oligomérico abarca desde el 90% a pH 4 hasta el 20% a pH 2.5 (Figura 7.22). De esta forma, queda claro que, al igual que ocurre en el caso del dominio PDZ3-PSD95, el intermedio observado en el dominio PDZ3-h-PSD95 se desestabiliza a valores menores de pH.

Tabla 7.4. *Parámetros termodinámicos del desplegamiento térmico del dominio PDZ3-h-PSD95 en tampón acetato 50 mM pH 4 obtenidos del análisis de experimentos de DSC con el modelo de que se corresponde al Esquema 3.3.*

T_{N-I} (°C)	$\Delta H_{N-I}(T_{N-I})$ (kJ·mol ⁻¹)	T_{I-In} (°C)	$\Delta H_{I-In}(T_{I-In})$ (kJ·mol ⁻¹)	T_{I-D} (°C)	$\Delta H_{I-D}(T_{I-D})$ (kJ·mol ⁻¹)	$\Delta G_{N-D}(298K)$ (kJ·mol ⁻¹)	R^2
53 ± 1	250 ± 20	106 ± 8	-55 ± 24	93 ± 2	190 ± 15	58 ± 2	0.997

Los intervalos de error se han considerado como tres veces los errores del ajuste (correspondiente a un intervalo de confianza del 99%).

Tabla 7.5. Parámetros termodinámicos del desplegamiento térmico del dominio PDZ3-h-PSD95 en tampón glicina a una concentración de 50 mM obtenidos del análisis de experimentos de DSC según el modelo $3N \rightleftharpoons I_3 \rightleftharpoons 3D$ para tres valores de pH ácido.

pH	T_{N-D} (°C)	$\Delta H_{N-D}(T_{N-D})$ (kJ·mol ⁻¹)	T_{D-In} (°C)	$\Delta H_{D-In}(T_{D-In})$ (kJ·mol ⁻¹)	$\Delta G_{N-D}(298K)$ (kJ·mol ⁻¹)	R^2
3.5	56.5 ± 0.4	233 ± 5	73.2 ± 0.4	-107 ± 3	18 ± 2	0.997
3	52.4 ± 0.6	236 ± 5	67 ± 3	-55 ± 9	16 ± 2	0.998
2.5	45.9 ± 0.4	196 ± 3	57 ± 3	-33 ± 6	10 ± 2	0.998

Los intervalos de error se han considerado como tres veces los errores del ajuste (correspondiente a un intervalo de confianza del 99%).

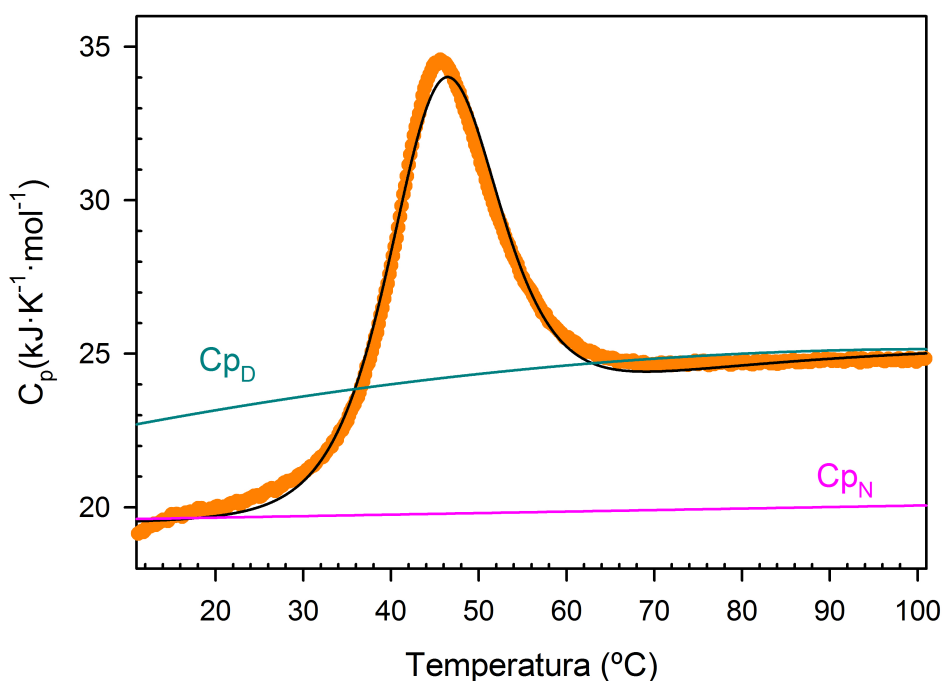


Figura 7.21. Curva de desplegamiento térmico del dominio PDZ3-h-PSD95 a $1.3 \text{ mg} \cdot \text{mL}^{-1}$ seguida por DSC en glicina 50 mM pH 2.5. Los puntos experimentales son los círculos en color naranja, mientras que el ajuste al modelo $N \rightleftharpoons D$ se representa mediante una línea negra. También se han representado las capacidades caloríficas tanto del estado nativo C_{pN} (en color rosa), como del desnaturalizado, C_{pD} (en color azul).

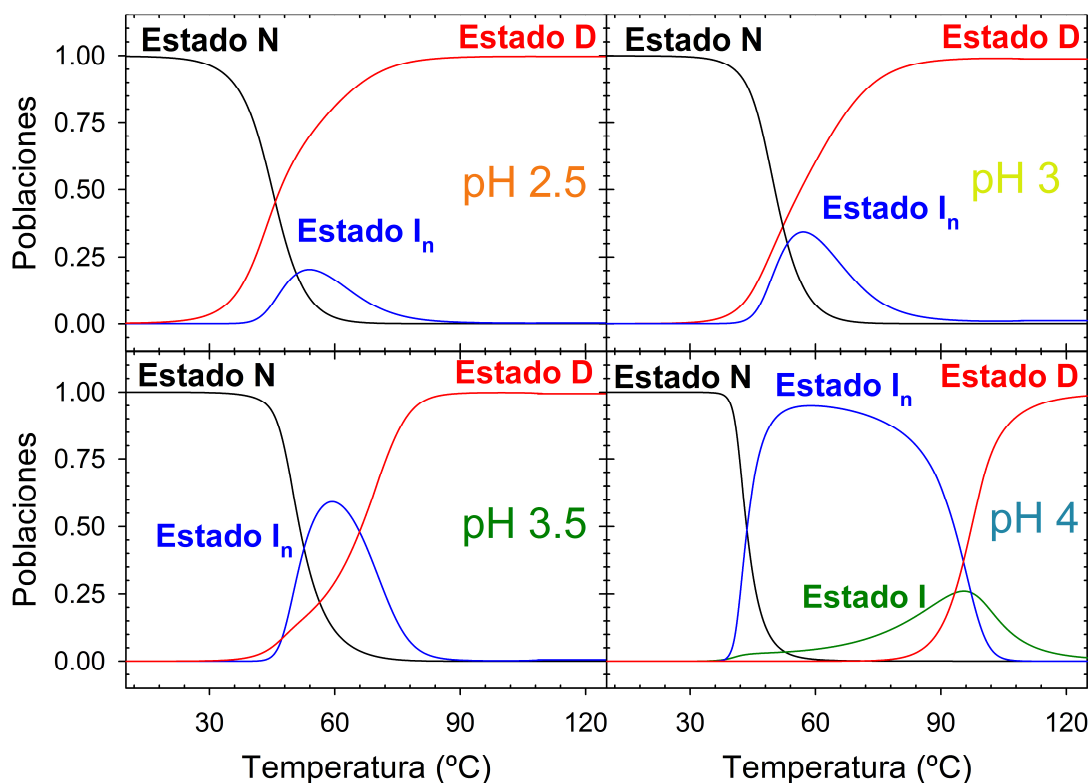


Figura 7.22. Distribución de las poblaciones de los diferentes estados conformacionales presentes en el desplegamiento térmico del dominio PDZ3-h-PSD95 a $1.3 \text{ mg} \cdot \text{mL}^{-1}$ y a distintos valores de pH. En color negro se representa la población del estado nativo (N), en rojo el desnaturalizado (D), en azul el estado intermedio asociado (I_n) y en verde el estado intermedio monomérico (I), descrito únicamente en el caso de pH 4.

Paralelamente a los experimentos realizados con el dominio PDZ3-PSD95 que se detallan en el Capítulo 6, se estudió si a pH ácido el intermedio de la construcción PDZ3-h-PSD95 puede formar también fibras amiloides. Se incubaron varias muestras a 60 °C de PDZ3-h-PSD95 a $7.2 \text{ mg} \cdot \text{mL}^{-1}$ en glicina 50 mM pH 3 durante distintos tiempos de incubación desde cuatro horas hasta una semana, para examinar la posible formación de fibras mediante TEM. Se eligió este valor ácido de pH con objeto de comparar los experimentos realizados con el dominio PDZ3-PSD95. Según las micrografías obtenidas, aparecen protofibras a las 4 horas de incubación (micrografía señalada con “4 horas” de la Figura 7.23), tal y como ocurría en el caso del dominio PDZ3-PSD95. Además, al cabo de un día de incubación aparecen fibras más maduras (micrografía señalada con “1 día” de la Figura 7.23), de nuevo de forma similar al dominio PDZ3-PSD95. Finalmente, tanto el aspecto enmarañado de las fibras de ambas construcciones, como la abundancia de fibras que presentan en una semana, hace concluir que la formación de fibras es similar en ambas construcciones del dominio a pH ácido.

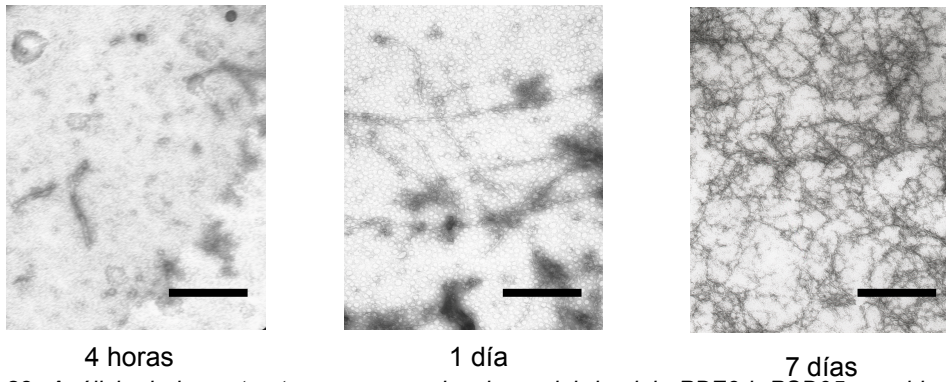


Figura 7.23. Análisis de las estructuras macromoleculares del dominio PDZ3-h-PSD95 seguidas por TEM. En el pie de cada micrografía está indicado el tiempo de incubación de la disolución de proteína a $7.2 \text{ mg}\cdot\text{mL}^{-1}$, en glicina 50 mM pH 3.0 a 60°C . La barra de escala corresponde a un tamaño de 200 nm

Por otro lado, se siguió la cinética de formación de las fibras mediante la incubación con ThT y ANS. Se incubaron dos muestras de PDZ3-h-PSD95 a $7.2 \text{ mg}\cdot\text{mL}^{-1}$ a 60°C , una en presencia de $12.5 \mu\text{M}$ de ThT y la otra con ANS a una concentración de $20 \mu\text{M}$. Los resultados indican la presencia de una fase de retardo cuando se sigue la fluorescencia de ThT, mientras que no aparece dicha fase de retardo en el caso del ANS (Figura 7.24). Esta diferencia se puede deber simplemente a que el intermedio oligomérico expone superficies hidrofóbicas al medio que aumentan a medida que dicho estado aumenta su población, de acuerdo con el incremento observado en la fluorescencia del ANS. Sin embargo, este intermedio oligomérico no une a ThT, tal y como se deduce por la fase de retardo que presenta. Sin duda, han de realizarse más pruebas para profundizar en el mecanismo de fibrilación, aunque los datos presentados permiten concluir que la desestabilización del intermedio con el pH se da de forma paralela en ambas construcciones.

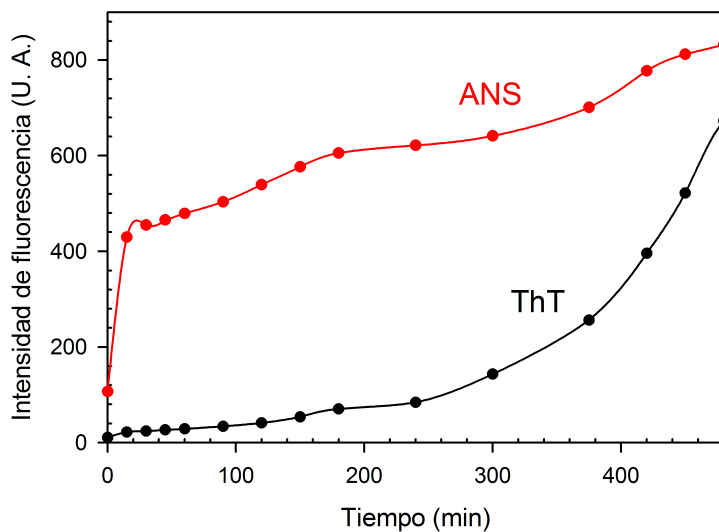


Figura 7.24. Cinéticas de crecimiento de la emisión de fluorescencia de ThT (negro) y ANS (rojo). Se incubó una muestra de PDZ3-h-PSD95 a $7.2 \text{ mg}\cdot\text{mL}^{-1}$ en la presencia de $12.5 \mu\text{M}$ de ThT o $20 \mu\text{M}$ de ANS en glicina 50 mM pH 3 y a 60°C de temperatura.

7.2 Discusión

Según los resultados obtenidos a partir de las distintas técnicas usadas en esta Memoria, la hélice α_3 situada en el extremo C-terminal del dominio PDZ3-PSD95 es un elemento regulador del plegamiento de este dominio. Cuando esta hélice se elimina, en la construcción PDZ3-h-PSD95 sigue detectándose la presencia de un intermedio que presenta además fenómenos de asociación, aunque no es posible conocer de forma precisa su estequiometría de asociación. De esta manera, aunque los análisis de DSC muestran una estequiometría correspondiente a un trímero, los resultados de DLS y de SEC no muestran indicios de la presencia de tales partículas, sino que se observan partículas de tamaños mayores. Esto podría deberse a la inter-asociación de los trímeros, obedeciendo al modelo descrito por Radford (Smith et al., 2006), en el que se describe que las moléculas van asociándose en diferentes grados para, en un paso final, formar fibras amiloides.

En todo caso, todas las evidencias revelan que la eliminación de la hélice α_3 de este dominio produce una estabilización del intermedio. Además, esa estabilización se confirma por la mayor tendencia a la fibrilación del dominio PDZ3-h-PSD95. Además, no se observa la formación de las estructuras anulares que se dan en el dominio PDZ3-PSD95.

7.2.1 Causas estructurales de la estabilización del intermedio de plegamiento del dominio PDZ3-PSD95 cuando se elimina su tercera hélice α

Para intentar elucidar la naturaleza molecular de la estabilización del intermedio, se realizó de nuevo un análisis de la tendencia a la agregación hecho con el programa TANGO (Fernandez-Escamilla et al., 2004) (Figura 7.25). Dicho análisis no revela ninguna diferencia significativa con respecto a la construcción que contiene la hélice, por lo que en principio los residuos responsables de la agregación han de ser los mismos. Así, al recordar de nuevo la estructura del dominio (Figura 7.1) se puede observar que la hélice interacciona con la zona β_3 - α_1 (residuos 340 a 350) y se sitúa a continuación de la zona β_5 - α_3 (residuos 392 a 397), que son precisamente las regiones con tendencia a formar agregación β que se predicen por el programa TANGO. Por lo tanto, la ausencia de dicha hélice puede afectar a ambas regiones del dominio, haciendo en ambas más fácil la accesibilidad al disolvente, lo cual explica la mayor tendencia a la asociación.

El hecho de que al disminuir el pH se produzca una desestabilización del intermedio, implica de nuevo la presencia de un proceso de protonación de residuos cuyo pK_a está de nuevo alrededor de 3. Debido a que no están presentes los residuos glutamato de la hélice α_3 , puede que los responsables de ese comportamiento sean los residuos ácidos presentes en el giro entre las láminas β_2 y β_3 .

Puesto que los espectros de CD realizados muestran una reestructuración del intermedio que implica un aumento de lámina β , puede ocurrir que sean tanto la región de las láminas β_2 y β_3 como la región de lámina β_5 las que se reestructuren en el intermedio. Ya se ha descrito en el Capítulo 3 que la succinimida de la posición 322 presente en el giro

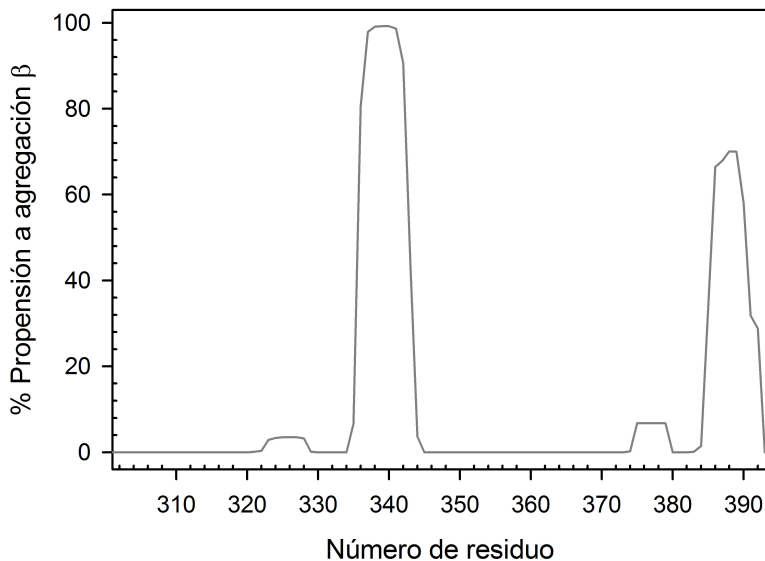


Figura 7.25. Gráfica del cambio del % de propensión de agregación en función del número de residuo (desde el 302 al 393 según la numeración de (Doyle et al., 1996)) obtenida a partir del programa TANGO al introducir la secuencia de aminoácidos del dominio PDZ3-h-PSD95 en un tampón estándar 50 mM pH 7.5.

entre las láminas $\beta 2$ y $\beta 3$ juega un papel importante en la interacción con la hélice $\alpha 3$, ya que interacciona con algunos residuos de la hélice, como el Glu401. Por lo tanto, la eliminación de la hélice $\alpha 3$ provoca que no exista esa interacción y que esa zona esté más desestructurada. Es posible así que que la región de la zona $\beta 3$ - $\alpha 1$, que presenta residuos con mayor tendencia a agregación β , se exponga al disolvente. Por otro lado, con respecto a la región de la lámina $\beta 5$, hay que tener en cuenta que a continuación de dicha lámina, está presente la hélice $\alpha 3$. Por ello, al truncar este elemento estructural, la lámina $\beta 5$ estaría menos empaquetada en el dominio, quedándose en el extremo C-terminal de la proteína. De esta forma, tendría una mayor facilidad para exponerse al disolvente en la reestructuración del intermedio. Además, puesto que dicho intermedio se desestabiliza en un intervalo de pH alrededor de 2.5, nuevamente han de ser algunos residuos cuyos equilibrios de protonación los que estén afectando al plegamiento. Debido a esto, se hace necesario el diseño de otros mutantes y el uso de técnicas adicionales para poder establecer qué está ocurriendo a nivel molecular.

7.2.2 Relevancia biológica de la presencia de la tercera hélice α del dominio PDZ3-PSD95

En el apartado anterior, se ha hablado de la trascendencia molecular que la eliminación de la hélice $\alpha 3$ podría tener en la capacidad del dominio PDZ3-PSD95 para unir sus ligandos. Las consecuencias que dicha eliminación conlleva en el mecanismo de plegamiento son evidentes, y además pueden tener consecuencias de relevancia biológica.

Al trincar dicha hélice, la tendencia a la asociación es mucho más fuerte en el dominio. Tanto es así, que la asociación comienza a tener lugar a temperaturas mucho menores que cuando está presente la hélice, además de que ocurre a concentraciones mucho más bajas. Ya se ha comentado la estabilización del intermedio, la cual, por otro lado, puede estar apoyada estructuralmente por lo que se ha dicho en la Sección anterior. Además, el intermedio tiende a formar fibras amiloides a una velocidad mayor a la que ocurre cuando está presente la hélice. De hecho, ya se vio en la Sección 5.1.2 que el dominio PDZ3-PSD95 a pH 3 no forma estructuras anulares y que la formación de las fibras se da por ello más rápidamente. En el caso del dominio PDZ3-h-PSD95 ocurre lo mismo, se acelera la formación de las fibras, que puede ser debida, entre otras razones, a que no se forman las estructuras anulares. Por otro lado, y de nuevo al igual que ocurría a pH 3.0 en el dominio PDZ3-PSD95, la fibrilación que se da en el dominio PDZ3-h-PSD95 es irreversible, en contraposición a la reversibilidad que presenta la construcción con la hélice a pH 7.5.

Hay asimismo un dato que pone de manifiesto la posible relevancia biológica de este elemento estructural. La asociación comienza a unos 34-36°C (Figura 7.8) y el dominio termina formando a lo largo del tiempo fibras amiloides a 37°C, temperatura fisiológica (micrografía señalada con “37°C” de la Figura 7.13). Hay que tener en cuenta que la proteína PSD95 es un elemento clave en la regulación del tráfico de señalización en la sinapsis neuronal (Sheng & Sala, 2001; Kim & Sheng, 2004; Dalva et al., 2007). Por otro lado, según Nussinov y colaboradores (Tsai et al., 2009), a la hora de hablar de proteínas *hub* – tal es el caso de PSD95 – la capacidad de unir tantas moléculas diana que dichas proteínas presentan no se debe a la flexibilidad o desorden intrínsecos de la proteína. Más bien, esta propiedad puede deberse a la posibilidad de la existencia de distintos productos génicos de una misma proteína, en otras palabras, la posibilidad de que el gen codifique varias formas de la misma proteína. Por ejemplo, que se expresen los dominios aislados.

De esta forma, podría darse la expresión del dominio PDZ3-PSD95 aislado para ejercer alguna función celular. Si tal fuese el caso, el papel de la tercera hélice α sería clave, puesto que en función de su presencia se podrían formar unas estructuras macromoleculares u otras, prevenir su asociación o fomentarla, etc.

Además, no solamente la presencia o ausencia de la hélice podría modular esto, sino que, debido a la dependencia del pH que presenta el plegamiento de este dominio (Capítulo 6), su electrostática podría ser modulada por el micro-ambiente generado alrededor de la membrana celular (Ma & Dougherty, 1997; Lund & Jönsson, 2005) en donde la proteína ejerce su función (para más detalles el Capítulo 2). Se podría controlar así, por ejemplo, la liberación de ligandos unidos a la proteína mediante un cambio conformacional en la hélice $\alpha 3$ (Petit et al., 2009).

De hecho, las interacciones electrostáticas entre la hélice $\alpha 3$ y el giro $\beta 2$ - $\beta 3$ pueden ser las responsables de la organización reversible de las estructuras anulares vistas únicamente a pH neutro (Capítulo 5). Sería necesaria más investigación y con otro tipo de técnicas diferentes a las usadas en esta Memoria para saber si dichas estructuras anulares pueden desarrollar toxicidad *in vivo* que parece darse mediante la formación de

canales iónicos en la membrana plasmática que provoca la pérdida de la homeostasis (Jang et al., 2008; Lashuel & Lansbury, 2006).

De hecho, en colaboración con la Dra. Sandra Villegas de la Universidad Autónoma de Barcelona, estamos haciendo ensayos con líneas celulares para establecer la posible toxicidad *in vivo* de las estructuras observadas, así como el uso de la espectroscopía de infrarrojo, para poder conocer con más detalle los cambios estructurales tipo β que se producen en los procesos descritos. Además pretendemos hacer uso de otras microscopías (como por ejemplo, la microscopía de fuerza atómica, *AFM*) para conocer con más detalle las estructuras observadas. Pensamos por otro lado que el uso de la resonancia magnética nuclear (*RMN*) sería idóneo para conocer mejor la conformación del intermedio.

También pretendemos completar estos estudios con el análisis de varios mutantes de este dominio en los que se cambie la naturaleza de los residuos ácidos que pueden ser los responsables de la variación del comportamiento experimental en función del pH. Pensamos que la mutación del resto Glu334 o el resto Glu401, o de ambos, que interaccionan entre sí conectando el giro entre las cadenas $\beta 2$ y $\beta 3$ y la hélice $\alpha 3$ puede ser clave para entender el mecanismo del plegamiento. Para ello, estamos pensando en eliminar alguna de las dos cargas, o bien introducir un resto con una carga positiva en su cadena lateral. Asimismo, el Asp332 que forma la succinimida puede tener también influencia, por la interacción descrita con el residuo Glu334 presente en la hélice $\alpha 3$. Sería conveniente sustituirlo por mutaciones no formadoras del anillo succinimida como glutamato que preservaría la carga negativa del aspartato, o bien glicina o alanina que eliminarían completamente la cadena lateral donde se forma la succinimida y teóricamente dotarían a esta zona de la proteína de una mayor flexibilidad conformacional. Contrariamente, la mutación por prolina emularía a la succinimida en disolución, lo que nos llevaría a observar qué efecto tiene en el plegamiento la rigidez que confiere la ciclación de una cadena.

En todo caso, y al igual que numerosos ejemplos de otras proteínas, (Gokhale & Khosla, 2000; Tanaka et al., 2003; Wriggers et al., 2005; Durand et al., 2010), en las proteínas multidominio el papel que presentan los elementos de unión entre sus dominios es clave, tal y como se ha visto con la proteína PSD95 en esta Memoria y en otros estudios previos (Long et al., 2003; Korkin et al., 2006).

CAPÍTULO 8

ANÁLISIS DE LA ESPECIFICIDAD DE UNIÓN
DE LOS DOMINIOS PDZ MEDIANTE LA
METODOLOGÍA DE EXPRESIÓN EN FAGOS

8.1 Resultados

Los estudios de la especificidad de unión de los dominios PDZ realizados hasta la fecha han conducido a conocer qué restos situados en los distintos elementos secundarios de un dominio PDZ son los que afectan a su especificidad de unión. Gracias a dichos estudios, comentados con más detalle en el Capítulo 2, y desarrollados en su mayoría en el laboratorio del Dr. Sidhu, se han podido establecer tanto un mapa de especificidad de dominios PDZ naturales (Tonikian et al., 2008), como generar familias sintéticas de dominios PDZ que han permitido simular los posibles mecanismos evolutivos de una familia de dominios modulares (Ernst et al., 2009; Ernst et al., 2010). Como continuación de estos estudios, y en colaboración con el Dr. Sidhu, en esta Memoria se describe un protocolo desarrollado mediante la metodología de expresión en fagos para poder establecer los determinantes moleculares de la especificidad de unión de dominios PDZ para cada posición del ligando. Esta metodología permite co-expresar una secuencia aminoacídica junto a una proteína de la cápside de un fago mediante la introducción en su genoma de la secuencia de ADN que codifica dicha secuencia aminoacídica. Esto conlleva que este polipéptido se exprese en la superficie del fago y que se pueda conocer su secuencia aminoacídica mediante la secuenciación de su ADN. Además, mediante ingeniería genética es posible introducir diversas mutaciones en varios restos de la secuencia aminoacídica co-expresada, lo que permite generar librerías polipeptídicas a partir de las cuales se pueden seleccionar las variantes más afines a determinadas moléculas diana (para ver más detalles de esta metodología, ver la Sección 3.13).

De acuerdo con esto, el protocolo descrito en esta Memoria ha consistido en construir una librería de dominios PDZ en los que se varíen los restos que determinan la especificidad de unión para cada una de las posiciones del ligando y seleccionar aquellos dominios con cuyos residuos se optimiza la interacción con sus ligandos. La estrategia consistió en co-expresar, fusionado a la proteína III del fago M13, un dominio PDZ cuyas propiedades de unión son bien conocidas, el dominio PDZ-Erbin. Los residuos responsables de la especificidad de unión para una posición determinada del ligando se mutaron para obtener una librería de variantes de este dominio. De esta forma se obtiene una librería de dominios PDZ en la que están todas las posibles variantes existentes en la naturaleza relativas a la interacción con el residuo presente en esa posición del ligando. Es decir, no se cambia el esqueleto básico del dominio PDZ-Erbin, pero al introducir variabilidad en la región responsable de la interacción, se están recogiendo todas las posibilidades que se dan en la naturaleza referidas a esta interacción concreta. Puesto que las librerías generadas por la metodología de expresión en fagos se pueden obtener con un límite experimental de variabilidad (hasta 10^{10} variantes), para simplificar los estudios se decidió estudiar cada una de las posiciones por separado.

Debido a que la región que interacciona con el residuo de la posición 0 del ligando es la que está mejor localizada, el estudio se realizó para dicha posición, para luego extenderlo en estudios posteriores a las posiciones -1, -2 y -3. Tal y como se menciona en el Capítulo 2, en la posición 0 del ligando, que es el extremo C-terminal de la molécula

diana, el grupo carboxilato del residuo presente en dicha posición se acomoda en un bolsillo hidrofóbico en el que suele estar presente la secuencia aminoacídica GLGF, muy conservada en estos dominios y que se encuentra en el giro entre las cadenas $\beta 1$ y $\beta 2$ ($\beta 1:\beta 2$, posiciones 4 a 7). Completan el bolsillo hidrofóbico el primer resto de la cadena $\beta 2$ ($\beta 2-1$) y los dos últimos restos que forman la hélice $\alpha 2$ ($\alpha 2-9$ y $\alpha 2-10$). En la interacción más común con ligandos, la carga negativa del grupo carboxilato del extremo C-terminal se coordina con los protones amida del motivo GLGF (Figura 8.1). A su vez, es muy importante la interacción de la cadena lateral del aminoácido presente en el ligando con las cadenas laterales de los residuos $\beta 1:\beta 2-6$, $\beta 2-1$, $\alpha 2-9$ y $\alpha 2-10$ (Figura 8.1). Estas interacciones tienen generalmente un carácter marcadamente apolar y son las determinantes de la especificidad de unión de estos dominios en dicha posición (Skelton et al., 2003; Appleton et al., 2006; Tonikian et al., 2008).

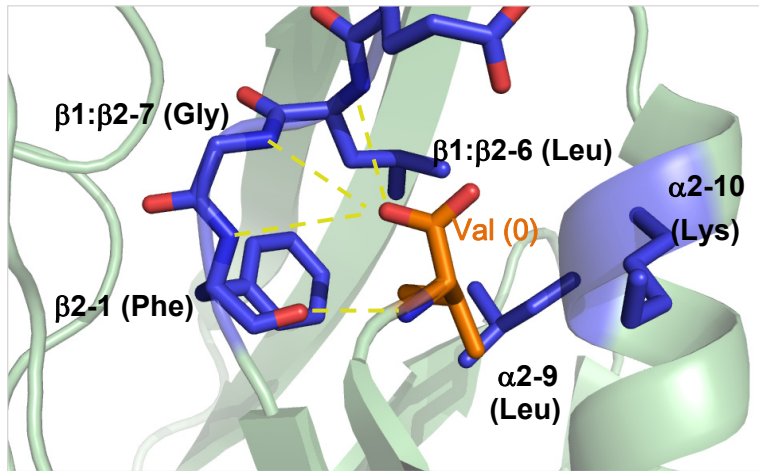


Figura 8.1. Interacción en la región del bolsillo hidrofóbico del dominio PDZ-Erbin unido al péptido de alta afinidad TGWETWV. El extremo C-terminal del péptido está en color naranja, mientras que los residuos del dominio PDZ responsables de la interacción están en color azul. Las interacciones polares, entre las que se encuentran las correspondientes al grupo carboxilato del extremo C-terminal del péptido, están representadas en líneas discontinuas de color amarillo.

En el caso del dominio PDZ elegido para generar la librería, el dominio PDZ-Erbin, el aminoácido que interacciona de forma óptima en dicho bolsillo es la valina, que tiene el tamaño adecuado para la interacción con los residuos que forman ese bolsillo, el cual presenta los residuos Leu26, Gly27 y Phe28, a los cuales le falta una glicina en la posición 25 para constituir el motivo GLGF, mencionado previamente como un motivo muy conservado en estos dominios. En el dominio PDZ-Erbin no está conservado propiamente, presentando un glutámico en vez de la primera glicina, siendo por tanto ELGF. La valina de la posición 0 del ligando también interacciona con los dos restos $\alpha 2-9$ y $\alpha 2-10$ de la hélice $\alpha 2$ del dominio, que son la Leu86 y la Lys87. Completan la región de unión al residuo de la posición 0 del ligando los residuos Glu19, Lys20, Ser21 y Pro22, presentes todos en el giro $\beta 1:\beta 2$ (Figura 8.2).

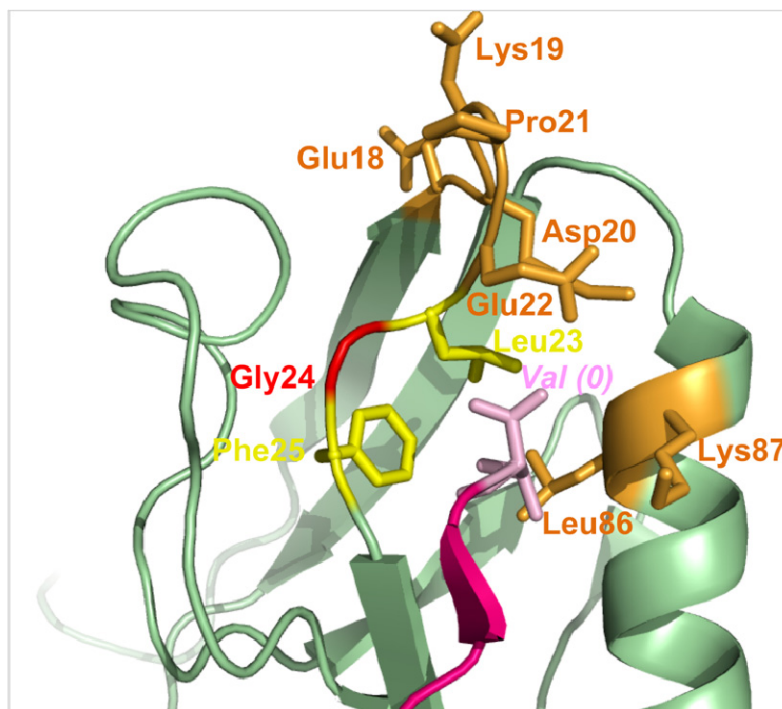


Figura 8.2. Diseño de la librería del dominio PDZ-Erbin. La estructura del dominio está coloreada en función de los cambios introducidos. En rojo está la Gly24, que se mantiene por razones estructurales. La Leu23 y la Phe25, que están en amarillo, han sido sustituidas por aminoácidos de carácter hidrofóbico, mientras que los residuos coloreados en naranja han sido cambiados por los veinte aminoácidos presentes en la naturaleza. En color rosa se representa el péptido de alta afinidad TGWETWV, cuyo aminoácido presente en el extremo C-terminal (posición 0) ha sido coloreado en violeta.

De acuerdo con lo que se ha venido mencionando hasta ahora, la estrategia que se ha seguido en esta Memoria ha sido construir una librería de variantes del dominio PDZ-Erbin que se diferencian entre sí en los residuos mencionados que forman el bolsillo hidrofóbico. Por otro lado, se obtuvieron las moléculas diana a partir de las cuales se pudiesen seleccionar las variantes óptimas del dominio PDZ-Erbin que nos den la información de la especificidad de unión en función de sus secuencias. Para el diseño de estas moléculas diana se eligió como base la secuencia peptídica TGWETWV, que es una de las más afines descritas hasta la fecha (Skelton et al., 2003) y de la cual se tiene la estructura cristalográfica (código PDB 1N7T). Se diseñaron variantes de este péptido al que se cambió únicamente el aminoácido de la posición 0, valina en el original. De esta forma se obtuvieron un total de 20 secuencias distintas de dicho péptido, una con cada aminoácido natural en la posición 0 del ligando.

Por otro lado, es conveniente recordar que, tal y como se detalla en el Capítulo 2, los dominios PDZ pueden también presentar reconocimiento interno de secuencias peptídicas, tanto una cadena β de otra proteína, como el reconocimiento de otro aminoácido en una hipotética posición +1. Para mimetizar este tipo de reconocimiento, se diseñaron otras 20 variantes del péptido TGWETWV, al cual se añadió a su extremo C-

terminal la secuencia de aminoácidos **XGGGKK**. En la posición en la que se sitúa la letra **X** se introdujeron cada uno de los veinte aminoácidos naturales. Se añadió la secuencia **XGGGKK** porque la presencia de las dos lisinas provoca que, mediante la repulsión de sus cargas positivas con la región hidrofóbica del dominio, no se establezca la posible coordinación del carboxilato del extremo C-terminal del nuevo péptido. Por otro lado, las tres glicinas confieren una alta flexibilidad para que las dos lisinas se puedan acomodar de la mejor forma. De esta manera se permitiría que la secuencia **TGWETWV** teóricamente se acomodara en su zona de unión, dejando la siguiente posición, la **X**, como la posible posición +1 del ligando.

Una vez construidas tanto la librería como las moléculas diana, se han realizado experimentos de selección de las variantes más afines a cada una de las moléculas diana. Posteriormente, se realizaron más experimentos de selección partiendo de estas variantes ya seleccionadas incluyendo un proceso de competición con los péptidos que presentan los aminoácidos más frecuentes en la posición 0 y así mejorar la selección (Figura 8.3).

8.1.1 Construcción de la librería del dominio PDZ-Erbin para el estudio de la especificidad relativa a la posición 0 del ligando

Para la construcción de una librería de proteínas mediante la metodología de expresión en fagos inicialmente se ha de tener un fagémido que codifique la proteína de interés fusionada a la proteína III de la cápside, puesto que tal y como se menciona en la Sección 3.13.2, es la proteína del fago que permite la co-expresión de proteínas de tamaños grandes. Para ello se disponía en el laboratorio del Dr. Sidhu de un fagémido que contiene el gen con la secuencia de nucleótidos que codifica el dominio PDZ-Erbin y el gen de la proteína III del fago M13.

Antes de la generación de la librería, hay que introducir en dicho fagémido codones de parada en las regiones que se van a mutar puesto que el método de mutagénesis que se ha usado, descrito originariamente por Kunkel (Kunkel, 1985), tiene un rendimiento de un 80%. Esto implica que de las 10^{10} variantes que suelen contener este tipo de librerías, el 20% corresponde al molde original, lo que provocaría un proceso de competición que no estaría ponderado, puesto que habría $2 \cdot 10^9$ copias de la proteína original, mientras que habría pocas copias del resto de variantes, que todas juntas sumarían el total de las $8 \cdot 10^9$ restantes. De esta forma, al introducir los codones de parada, la proteína natural no se expresa completa, por lo que no va a unir y no va a competir con el resto de variantes.

Para la síntesis del fagémido que incluye los codones de parada, se usó el método de mutagénesis de Kunkel, para lo cual se diseñaron dos oligonucleótidos de tal forma que se haya quince bases que se alineen con las regiones contiguas a la zona en que se introducen los codones de parada. Una de las regiones es el giro entre las cadenas $\beta 1$ y $\beta 2$ (residuos Glu18 hasta Phe25) y la otra es una parte de la hélice $\alpha 2$ (residuos Leu86 y Lys87). Poner tres codones de parada resulta suficiente como para asegurar que la proteína original no se va a co-expresar. Los oligonucleótidos así diseñados, que se denominan “de parada” en la Figura 8.4, se unieron a una cadena de ADN monocatenario

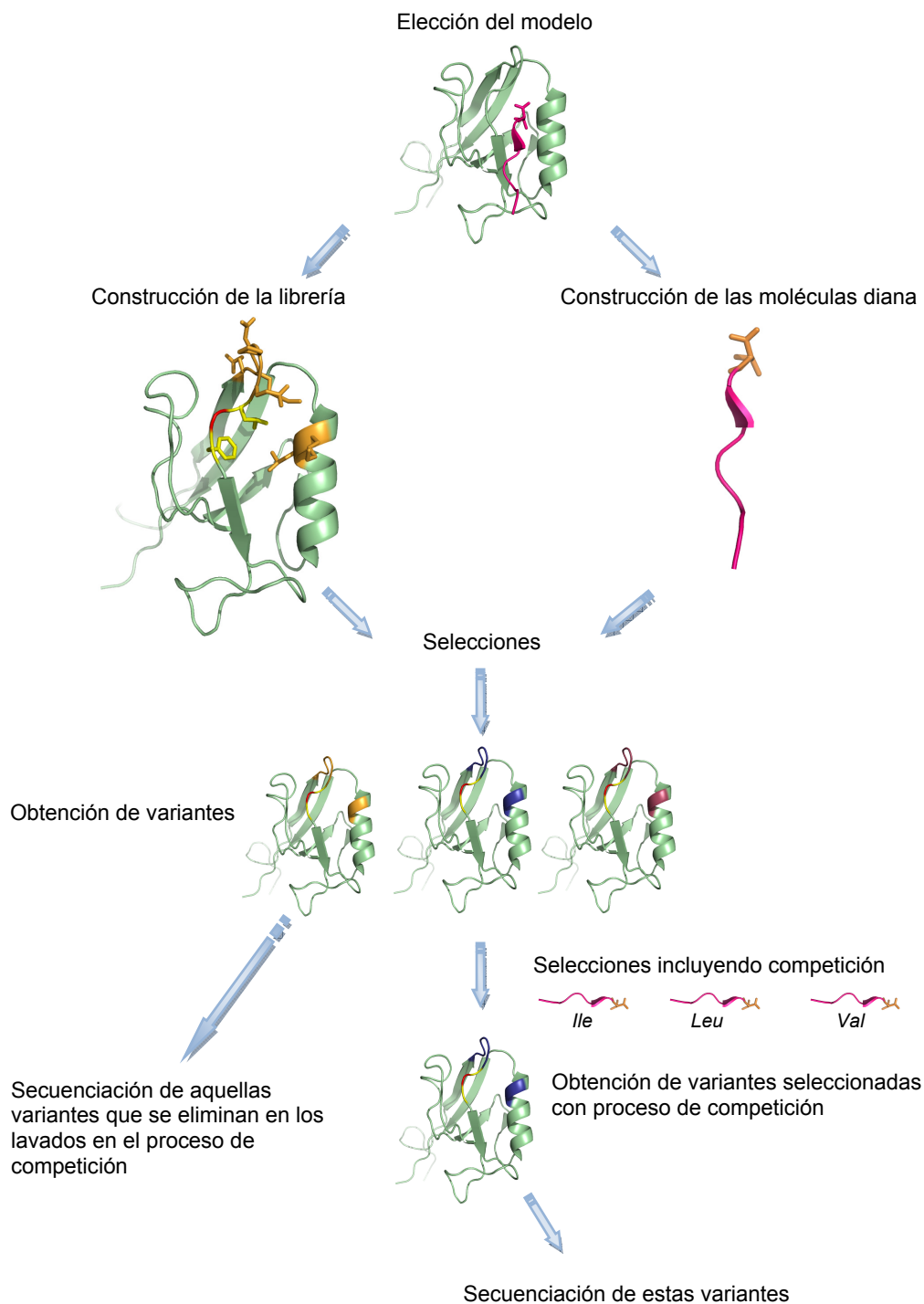


Figura 8.3. Esquema del protocolo desarrollado en esta Memoria para el estudio de la especificidad de los dominios PDZ correspondiente a la posición 0 del ligando.

molde proveniente del fagémido que contiene la secuencia que codifica Erbin y la proteína III del fago M13. A continuación, se añadió la polimerasa T7, que aporta al oligonucleótido que tiene los codones de parada las bases complementarias al molde, quedando dos extremos romos, los cuales se unieron mediante la acción de la ligasa T4, obteniéndose el mismo fagémido ya de doble cadena, pero con los codones de parada en la zona que se desea mutar (para más detalles sobre el método de mutagénesis de Kunkel, ver Sección 3.13.3).

A continuación se diseñaron dos oligonucleótidos mediante los cuales se introdujeron las mutaciones deseadas para la generación de la librería. Para el diseño de estos dos nuevos oligonucleótidos, se tomó la secuencia original del gen que codifica el dominio PDZ-Erbin en la que se sustituyeron aquellos codones que codifican los residuos que se querían mutar por codones que introducen las mutaciones deseadas. En cuanto a la región $\beta 1:\beta 2$, la Gly24 no se mutó, es decir, se mantuvo el codón GGA, puesto que dicha glicina es necesaria para permitir el giro estructural entre las cadenas $\beta 1$ y $\beta 2$ del dominio. Se ha señalado en la Figura 8.2 ese residuo en rojo, para indicar dicha restricción. Por otro lado, los residuos Leu23 y Phe25, coloreados en amarillo, se sustituyeron por aminoácidos apolares mediante el codón degenerado *NYT* (ver Tabla 3.2), que codifica los aminoácidos fenilalanina, isoleucina, leucina, valina, prolina, alanina, serina y treonina. Finalmente, el resto de residuos del giro $\beta 1:\beta 2$, que son los restos Glu19, Lys20, Ser21 y Pro22 y los restos Leu86 y Lys87 que están en la hélice $\alpha 2$, coloreados en naranja en la Figura 8.2, se sustituyeron por los veinte aminoácidos naturales, mediante el codón degenerado *NNK*. Para completar el diseño de los oligonucleótidos, se añadieron quince bases correspondientes a la secuencia original que codifica el dominio PDZ-Erbin a ambos lados de los codones que codifican las regiones mutadas. Esas quince bases son las mismas que están situadas a cada lado de los codones de parada en los oligonucleótidos “de parada” (Figura 8.4), con los que se sintetizó el molde para la generación de la librería, y proporcionan la complementariedad necesaria para la unión al molde. A continuación, los nuevos oligonucleótidos diseñados, denominados “mutagénicos” en la Figura 8.4, se unieron al molde que contenía los codones de parada y se completó el proceso de mutagénesis con la acción de la polimerasa T7 y de la ligasa T4 como ya se ha mencionado.

Cuando se completó el proceso de mutagénesis, el ADN resultante se introdujo mediante electroporación en células SS320, que, como ya se ha comentado en la Sección 3.13.3.1, son células *E. coli* obtenidas mediante la hibridación de los genomas del tipo MC 1061 y XL1-Blue, lo que les proporciona una alta capacidad de transformación y el pilus bacteriano para poder ser infectadas por los fagos. Después de la electroporación, las células se crecieron en medio de cultivo 2YT para obtener una concentración adecuada de fagos, del orden de 10^{12} por mililitro de cultivo. Finalmente, se sembraron varias diluciones de este cultivo en placas con el antibiótico ampicilina para conocer la variabilidad de la librería mediante un cálculo matemático a partir del contaje de colonias de la dilución sembrada en la placa, al que hay que aplicar el factor de dilución según la relación de volumen sembrado en placa y volumen total de cultivo. Cada una de esas colonias

Oligonucleótido de parada región $\beta 1:\beta 2$

ATA CGT GTG CGT GTG TAA TAA TAA AGC ATA TCA GGT
 I R V R V * * * S I S G

Oligonucleótido mutagénico región $\beta 1:\beta 2$

ATA CGT GTG CGT GTG NNK NNK NNK NNK NNK NYT GGA NYT AGC ATA
 I R V R V X X X X X ϕ G ϕ S I

TCA GGT
 S G

Oligonucleótido de parada región $\alpha 2$

CAA GCA GTG TCC TTG TAA TAA TAA ACT TTC CAG AAT ACA
 Q A V S L * * * T F Q N T

Oligonucleótido mutagénico región $\alpha 2$

CAA GCA GTG TCC TTG NNK NNK ACT TTC CAG AAT ACA
 Q A V S L X X T F Q N T

Figura 8.4. Secuencias de los oligonucleótidos usados para generar la librería. Se ha seguido el mismo código de colores para las mutaciones que el usado en la Figura 8.2: en color rojo el residuo que no se muta, en color naranja los 20 aminoácidos naturales (X) y en color amarillo los aminoácidos con cadena lateral apolar (ϕ). A ambos lados de la zona de mutación está la secuencia original del dominio PDZ-Erbin.

representa una estirpe de fagos con una única secuencia y por tanto es una única variante de la librería. La variabilidad experimental de la librería del dominio PDZ-Erbin resultó ser de $3 \cdot 10^9$ variantes. Puesto que en la librería se mutaron en total seis restos por los veinte aminoácidos naturales, y dos restos por ocho aminoácidos distintos, de naturaleza apolar, la librería presenta una variabilidad teórica de $20^6 \cdot 8^2 = 2 \cdot 10^8$ variantes, por lo que se puede concluir que la librería contenía todas las mutaciones posibles.

8.1.2 Construcción de las moléculas diana para el estudio de la especificidad relativa a la posición 0 del ligando

Una vez construida la librería de variantes del dominio PDZ-Erbin, el siguiente paso consistió en obtener las moléculas diana a partir de las cuales se pudiesen seleccionar los mutantes de dicha librería que interaccionan de forma óptima con cada uno de los veinte aminoácidos naturales en la posición 0 del ligando. Como ya se ha mencionado, se obtuvieron las 20 variantes del péptido TGWETWV, cada una de las cuales tiene un aminoácido natural en la posición 0 del ligando y otras 20 variantes del péptido TGWETWVXGGGKK, de nuevo con cada uno de los aminoácidos naturales en la posición X.

Todos los péptidos se obtuvieron unidos al extremo C-terminal de la proteína transferasa S de glutatión (*Glutathione S-transferase*, *GST*, en inglés). Para ello, también se usó el método de mutagénesis de Kunkel. Se diseñaron dos oligonucleótidos que contienen las secuencias de los péptidos de interés (TGWETW**X** y TGWETWV**X**GGGKK) en los que, para introducir los 20 aminoácidos naturales, se introdujo el codón degenerado NNK en la posición donde está situada la letra **X**. Ambos oligonucleótidos también contienen bases complementarias a un plásmido que contiene el gen que codifica la proteína GST, de tal forma que los nuevos plásmidos codifican los diferentes péptidos unidos a través de un enlace peptídico al extremo C-terminal de la proteína GST (las secuencias de los oligonucleótidos están recogidas en la Figura 8.5). De esta forma, los péptidos tienen libre su extremo C-terminal.

Una vez completada la mutagénesis, el nuevo DNA se introdujo en células de la cepa XL1-Blue mediante un proceso de transformación. Se seleccionaron varias colonias para la secuenciación del plásmido que contenían y se purificaron aquellos plásmidos que codifican las veinte variantes de ambos péptidos mediante una batería comercial de QIAGEN.

La fusión de estos péptidos a la proteína GST se hizo por dos razones. La primera, porque se pueden purificar con un protocolo estándar de purificación de proteínas de fusión a GST, tal y como se detalla en la Sección 3.2.2. La segunda, para que la inmovilización del péptido para las selecciones fuese por adsorción a la superficie *Maxisorp*. Esta superficie de la casa comercial *Nunc* tiene una alta capacidad de adsorción de macromoléculas con restos polares, es decir, la adsorción se da de forma bastante inespecífica. Así, cuando se quiso fijar el péptido para las selecciones en este tipo de placas, se decidió fusionarlo a una proteína grande para que, al ser de mayor tamaño que el péptido, la adsorción se produjese en mayor proporción por una región de la proteína que no fuese la correspondiente al péptido de interés, quedando éste accesible para el proceso de selección.

8.1.3 Selección de las variantes del dominio PDZ-Erbin más afines para cada uno de los veinte aminoácidos naturales presentes en el extremo C-terminal del ligando

Una vez obtenidas tanto la librería como las variantes del péptido TGWETW**X** fusionado a GST, se procedió a la selección de los mutantes de Erbin más afines para cada uno de los péptidos. Tal y como se describe en la Sección 3.13.3.2, se realizaron cinco rondas de selección para eliminar los mutantes que no unen a los péptidos y obtener aquellos que lo hacen con mayor afinidad. Dichas rondas consistieron en líneas generales en una etapa de incubación de la librería del dominio PDZ-Erbin con cada uno de los péptidos fusionados a GST fijados a la superficie de la placa para que se pueda producir la unión. Puesto que se fija poca cantidad de péptido a la placa, 1 µg, se usa una proteína estándar, la albúmina de suero bovino (*Bovine Serum Albumin*, *BSA*, en inglés), para bloquear el resto de las posiciones activas de la placa, a las que se podrían adsorber los

Oligonucleótido para introducir TGWETWX

CCG CGT GGT TCC GGT ACC ACC GGT TGG GAG ACC TGG NNK TAA TAA ACC
 P R G S G T T G W E T W X * * T
 GAT ACA ATT
 D T I

Oligonucleótido para introducir TGWETWVXGGGKK

CCG CGT GGT TCC GGT ACC ACC GGT TGG GAG ACC TGG GTG NNK GGG GGC
 P R G S G T T G W E T W V X G G
 GGT AAG AAG TAA TAA ACC GAT ACA ATT
 G K K * * T D T I

Figura 8.5. Secuencias de los oligonucleótidos usados para introducir los péptidos TGWETWX y TGWETWVXGGGKK en el extremo C-terminal de la proteína GST. Las secuencias correspondientes a dichos péptidos se han resaltado en gris. La secuencia previa a los péptidos corresponde al extremo C-terminal de la proteína GST.

fagos de forma inespecífica. A continuación, se realizó una etapa de lavado para eliminar aquellas variantes que no unen a los péptidos. Finalmente, para poder hacer una nueva ronda, se amplificaron los fagos conteniendo las variantes del dominio PDZ-Erbin que quedaron unidas después de la etapa de lavado.

Para cuantificar el proceso de selección y comprobar que se producía un enriquecimiento en los fagos que presentan las variantes del dominio PDZ-Erbin más afines para cada uno de los péptidos, se hizo un ensayo ELISA (que se describe con detalle en la Sección 3.13.3.3). Éste se realizó después de cada ronda con el conjunto de los fagos seleccionados, para tener una estimación de si se va produciendo un enriquecimiento en los fagos que unen a los péptidos conforme hay un mayor número de rondas. Este ensayo permite establecer si hay unión neta de la variante unida a la proteína del fago con el péptido diana, ya que se realiza de forma independiente tanto frente al péptido diana fusionado a GST como a la proteína GST sin el péptido, que actúa de control negativo.

Se añadieron los fagos a los péptidos diana fusionados a GST y la proteína GST sin dichos péptidos y se dejaron dos horas de incubación. A continuación se hicieron lavados con PBS y Tween 20 al 0.05% para eliminar aquellos fagos no unidos o unidos débilmente. Después se añadió el anticuerpo antiM13, se dejó un tiempo de incubación para a continuación hacer lavados con PBS y Tween 20 al 0.05% para eliminar el anticuerpo no unido. Posteriormente, se añadió el TMB, el cual es oxidado por la enzima peroxidasa de rábano común, la cual está conjugada al anticuerpo M13. Al oxidarse el TMB, este sustrato cromogénico dio un producto azul, el cual cambió a color amarillo cuando se añadió ácido fosfórico 1 M una vez que se había desarrollado por completo la reacción enzimática, lo que sucedió en unos dos minutos, según el color azul generado por

la reacción enzimática. El producto de color amarillo absorbe a 450 nm, por lo que la diferencia de absorbancia registrada a esa longitud de onda entre el ensayo con el péptido fusionado a GST y la proteína GST sin el péptido, dió una estimación de la diferencia de cantidad de anticuerpo presente, que es a la vez la diferencia de cantidad de fagos presentes. Así, esta diferencia nos permitió saber si hay unión entre la variante del dominio PDZ-Erbin al péptido diana, o si, la unión se debe a otra causa. Además, cuanto mayor fue la diferencia de absorbancia, indicó que hay más anticuerpo, lo que implica que haya un mayor número de fagos, por lo que se deduce que hay más variantes de Erbin que unen al péptido diana.

Se observó que conforme se va progresando con las rondas, aumenta la diferencia de absorbancia entre el ensayo con el péptido y el ensayo con la proteína GST únicamente realizado después de cada ronda, excepto para los péptidos *TGWETWI*, *TGWETWL* y *TGWETWV*, lo que significa que, excepto para estos tres últimos péptidos, el conjunto de los fagos se fue enriqueciendo en aquellos que presentan variantes del dominio PDZ-Erbin que unen con mayor afinidad a los péptidos diana frente a los cuales se seleccionan (Figura 8.6).

En este ensayo se observa también que la diferencia de absorbancia obtenida a partir del conjunto de las variantes del dominio PDZ-Erbin depende en gran medida del aminoácido presente en la posición 0 del péptido frente al cual fueron seleccionadas. Así, cuando este aminoácido es arginina, aspártico, cisteína, glutámico, glutamina, glicina o histidina el conjunto de los fagos dan una diferencia pobre de absorbancia, menor de 0.2, indicando que hay muy poca proporción de fagos que unen al péptido. Cuando el aminoácido es alanina, asparragina, lisina, serina, triptófano o tirosina, la diferencia de absorbancia es algo superior (entre 0.2 y 0.4), lo que está indicando que la proporción de fagos que unen al péptido es mayor frente a los que no. Finalmente, si el aminoácido presente en la posición 0 del péptido es isoleucina, leucina, metionina, fenilalanina, prolina, treonina o valina, la diferencia de absorbancia obtenida en el ensayo ELISA es mayor de 0.4, por lo que habrá más fagos que unan al péptido. El hecho de que no se produzca este aumento en la diferencia de absorbancia en el conjunto de los fagos que se seleccionaron frente a los péptidos que contienen isoleucina, leucina y valina (*TGWETWI*, *TGWETWL* y *TGWETWV*) no es especialmente significativo, puesto que los valores de diferencia de absorbancia correspondientes a estas variantes son altos en todas las rondas de selección (mayores de 0.5). Esto indica que hay muchas variantes de la librería del dominio PDZ-Erbin que unen a esos péptidos desde el principio de las selecciones, cumpliéndose ya desde el principio de las selecciones el objetivo perseguido, que es encontrar mutantes de Erbin que unan a cada uno de los péptidos.

Con estos ensayos ELISA se pretendía estimar únicamente el enriquecimiento en los fagos que unen a los péptidos diana frente a los cuales estaban siendo seleccionados a lo largo de las rondas. Para poder conocer en ese conjunto de fagos cuáles de ellos en concreto unen al péptido, primero hay que sembrar un cultivo de células XL1-Blue infectadas por esos fagos en una dilución tal que permita tener colonias bien diferenciadas

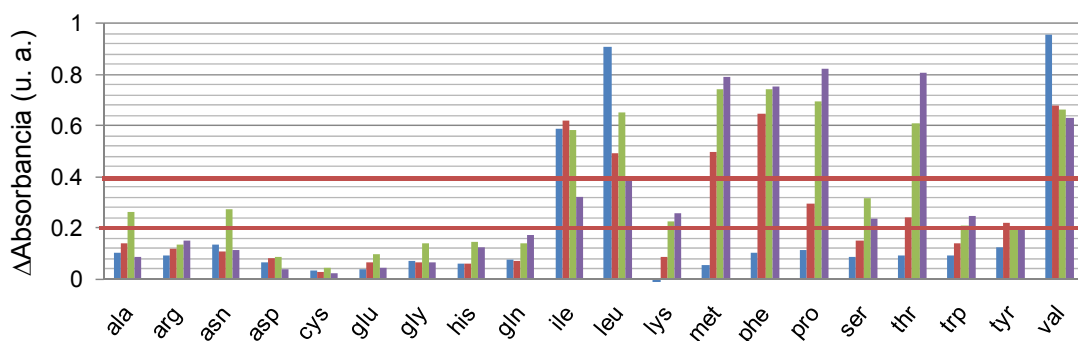


Figura 8.6. Cambio de la diferencia de absorbancia del ensayo ELISA a lo largo de las rondas. En azul está representada la ronda 2, en rojo la 3, en verde la 4 y en morado la 5. En la abscisa está escrito en código de tres letras el aminoácido que está presente en la posición 0 del péptido frente al cual se hicieron las selecciones. Las líneas rojas horizontales se han incluido como referencia para marcar los valores de absorbancia de 0.2 y 0.4.

para no tener un continuo de colonias en la placa. Cada una de estas colonias representa lo que se podría denominar una *estirpe de fagos monoparental*, en el sentido de que cada una de esas colonias contendría un fagémido con una única secuencia de la librería. A partir del crecimiento de estas colonias separadamente en un cultivo de 2YT en presencia del fago asistente M13/K07, que es un fago que permite el ensamblaje de fagos a partir de colonias de células que poseen un fagémido, se pueden amplificar los fagos correspondientes, los cuales poseen dicho fagémido, que co-expresa la variante correspondiente. Con cada conjunto de fagos, cuya secuencia va a ser única por tanto y que co-expresan una única variante de la librería, se puede hacer un ensayo ELISA para ver si hay unión de dicha variante al péptido diana frente al cual fueron seleccionados. En estos ensayos ELISA se considera como criterio general, debido a la experiencia acumulada, que hay unión al péptido cuando la diferencia de absorbancia es al menos mayor de 0.3.

La elección del número de colonias o estirpes de fagos monoparentales se hizo en función de los resultados del ensayo ELISA del conjunto de fagos seleccionados a lo largo de las rondas (Figura 8.6). Se eligió la ronda cuarta para todas las variantes, puesto que las diferencias de absorbancia del ensayo ELISA son las más altas, además de que se sabe que a partir de la tercera ronda aumentan las probabilidades de perder fagos que unen a la molécula diana a costa de aquellos que se amplifican en mayor cantidad (Clackson & Lowman, 2004). El criterio fue elegir tantas más colonias conforme el ensayo ELISA con el conjunto de los fagos proporcionó los resultados más negativos, puesto que hay menor probabilidad de encontrar fagos que unan al péptido, tal y como ya se ha explicado. Así, para aquellos que las señales de absorbancia fueron menores de 0.2 (arginina, aspártico, cisteína, glutamina, glutámico, glicina e histidina), se seleccionaron 48 colonias; cuando la señal de absorbancia fue entre 0.2 y 0.4 (alanina, asparragina, lisina, serina, triptófano y tirosina) se seleccionaron 24 y cuando la señal de absorbancia fue

mayor de 0.4 (isoleucina, leucina, metionina, fenilalanina, prolina, treonina y valina) se seleccionaron 12 colonias. Todas estas colonias (564 en total) se crecieron por separado en el medio de cultivo 2YT para obtener los fagos monoparentales correspondientes, con los que se hizo un nuevo ensayo ELISA.

La primera conclusión de este ensayo es que se han seleccionado variantes únicas del dominio PDZ-Erbin que unen a cada una de las veinte variantes del péptido **TGWETW \mathbf{X}** , ya que aparecen estirpes de fagos monoparentales seleccionadas frente a cada uno de los péptidos que dan una diferencia de absorbancia del ensayo ELISA superior a 0.3 en todos los casos (Figura 8.7). Además, los resultados de este ensayo revelan que la frecuencia de variantes únicas del dominio PDZ-Erbin que unen al péptido frente al cual han sido seleccionadas está en función del aminoácido que esté presente en la posición 0 del péptido. Así, cuando se seleccionaron 12 colonias, al menos el 60% de los fagos seleccionados presentan unión (panel A de la Figura 8.7). Por otro lado, cuando se seleccionaron 24, la proporción fue menor y bastante variable en función del péptido frente al cual habían sido seleccionadas las variantes. Sólo se encontró una variante del dominio PDZ-Erbin que une al péptido **TGWETW \mathbf{A}** , mientras que hubo hasta ocho variantes que unen a los péptidos **TGWETW \mathbf{Y}** y **TGWETW \mathbf{W}** (panel B de la Figura 8.7). Finalmente, cuando se eligieron 48 colonias, hubo alrededor de un 50% de estos fagos que unen, excepto los que se seleccionaron frente a **TGWETW \mathbf{G}** , que hubo hasta un 75% (paneles C y D de la Figura 8.7).

8.1.3.1 Selección de las variantes del dominio PDZ-Erbin más específicas mediante un proceso de competición

Hasta este punto, para cada uno de los veinte péptidos fusionados a GST se habían conseguido identificar variantes del dominio PDZ-Erbin afines al péptido frente al cual habían sido seleccionadas. Sin embargo, el hecho de que presenten afinidad no implica que esa unión sea específica a ese péptido. El objetivo perseguido es obtener unas variantes del dominio PDZ-Erbin que unan con la mayor especificidad posible el aminoácido de la posición 0, es decir, que por la secuencia de aminoácidos que presenta el dominio pueda distinguir uno de los 20 aminoácidos frente al resto. Para conseguir esto, se realizaron más rondas de selección incluyendo un proceso de competición.

Como ya se ha mencionado, en los estudios previos se ha visto que todos los aminoácidos presentes en la posición 0 del ligando son apolares, excepto la cisteína. Además, los más frecuentes en la naturaleza son los aminoácidos isoleucina, leucina y valina (Tonikian et al., 2008). Debido a esto, se eligieron para la competición los péptidos **TGWETW \mathbf{I}** , **TGWETW \mathbf{L}** y **TGWETW \mathbf{V}** fusionados a GST ya que por su naturaleza podrían unir mejor que el resto de los péptidos, incluso aunque la variante del dominio PDZ-Erbin haya sido seleccionada frente a un péptido con otro aminoácido en su posición 0.

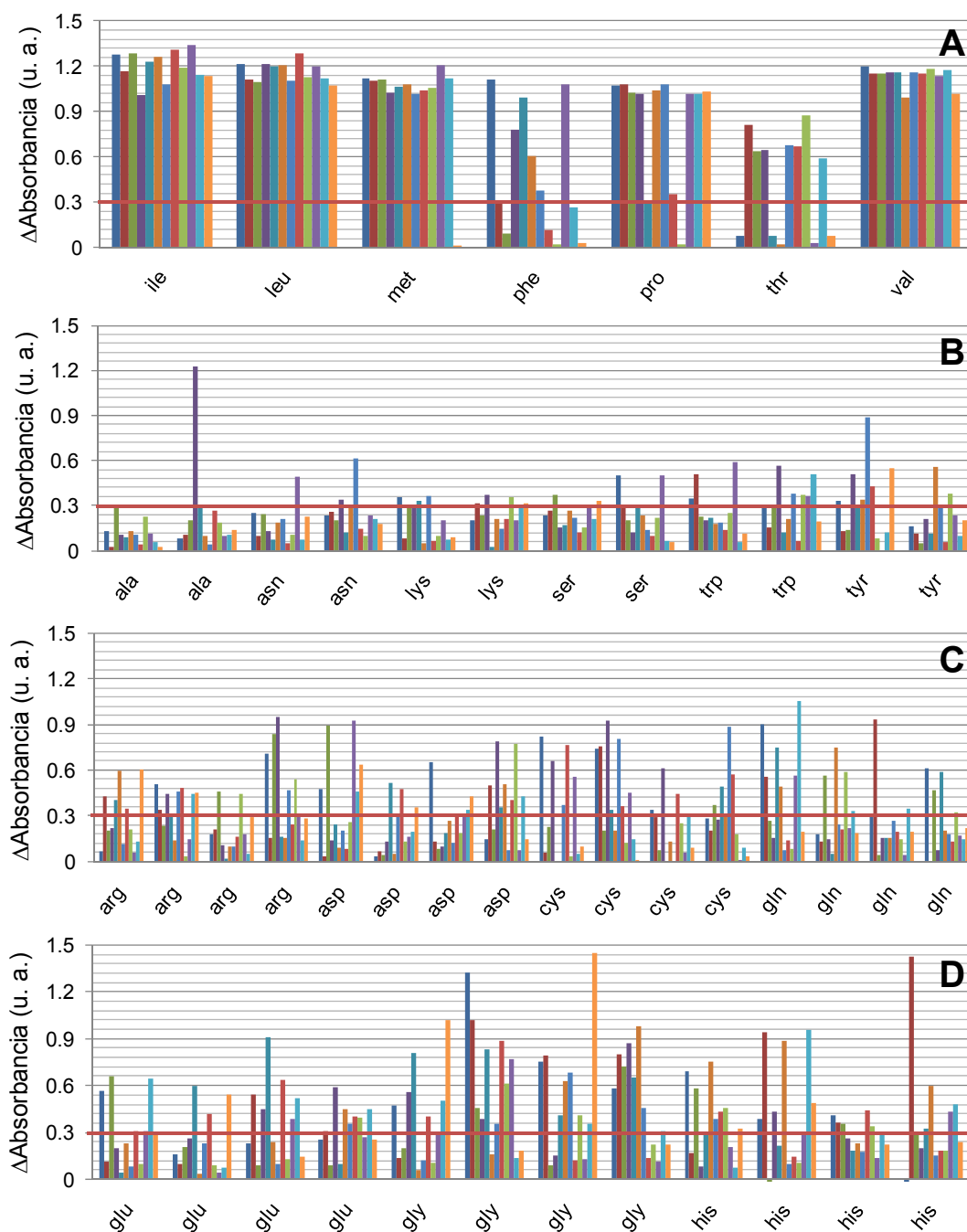


Figura 8.7. Análisis de estirpes de fagos monoparentales seleccionados frente a cada uno de los péptidos mediante ensayo ELISA. Como se indica en el texto, en función del péptido frente al cual se hicieron las selecciones, se eligieron 12 (panel A), 24 (panel B) o 48 (paneles C y D). Debajo del conjunto de fagos seleccionados se indica el aminoácido que está presente en la posición 0 del péptido frente al cual se hicieron las selecciones. La línea roja horizontal se ha incluido como referencia para marcar el valor de absorbancia de 0.3.

La forma de hacerlo fue tomar los fagos que se habían seleccionado en la tercera ronda de las selecciones previas y no de la cuarta, por la razón ya citada de que a partir de la tercera aumenta la probabilidad de que haya un exceso de fagos que tengan mayor facilidad de proliferación, además de que se busca el compromiso entre tener una diversidad elevada de variantes junto con la eliminación de la mayoría de variantes que no unen a la molécula diana, algo que ocurre a partir de la tercera ronda (Sidhu, 2005). De esta forma, se hicieron tres rondas más en las que se incluyó una etapa de incubación de las variantes del dominio PDZ-Erbin con esos tres péptidos en una concentración final de 1 μM . Así se conseguiría ejercer una presión selectiva mayor sobre las variantes ya seleccionadas. Para el caso de las nuevas selecciones con el péptido que presenta leucina en la posición 0, la mezcla competidora fueron los péptidos que presentan isoleucina y valina, de nuevo en una concentración total de 1 μM ; para las nuevas selecciones con el péptido que presenta isoleucina en la posición 0 la mezcla competidora fueron los péptidos que presentan leucina y valina, en una concentración total de 1 μM ; finalmente, para las nuevas selecciones con el péptido que presenta valina en la posición 0 la mezcla competidora fueron los péptidos que presentan isoleucina y leucina, en una concentración total de 1 μM .

En este caso, no se siguió ninguna de las tres rondas realizadas incluyendo el proceso de competición con un ensayo ELISA del conjunto de los fagos seleccionados, puesto que ya se había visto que el proceso funcionaba y únicamente interesaba encontrar estirpes de fagos monoparentales que unan específicamente al péptido frente al cual han sido seleccionados. Por eso, al acabar la tercera ronda con competición, se sembraron en placa cultivos con el conjunto de los fagos seleccionados para poder seleccionar colonias independientes, que, como se ha mencionado, representan a una estirpe de fagos monoparental. Así, se eligieron 48 colonias para cada una de las variantes de los péptidos. Los resultados de este ensayo ELISA a partir de estas 960 colonias están recogidos en la Figura 8.8.

Estos resultados muestran que ninguna de las 48 variantes del dominio PDZ-Erbin seleccionadas frente a los péptidos que tienen en la posición 0 los aminoácidos alanina, arginina, asparragina, aspártico, cisteína, glutámico, glicina, glutamina, lisina, serina, triptófano y tirosina, da una diferencia de absorbancia mayor de 0.3 (Figura 8.8). Ya se ha descrito que en las selecciones en que no había una etapa de incubación con los péptidos TGWETWI, TGWETWL y TGWETWV, el ensayo ELISA de las colonias seleccionadas mostró que a partir de algunas de éstas se obtenían valores de diferencia de absorbancia superiores a 0.3 (Figura 8.7). Esto implica que no se unieron al péptido frente al cual se habían seleccionado inicialmente y sí a los péptidos con los que se realizó la competición, de esta forma, se eliminaron en la etapa de lavado y no se seleccionaron para los siguientes pasos. Así, se podría decir que estas variantes unen con mayor afinidad a los péptidos con los que se realizó la competición que aquellos con los que habían sido seleccionadas inicialmente.

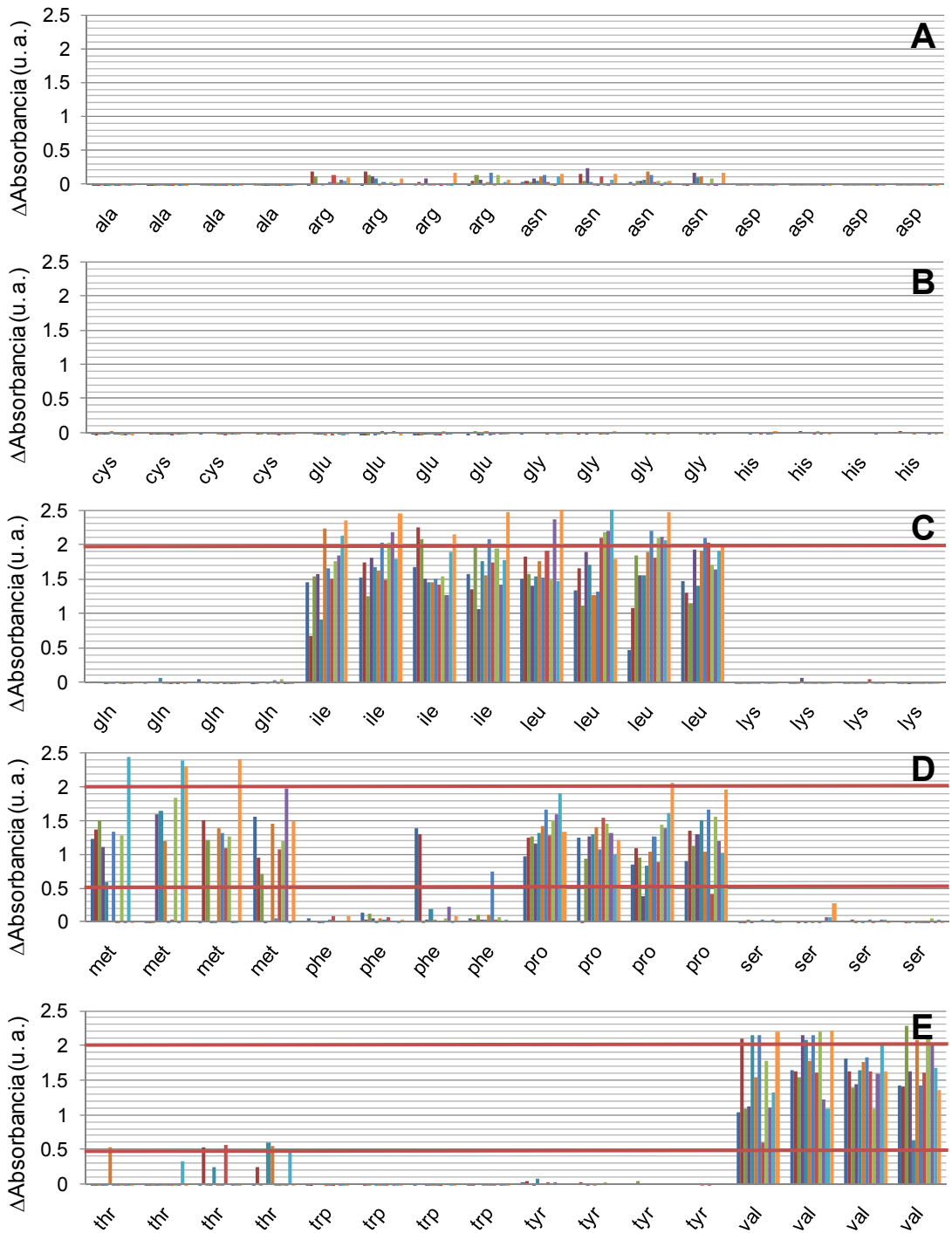


Figura 8.8. Análisis de estirpes de fagos monoparentales seleccionados en la tercera ronda hecha incluyendo el proceso de competición frente a cada uno de los péptidos mediante ensayo ELISA. Como se indica en el texto, hay 48 estirpes de fagos monoparentales para cada uno de los péptidos frente a los que fueron seleccionados, los cuales que aparecen agrupados en la Figura de 12 de 12. Debajo de cada uno de estos grupos de 12 se indica el aminoácido que está presente en la posición 0 del péptido frente al cual se hicieron las selecciones.

En cambio, sí ha habido variantes del dominio PDZ-Erbin que siguen uniendo al péptido frente al cual fueron seleccionadas cuando se incluye la competición, en concreto los que presentan isoleucina, leucina, metionina, fenilalanina, prolina, treonina y valina en la posición 0 del péptido. Esto se deduce al observar que en los paneles C, D y E de la Figura 8.8 de entre las 48 colonias que se seleccionaron, hay algunas a partir de las cuales se obtuvieron valores de diferencia de absorbancia en el ensayo ELISA mayores de 0.3. Además, los valores de diferencia de absorbancia son más altos que los de las variantes seleccionadas sin el proceso de competición. De esta forma, se puede concluir que estas variantes seleccionadas presentan una mayor selectividad con respecto a los péptidos con los que se hizo la competición ya que unen al péptido frente al cual están siendo seleccionadas aun en la presencia de otros a los que también podrían unirse.

8.1.3.2 Secuenciación de las variantes del dominio PDZ-Erbin seleccionadas

Llegados a este punto, se ha de destacar que se han realizado dos procesos de selección. Por un lado, el inicial en el que se realizaron cinco rondas; y, por otro, aquel en el que, con los fagos seleccionados de la tercera ronda de la selección inicial, se hicieron tres rondas más en las que se incluyó en cada una de ellas un proceso de competición. Los ensayos ELISA realizados a partir de las colonias seleccionadas, que son las que al fin y al cabo representan variantes únicas del dominio PDZ-Erbin de la librería, indicaron que hay variantes que unen a todos y cada uno de los péptidos. Por otro lado, cuando se introdujo el proceso de competición, los ensayos ELISA indican que muchas de las variantes unen mejor a los péptidos con los que se realizó la competición, por lo cual se eliminan en los lavados. Todo esto hay que tenerlo en cuenta para la secuenciación de las variantes del dominio PDZ-Erbin, puesto que ésta se hace según estos resultados provenientes de las selecciones.

De esta forma, se hicieron dos bloques distintos de secuenciación. Por un lado, aquellas variantes seleccionadas sin proceso de competición; por otro, aquellas que se seleccionaron con proceso de competición. Esto se hizo porque se quería conocer la secuencia de las variantes del dominio PDZ-Erbin que habían sido seleccionadas frente a cada uno de los péptidos, independientemente de que sean más específicas o menos. Debido a esto, en los casos en los que la competición no fue válida, se secuenciaron aquellas que habían sido seleccionadas sin realizar dicho proceso de competición. Por otro lado, las variantes seleccionadas mediante el proceso de competición se secuenciaron, puesto que se unen específicamente al péptido frente al cual han sido seleccionadas. A la hora de elegir qué secuenciar, se establecieron varios límites en los valores de diferencia de absorbancia de los ensayos ELISA para tener un número homogéneo de variantes seleccionadas frente a cada uno de los péptidos.

Las secuencias de aquellas variantes seleccionadas sin el proceso de competición, que son aquellas que se seleccionaron frente a los péptidos que tienen en su posición 0 los aminoácidos alanina, arginina, asparragina, aspártico, cisteína, glutámico, glicina, glutamina, lisina, serina, triptófano y tirosina están recogidas en la Tabla 8.1. Para su

secuenciación, se tuvo en cuenta que, tal y como se ha mencionado, el valor mínimo de diferencia de absorbancia que se consideró fue de 0.3, por lo que no se secuenció ninguna variante con un valor menor. Debido a esto, sólo se secuenció una variante del dominio PDZ-Erbin seleccionada frente al péptido TGWETWA, que da un valor de diferencia de absorbancia de 1. En el resto de las variantes, se eligieron tres valores de absorbancia de tal forma que se secuenciase un número homogéneo de éstas: 0.5 para las variantes seleccionadas frente a los péptidos que tienen en su posición 0 arginina, glutámico, histidina, triptófano y tirosina; 0.75 para las variantes seleccionadas frente a los péptidos que tienen aspártico, cisteína, glutamina y glicina; 1 para las variantes seleccionadas frente al péptido que tiene fenilalanina.

Por otro lado, respecto a las variantes que se seleccionaron mediante competición, aquellas que fueron seleccionadas frente a los péptidos que tienen en la posición 0 isoleucina, leucina, metionina, prolina y valina, se eligieron cuando en el ensayo ELISA el valor de absorbancia fue superior a 2. En el caso de las que se seleccionaron frente a los péptidos que tienen treonina y fenilalanina, puesto que ninguna llega a ese valor, se decidió poner un límite más bajo, 0.5. Las secuencias de todas estas variantes están recogidas en la Tabla 8.2.

Finalmente, las secuencias obtenidas, agrupadas por los péptidos frente a los que fueron seleccionadas, se procesaron con el programa Weblogo (Crooks et al., 2004), que es muy útil para poder visualizar perfiles consenso. La altura de un aminoácido en el *logo de secuencia* que genera el programa indica la frecuencia relativa con que aparece en el conjunto de todas las secuencias en esa posición. De esta forma, las secuencias recogidas en las Tablas 8.1 y 8.2 se pueden analizar de forma más intuitiva en la Figura 8.9.

Tabla 8.1. Secuencias de los mutantes del dominio PDZ-Erbin seleccionados frente a las variantes del péptido TGWETW~~X~~ sin el proceso de competición.

X	<i>Región β1-β2</i>	<i>α2</i>	X	<i>Región β1-β2</i>	<i>α2</i>
Ala	LHRRSFGF	LR	Gly	TSYRGPGS	RY
	KRFRWFGS	RC		LLLSETGS	RM
	NTVRGFGT	GR		YSRRLFSGS	RY
Arg	NRFRGFGT	GW		GMRHFSGS	VW
	TRPFIFGS	VY		PKGVTSGP	RY
	WPHARIGI	LR		LILYRVGF	RC
	AQRKAFGF	YV		GMRHFSGS	LR
Asn	VVASIIGT	CV		SKRIMSGA	VY
	KRYWLPGP	YV		LHIYYIGF	RC
	KRYWLPGP	RC		RRWLRPGT	RY
Asp	SLFHTPGT	LY		NSHDRVGS	VY
	KRFRVSGS	HY		LFANGFGP	QA
	RFGWVGV	RY		RTGDAAGA	VL
	WIWRLSGT	RV		TNCCASGT	QA
	RFLSPPGV	YA		KSYSAGT	VY
Cys	RIPNYIGA	VY	RYKIRIGF	KC	
	LNIFTFGA	WP	LFFTSIGS	RW	
	GLGRRFGS	RY	IYFNRFPG	VY	
	GGHGKVG	KT	FDKPKVGA	TY	
	IYLWYSGS	CV	LDKPKVGA	TY	
	LGAGEVGA	YV	LTMWRIGT	RL	
	AQRKAFGF	FR	AQRKAFGL	FR	
Gln	FTFQKSGS	RY	FTRYHFGV	GV	
	AQRKAFGF	RC	ENFYIFGL	CI	
	LVKRRTGT	VY	ENFYIFGL	CV	
Glu	PIRSCVGP	FV	RRLRWPGT	VY	
	HNYLSAGP	MY	CLMIYPGL	YV	
	PNFWVGP	VY	NLHESLGP	YW	
	YRFIYSGA	YA	RSELDLGS	RL	
	RRVMYLG	YV	LRPLVFGV	RL	
Tyr	HNYLSAGP	YV	RLGGWVGS	YV	
	QNRRLTGF	GY	LTASHSGT	MW	
	RRVMYLG	RY	APLNLFG	VY	
	SRKRWSGA	YA	WTKRRLGP	VY	
	QNRRLTGF	RY	LFFVGFGA	RI	
	SRKRWSGA	MY	MMFGSFGP	VY	

Tabla 8.2. Secuencias de los mutantes del dominio PDZ-Erbin seleccionados frente a las variantes del péptido TGWETW \mathbf{X} obtenidos con el proceso de competición.

X	Región $\beta 1$ - $\beta 2$	$\alpha 2$
Ile	GSPRGFGF	LR
	ASRGRFGF	LR
	VPSSRFGF	LR
	TTQRNFGF	LR
	FPPRGFGF	LR
	LDRGRFGF	LR
	RVHSRFGF	LR
	PRGSRFGF	LR
	TQPRNFGF	LR
	SLKRNFGF	LR
	LGRGRFGF	LR
	RARRDFGF	LR
	THRRGFGF	LR
	Leu	HRPCGLGF
PRSPRFGF		FR
PGGRYFGF		MR
VPRKYFGF		IN
HRPCGLGF		LC
GRDRKFGF		FR
WRPCGIGA		LC
NRNAVFGF		FR
GRDRKFGF		IR
SMKPRFGF		IR
PRKGCFGF		IR
RSRPVFGF		FK
ANRGRFGF		FR
Met		CTRPCFGF

X	Región $\beta 1$ - $\beta 2$	$\alpha 2$
Met	CTRACFGF	IR
	CLRRCFGF	MR
Phe	GCRPWAGF	WC
	RWCPWAGF	WC
Pro	RWCPWAGF	WK
	NVRGRFGF	IR
Thr	RVRGRFGF	LR
	WPRARAGF	LR
	SARGCLGV	LR
	RRKGCLGA	LR
	RARGCLGV	LR
	WPHARIGN	LR
	CMSKCLGI	LR
	GDRRSFGF	MR
Val	MCRKTFGF	LR
	GSGRMLGF	LR
	RKGGCIGV	LR
	CPRGYLGI	LR
	CGRRLFGE	KG
	RTRASLGF	LR
	PSRPGFGF	LR
	CPRGYLGI	MG
	MTGGLFGF	FR
	MNRPKFGF	LR
VPRRTFGF	LR	
FCGARLGF	LR	



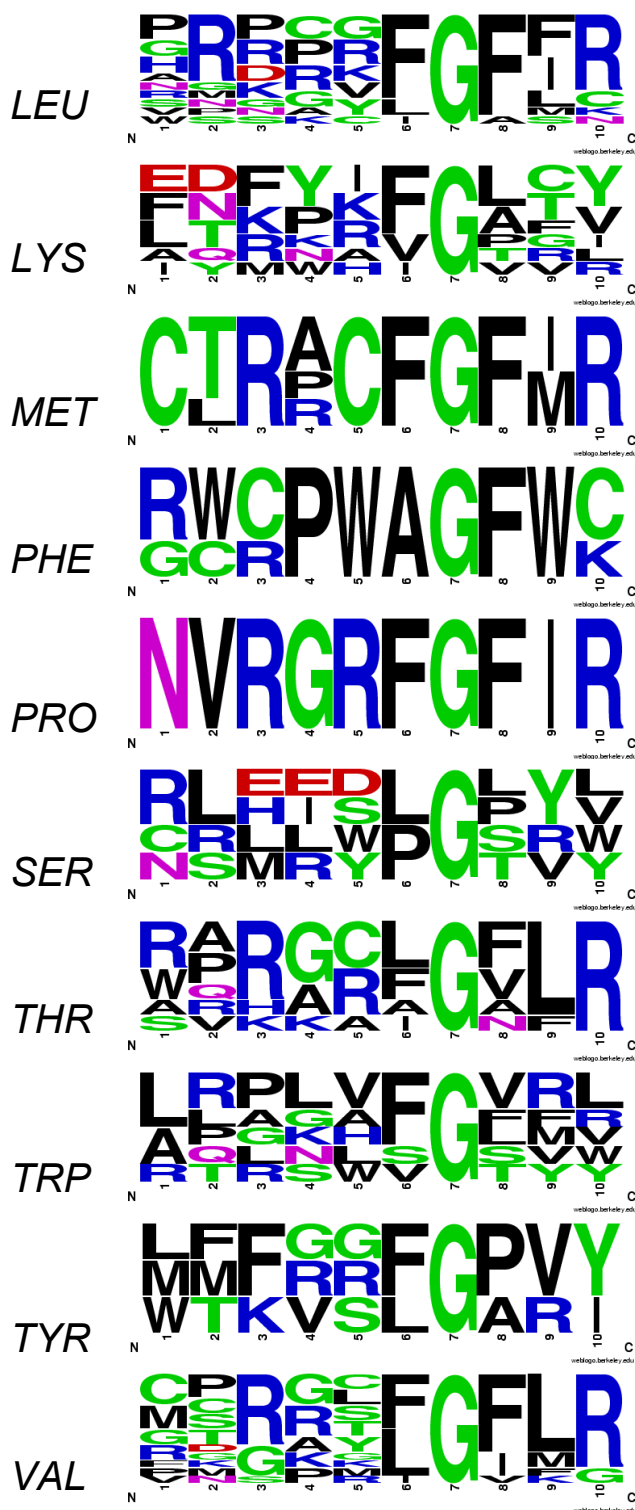


Figura 8.9. Figuras generadas por el programa Weblogo (Crooks et al., 2004) al introducir las secuencias de los mutantes del dominio PDZ-Erbin seleccionados frente a las veinte variantes del péptido TGWETWV. Junto a cada figura está indicado el aminoácido en la posición 0 del péptido correspondiente.

8.1.4 Selección de variantes del dominio PDZ-Erbin que unen secuencias peptídicas en un reconocimiento de tipo interno

Se usó la misma librería de mutantes del dominio PDZ-Erbin para las selecciones frente a las variantes del péptido TGWETWVXGGGKK para ver si existía en dicha librería algún mutante que presentase reconocimiento de tipo interno. Se hicieron de nuevo cinco rondas de selección cuya evolución se siguió mediante un ensayo ELISA con el conjunto de los fagos seleccionados en cada ronda para saber si se producía un enriquecimiento en los fagos que unen a los péptidos diana. En este caso las diferencias de absorbancia entre los ensayos con el péptido unido a GST y la proteína GST fueron mayores de 0.3 a lo largo de las rondas con el conjunto de los fagos seleccionados únicamente frente a los péptidos que presentan metionina, fenilalanina, triptófano y tirosina en la hipotética posición +1 (Figura 8.10). En el resto, los valores de absorbancia fueron menores.

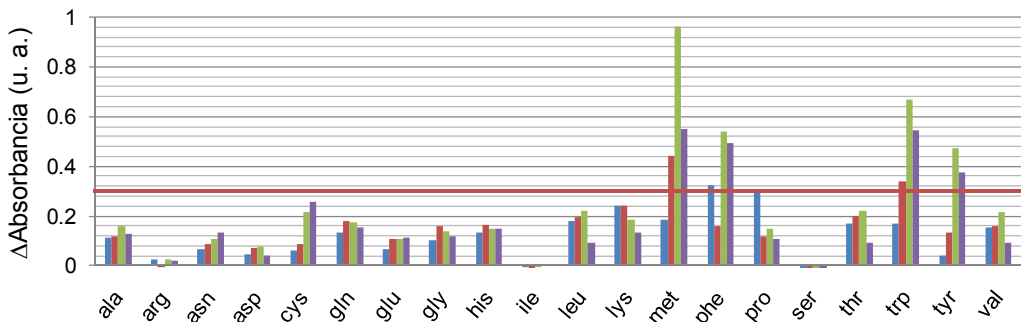


Figura 8.10. Cambio de la diferencia de absorbancia del ensayo ELISA realizado con el conjunto de los fagos seleccionados a lo largo de las cinco rondas realizadas frente a las veinte variantes del péptido TGWETWVXGGGKK. En azul está representada la ronda 2, en rojo la 3, en verde la 4 y en morado la 5. Debajo de cada conjunto de valores está indicado el aminoácido de la posición X del péptido frente al cual se hicieron las selecciones. La línea roja horizontal se ha incluido como referencia para marcar el valor de absorbancia de 0.3.

Se sembraron en placas cultivos infectados con los fagos seleccionados en la ronda 4 para poder hacer un ensayo ELISA a partir de los fagos monoparentales obtenidos de sus respectivas colonias. Se eligieron 24 colonias para cada uno de los péptidos frente a los cuales fueron seleccionadas (480 en total). De estas 24 variantes analizadas, únicamente algunas de las seleccionadas frente a los péptidos que tienen de nuevo metionina, fenilalanina, triptófano y tirosina en la posición +1, unen al péptido correspondiente, puesto que presentan incrementos de absorbancia mayores de 0.3 (Figura 8.11). Del resto de variantes seleccionadas no se obtuvo ninguna cuya diferencia de absorbancia llegase a 0.3, por lo que se puede concluir que no hay variantes del dominio PDZ-Erbin que se unan a esos péptidos.

Se secuenciaron las variantes del dominio PDZ-Erbin que sí unen a los péptidos con metionina, triptófano y tirosina y que presentan una absorbancia máxima, además de las cuatro mejores seleccionadas frente al que tiene fenilalanina (Tabla 8.3). Se obtuvieron asimismo con estas secuencias los respectivos *logos de secuencia* mediante el programa Weblogo (Figura 8.12).

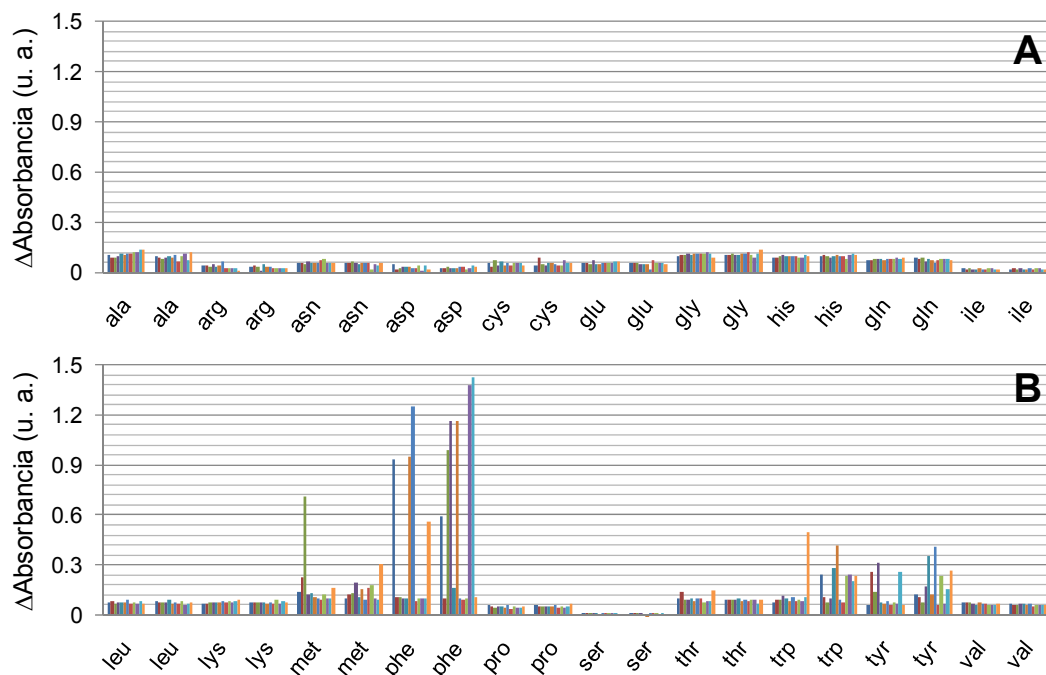


Figura 8.11. Análisis de estirpes de fagos monoparentales seleccionados frente a cada uno de los péptidos que mimetizan el reconocimiento de tipo interno mediante ensayo ELISA. Debajo del conjunto de valores obtenidos a partir de los clones seleccionados se indica el aminoácido que está presente en la hipotética posición +1 del péptido frente al cual se hicieron las selecciones.

Tabla 8.3. Secuencias de los mutantes del dominio PDZ-Erbin seleccionados frente a las variantes del péptido TGWETWVXGGGKK que dieron unión específica.

X	<i>Región β1-β2</i>	α 2
Met	AQRKAFGF	FR
Phe	LTRWSLGF	FT
	LTGKCLGF	LR
	LWKKTVVL	FR
	PRSGGFGF	FR
Trp	LTGWPLGF	FR
Tyr	LTRWSLGF	FT



Figura 8.12. Figuras generadas por el programa Weblogo (Crooks et al., 2004) al introducir las secuencias de los mutantes del dominio PDZ-ErbB seleccionados frente a las variantes del péptido TGWETWVXGGGKK que dieron unión específica. Junto a cada Figura está indicado el aminoácido en la posición 0 del péptido correspondiente.

8.2 Discusión

El conjunto de secuencias generadas por selección de fagos abre las vías para una comprensión a nivel molecular de la especificidad de la familia de los dominios PDZ. Tal y como se explica en el Capítulo 2, las familias de dominios modulares son claves en el funcionamiento celular y es especialmente interesante conocer cómo varían las propiedades de unión de una forma tan mínima, a nivel de pocos residuos. Según los estudios mutacionales que acompañaron al mapa de especificidad de dominios PDZ naturales, se demostró que sería posible determinar la especificidad de un dominio cualquiera sabiendo su secuencia aminoacídica, puesto que parecen suficientes determinadas posiciones en ciertas regiones del dominio para seleccionar específicamente diferentes aminoácidos en las distintas posiciones del ligando (Tonikian et al., 2008).

Con los nuevos mutantes seleccionados en esta Memoria para el dominio PDZ-ErbB se han rastreado prácticamente todas las posibilidades que se puedan dar en la naturaleza, por lo que, extendiendo los estudios iniciados a las posiciones -1, -2 y -3, se podría tener un conocimiento a nivel molecular de las causas de las distintas clases de especificidad. Esta información, debidamente racionalizada con herramientas computacionales permitirá establecer la especificidad de cualquier dominio PDZ a partir de su secuencia aminoacídica, o, dicho en otras palabras, permitirá conocer *in silico* las preferencias por determinadas secuencias aminoacídicas.

Ya se ha hecho un primer intento de este tipo de bases de datos que intentan predecir especificidades en los dominios PDZ (Smith & Kortemme, 2010). En ese caso, basados en las estructuras de los diferentes dominios, mediante algoritmos computacionales se han predicho con éxito los perfiles de especificidad de 17 dominios PDZ presentes en la naturaleza. Se pudo comprobar que los resultados son fiables mediante la comparación con los datos experimentales obtenidos en el mapa de especificidad previamente establecido (Tonikian et al., 2008). La gran limitación de este método es que se requiere tener la estructura del dominio; de hecho, el número predicho en el trabajo descrito es el número de estructuras de tales dominios que hay en la base de datos del *Protein Data Bank* de los cuales se obtuvieron los perfiles de especificidad de forma experimental.

Sin embargo, la información que se podrá obtener a partir del trabajo iniciado en esta Memoria será de una naturaleza distinta, ya que la información es obtenida directamente de la secuencia aminoacídica, no de la estructura del dominio. Por lo tanto, toda la información obtenida de los experimentos de selección, debidamente completada, permitirá, mediante un cribado de los genomas, determinar y localizar nuevas interacciones naturales de interés y hacer nuevos descubrimientos sobre los sistemas de señalización en los que estos dominios son clave. Se podría usar incluso con fines terapéuticos, ya que se ha visto que los dominios PDZ son proteínas diana ideales para poder modular las interacciones patógenas proteína-proteína que se dan en los mecanismos de algunas enfermedades en los que dichos dominios están implicados (Dev, 2004).

Hasta ahora, para conocer las diferencias de especificidad de unión entre estos dominios ha habido que realizar estudios a nivel molecular y estructural. Así, por ejemplo, se realizó un estudio en el que se comparó el dominio PDZ-Erbin, cuyo ligando óptimo presenta en la posición 0 una valina, con el dominio PDZ1 de la proteína ZO-1, en cuyos ligandos son preferidas la isoleucina o la leucina en dicha posición (Appleton et al., 2006). La interacción del grupo carboxilato del extremo C-terminal del péptido es básicamente la misma en los dos dominios. En ella, la carga negativa es coordinada por los protones amida en la zona del motivo GLGF, altamente conservado en los dominios PDZ. Sin embargo, las interacciones que se dan entre las cadenas laterales son distintas, debido a que presentan residuos diferentes. En el primer residuo que forma parte de la lámina β_2 (β_2 -1) del dominio PDZ-Erbin hay una fenilalanina (Figura 8.13), mientras que en la posición equivalente del dominio PDZ1 de ZO-1 hay una leucina. Ello explica las diferencias preferenciales tienen los dos dominios en la unión al ligando en su posición 0, puesto que el mayor tamaño de la fenilalanina hace que no estén favorecidos aminoácidos con cadenas laterales tan grandes como la leucina o la isoleucina y hace que se acomode mejor una valina. En cambio, en el dominio PDZ1 de ZO-1, como tiene una leucina, que es menos voluminosa que la fenilalanina, se pueden acomodar los residuos un poco más grandes.

Con la información recogida en las Tablas 8.1 y 8.2 se pretenden obtener este tipo de datos que determinan la especificidad de unión de estos dominios. De hecho, ya se puede obtener algo de información. Respecto al residuo presente en la posición β_2 -1 que

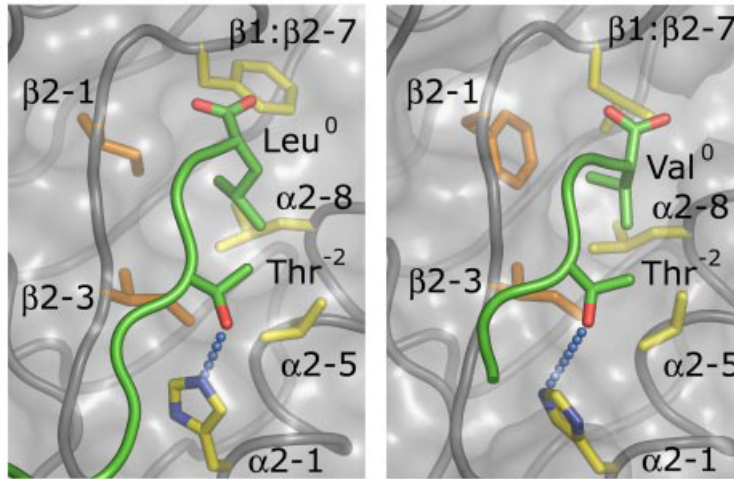


Figura 8.13. Detalle de la interacción entre los residuos del dominio PDZ que interaccionan con los aminoácidos presentes en los sitios 0 y -2 de los ligandos WRRTTYL y TGWETWV unidos respectivamente a los dominios PDZ1 de la proteína ZO-1 (A) y PDZ-Erbin (B). Figura tomada de (Appleton et al., 2006).

se ha mencionado en el párrafo anterior, se observa que cuando el péptido frente al cual se seleccionan las variantes tiene valina en la posición 0, las variantes del dominio PDZ-Erbin tienen en $\beta 2-1$ tanto fenilalanina (un 50% de los mutantes) como isoleucina y valina (cada aminoácido está presente en un 25% de los mutantes aproximadamente). Llama aún más la atención que cuando los péptidos frente a los cuales se hacen las selecciones tienen en la posición 0 leucina o isoleucina, los mutantes del dominio PDZ-Erbin presentan en la posición $\beta 2-1$ mayoritariamente fenilalanina.

En un principio, estos resultados pueden parecer contradictorios con lo que ya se ha explicado sobre las características de especificidad que confiere el aminoácido en posición $\beta 2-1$ de un dominio PDZ. Hay que tener en cuenta que, en el caso presente en esta Memoria, se está variando toda la región del bolsillo hidrofóbico que está interaccionando con el ligando en la posición 0, por lo que se están seleccionando los mejores mutantes, pero entendidos globalmente. Esto quiere decir que se está afectando la zona de unión en su conjunto. Por eso, aunque en un principio una leucina en la posición $\beta 2-1$ del dominio PDZ favorecería la unión a isoleucina en la posición 0 del ligando por razones de tamaño, al estar afectando más residuos de esa región, ese efecto se puede compensar por otros restos y seleccionar una fenilalanina en esa posición del dominio. Podría estar ocurriendo que las interacciones de tipo apolar que se dan se vean más favorecidas con la cadena lateral de la fenilalanina y que el resto de los aminoácidos del bolsillo de unión sean menos voluminosos para que pueda estar presente la fenilalanina.

De hecho, ya se ha demostrado en los dominios PDZ que cuando se muta la zona de unión buscando los mejores residuos del dominio para una interacción determinada, la nueva configuración de la región provoca que haya propiedades de unión diferentes (Ernst et al., 2010). Así, los mutantes seleccionados han de estudiarse tanto particularmente

como en su conjunto, por lo que los resultados obtenidos en esta Memoria han de completarse con un estudio más detallado de los mutantes seleccionados. De hecho, hemos purificado ya todas las variantes del dominio PDZ-Erbin que se han secuenciado y se están obteniendo los perfiles de especificidad con librerías peptídicas para cada una de ellas, para ver cómo ha variado su especificidad de unión al cambiar la configuración del bolsillo hidrofóbico. De hecho, una vez obtenidos estos perfiles de especificidad se les asignará una Matriz de Peso de cada Posición (*Position Weight Matrix, PWM*) (Obenauer et al., 2003). Cada columna de esta matriz representa la preferencia del aminoácido para unir en cada una de las posiciones del ligando como una distribución de probabilidad. Por ejemplo, se obtienen diez ligandos peptídicos para una variante de Erbin, y en esos diez ligandos aparecen dos tipos de aminoácidos en la posición 0, uno 8 veces y otro 2 veces. Eso correspondería a los valores 0.8 y 0.2 en la columna de la matriz para esos dos aminoácidos. Una vez construidos estos PWM, se podrá asignar un potencial de especificidad (*specificity potential, SP*) a cada una de las variantes, cuyo valor va desde nada selectivo (puede reconocer cualquier aminoácido, $SP = 0$) hasta específico (reconoce un único aminoácido, $SP = 1$) (Tonikian et al., 2008). Con estos datos, probablemente se podrán elucidar reglas generales para conocer la especificidad en función de la secuencia, porque cuando más alto sea el SP de la variante, más específica será, y se podrá relacionar con la secuencia de ésta. Además, se pretende hacer un estudio termodinámico-estructural de aquellas variantes con los valores más altos de SP, que permitirá la racionalización de los resultados.

Por otro lado, en la Sección 8.1.4 se describe cómo hay variantes del dominio PDZ-Erbin que unen frente a las variantes del péptido TGWETWV**X**GGGKK que presentan los aminoácidos metionina, fenilalanina, triptófano y tirosina en la posición **X** del ligando. Ya que este péptido mimetiza interacciones de tipo interno, se deduce que existen variantes de este dominio PDZ capaces en principio de presentar reconocimiento de este tipo, lo cual es significativo en sí mismo por las propiedades de unión generales de estos dominios.

De hecho, lo es más aún si se tiene en cuenta una característica común a los dos casos de reconocimiento interno con una posición +1 descritos en bibliografía (Penkert et al., 2004; Zhang et al., 2009). En ambos, hay un resto aspartato en el péptido cuyo grupo carboxilato de su cadena lateral presenta una interacción equivalente a la que tiene el grupo carboxilato del extremo C-terminal en la interacción clásica de estos dominios. De hecho, tal y como se ha diseñado el péptido frente al cual se han hecho las selecciones (TGWETWV**X**GGGKK) en un principio no se debería dar esa interacción equivalente, puesto que el único resto que podría darla, el glutámico de la hipotética posición -3, estaría muy alejado del bolsillo hidrofóbico. Ahora bien, se hace necesario un estudio más detallado de los mutantes para clarificar esto. En concreto sería necesario obtener los perfiles de especificidad con librerías peptídicas de las siete variantes secuenciadas, para ver si son capaces de establecer un reconocimiento de tipo interno y también de reconocer péptidos con el extremo C-terminal libre, tal y como se da en los dominios PDZ Dvl (Zhang et al., 2009). También sería necesario obtener algunas de las estructuras de las siete variantes con los distintos péptidos obtenidos de las selecciones con librerías peptídicas,

para poder averiguar si se produce una deformación del bolsillo hidrofóbico, y confirmar así una vez más la plasticidad que estos dominios tienen. De esta forma, se podría seguir confirmando la característica de estos dominios como sistemas en los que se consigue el compromiso existente entre la rigidez de una especificidad única y el caos de una promiscuidad falta de selectividad, que les ha hecho ser dominios clave en la organización celular como las sinapsis, uniones celulares, etc.

CAPÍTULO 9

RESUMEN Y CONCLUSIONES

El trabajo presentado en esta Memoria tenía como objetivo el estudio detallado del equilibrio conformacional del tercer dominio PDZ de la proteína PSD95, con el interés de abordarlo por ser un dominio modular de una proteína *hub*, además de por ser el primer estudio termodinámico-estructural del equilibrio conformacional que se ha hecho hasta la fecha de un dominio PDZ. Por otro lado, también se ha tratado de desarrollar un protocolo mediante la metodología de expresión en fagos para poder avanzar en el conocimiento de la especificidad de unión de ligandos de los dominios PDZ, lo que permitirá predecir sus propiedades funcionales a partir únicamente su secuencia aminoacídica.

De forma más detallada, el trabajo se resume en:

- I. Se clonó y purificó el dominio PDZ3-PSD95, al cual se le buscaron condiciones de cristalización mediante el uso de las técnicas de gota colgante y de contradifusión. Una vez obtenidos los cristales, se difractaron y se resolvió la estructura en dos de las condiciones usadas. Estos resultados se han publicado en la revista *Journal of Structural Biology* (Camara-Artigas et al., 2010).
- II. Se realizó un estudio del equilibrio conformacional del dominio PDZ3-PSD95 mediante DSC en condiciones de pH neutro. Puesto que los resultados de DSC revelaron la presencia de un intermedio de equilibrio que presenta fenómenos de asociación-disociación, según se dedujo a partir de la dependencia de las trazas calorimétricas en función de la concentración, se hicieron experimentos de DLS, SEC, CD y fluorescencia de ANS y ThT para caracterizar dicho intermedio. Además, éste resultó ser un precursor de formación de estructuras anulares y fibrilares, tal y como revelaron las diferentes espectroscopias y las micrografías realizadas mediante TEM. Todos estos resultados han sido objeto de una publicación en la revista *Biophysical Journal* (Murciano-Calles et al., 2010).
- III. Se hizo, asimismo, un estudio paralelo al descrito en el punto II del dominio PDZ3-PSD95 a valores ácidos de pH para profundizar en las características moleculares de los procesos observados. Se descubrió que el intermedio se desestabiliza energéticamente, no poblándose significativamente cuando el pH es menor de 3.5. Esto apuntaría a los residuos aspartato y/o glutamato como los responsables de dicho comportamiento, puesto que tienen un valor de pK_a alrededor de ese valor de pH. Estos resultados han sido recogidos en una publicación científica que está siendo revisada en la revista *Biophysical Chemistry*.
- IV. Se clonó y purificó el dominio PDZ3-PSD95 sin su hélice α_3 . Dicha hélice es un elemento de conexión entre el dominio PDZ3-PSD95 y el módulo SH3-GK-PSD95 y ya se había demostrado en estudios previos que es un elemento regulador de las características de unión de este dominio, hecho que se comprobó en esta Memoria mediante ITC. Se realizaron

experimentos en paralelo a los realizados con el dominio PDZ3-PSD95 para evaluar el papel de la hélice $\alpha 3$ en el plegamiento del dominio.

- V. Finalmente, se describe en esta Memoria el protocolo desarrollado para intentar establecer la especificidad de unión de los dominios PDZ a partir del único conocimiento de su secuencia aminoacídica mediante la metodología de expresión en fagos. Para ello, se eligió como sistema arquetípico el dominio PDZ-Erbin, mediante el cual se generó una librería en la cual se mutaron de manera aleatoria todos los residuos responsables de interaccionar con el aminoácido presente en la posición 0 del ligando. Se eligieron como moléculas diana para las selecciones veinte variantes del ligando de alta afinidad para PDZ-Erbin, TGWETWV, las cuales se diferenciaban cada una en el aminoácido presente en su posición 0. Seguidamente se seleccionaron los mutantes del dominio PDZ-Erbin de la librería frente a cada una de las variantes del péptido, y se secuenciaron aquellos que se unieron a cada uno de los péptidos.

Las conclusiones más importantes del trabajo presentado en esta Memoria se enumeran a continuación:

1. Las estructuras del dominio PDZ3-PSD95 han sido resueltas por difracción de rayos X a 1.4 Å de resolución. Éstas revelaron un cambio conformacional de los residuos presentes en la hélice $\alpha 3$ respecto a las estructuras conocidas de dicho dominio, las cuales fueron obtenidas a partir de secuencias del dominio PDZ3-PSD95 que contenía un artefacto de clonación que no pertenece a la secuencia natural de la proteína PSD95.
2. En dichas estructuras se observó la presencia de un anillo de succinimida en la posición Asp332, lo cual implica una reducción de la flexibilidad del giro entre las cadenas $\beta 2$ y $\beta 3$ del dominio donde se encuentra ese resto, siendo de esta forma la región más rígida en el global de la estructura. Este anillo de succinimida favorece la formación del cristal y participa directamente en los contactos cristalinos.
3. El dominio PDZ3-PSD95 presenta en su mecanismo de desplegamiento un intermedio de equilibrio que muestra fenómenos de asociación, tal y como revelaron los experimentos de CD, DSC, DLS y SEC. Además, los resultados experimentales apuntan a que el grado de asociación del intermedio corresponde a un trímero, aunque también podría ser una mezcla de dímeros, trímeros y tetrámeros, que podría compendiarse estadísticamente a un trímero.
4. Este intermedio de equilibrio es además un precursor de estructuras anulares y fibrilares, según se observó mediante TEM y la unión a Congo Red, ANS y ThT. Además, desde un punto de vista cinético no se detecta

- la presencia de una fase de retardo (*lag phase*), la cual se ha observado en la mayoría de los procesos de fibrilación de otras proteínas.
5. El proceso de formación de estas estructuras supramacromoleculares presenta un alto carácter reversible. La reversibilidad en la formación de estas estructuras a partir del intermedio se puede deber a la mayor estabilidad del estado nativo a temperatura ambiente respecto a la de aquellas, en contraposición a la mayoría de los ejemplos descritos hasta la fecha.
 6. Las conclusiones 4 y 5, junto con la formación de estructuras anulares de forma reversible sin la presencia de membranas o disolventes orgánicos, como habían sido observadas hasta la fecha, hacen del dominio PDZ3-PSD95 un modelo más que interesante para la caracterización y estudio de los mecanismos de agregación relacionados con el plegamiento de las proteínas.
 7. A valores de pH menores de 3.5 el intermedio de plegamiento observado en el dominio PDZ3-PSD95 se desestabiliza energéticamente. A esos valores de pH, además de que no se forman estructuras anulares, la formación de fibras es más rápida e irreversible que la que se da a pH neutro presentándose en su mecanismo una fase de retardo. La irreversibilidad se debería a la mayor estabilidad de las fibras con respecto al estado nativo.
 8. El estudio equivalente del dominio PDZ3-PSD95 sin su hélice α_3 , en donde están algunos de los posibles restos responsables de las diferencias descritas en función del pH, reveló que al eliminar dicha hélice se produce una estabilización del intermedio. El intermedio no forma estructuras anulares, aunque sí fibrilares, pero lo hace más rápidamente y de forma irreversible. Nuevamente, al bajar el pH, el intermedio se desestabiliza energéticamente, de forma paralela a como ocurre en la construcción con la hélice.
 9. Todos estos resultados nos hacen postular una organización estructural para el intermedio de plegamiento del dominio PDZ3-PSD95 que presentaría un contenido en estructura secundaria semejante al del estado nativo, pero donde algunos elementos estructurales, principalmente las cadenas β_3 y β_5 podrían encontrarse más o menos desorganizadas.
 10. Los resultados obtenidos a partir de las selecciones de la librería del dominio PDZ-Erbin confirmaron que la metodología de expresión en fagos es una herramienta muy potente para el estudio de la especificidad de unión de dominios PDZ, debido a su previamente conocida potencialidad a la que cuando además se incluyó un proceso de competición generó la obtención de variantes más selectivas con respecto a sus moléculas diana.

CHAPTER 9

SUMMARY AND CONCLUSIONS

The work presented in this Thesis project has been focused the in depth study of the conformational equilibrium of the third PDZ domain of the hub PSD95 protein. More interestingly, this piece of work constitutes the first thermodynamic-structural study of the conformational equilibrium of a PDZ domain achieved up to date. In addition, by using the phage display methodology we have tried to develop a protocol which allows us to advance in the understanding the rules governing PDZ domains specificity upon ligand binding processes, in order to predict their functional properties from its aminoacid sequence.

In detail, the work is summarized in the following points:

- I. PSD95-PDZ3 domain was cloned and purified. Then, crystallization conditions were explored by using hanging drop and counter-diffusion methods. Obtained crystals were diffracted and the structure was solved under two of the studied conditions. These results were published in the *Journal of Structural Biology* (Camara-Artigas et al., 2010).
- II. The conformational equilibrium of PSD95-PDZ3 domain was characterized by DSC at neutral pH conditions. Since the DSC results revealed the presence of an equilibrium intermediate that presented association-dissociation phenomena, as inferred from the concentration dependence of the calorimetric traces, DLS, SEC, CD and ANS and ThT fluorescence experiments were conducted to characterize such intermediate state. In addition, this intermediate proved to be a precursor of annular and fibrillar structures, as revealed by different spectroscopic techniques and TEM micrographs. All these results have been the subject of a publication in *Biophysical Journal* (Murciano-Calles et al., 2010).
- III. A set of experiments achieved in parallel studies mentioned in point II, were with PSD95-PDZ3 domain at acidic pH, to get insight into the molecular aspects of the observed processes. It was found that the intermediate is destabilized energetically and becomes significantly unpopulated below pH 3.5. This should lead to aspartate and/or glutamate residues as the responsible for this behaviour, since they have a pK_a around this pH value. These results have been collected in a scientific publication that is being reviewed in the journal *Biophysical Chemistry*.
- IV. PSD95-PDZ3 domain was also cloned and purified without the $\alpha 3$ helix. This helix acts as a linker between the PSD95-PDZ3 domain and the PSD95-SH3-GK module. It has been shown previously its role as a regulator of the binding properties of this domain, a fact confirmed in this Thesis by ITC experiments. Equivalent experiments to those made with PSD95-PDZ3 domain were conducted with this construction to evaluate the $\alpha 3$ helix role in the folding of the domain.
- V. Finally, this Thesis describes a protocol that tries to establish the binding specificity of PDZ domains from its amino acid sequence by using the

phage display method. To achieve this goal, a library was generated using the PDZ-Erbin domain as a template. In such library, the aminoacids interacting with the residue located in position 0 of the ligand were randomly mutated. To perform the selections, twenty variants of a high-affinity ligand of the Erbin-PDZ domain, TGWETWV, which differed in each natural aminoacid in position 0, were chosen as target molecules. Then, mutants of the Erbin-PDZ domain of the library were selected against each one of the variants of the peptide. Only those variants that bind to each one of the peptides were sequenced.

The conclusions of the work presented in this Thesis project are listed below:

1. PSD95-PDZ3 structures have been solved by X-ray diffraction at 1.4 Å resolution. These structures have revealed a conformational change of residues in the $\alpha 3$ helix compared to the previously known structures for this domain, which contain a cloning artefact that does not belong to the natural sequence of PSD95 protein.
2. Our structures have shown the presence of a succinimide ring at position Asp332, which implies a reduction in the flexibility of the turn between $\beta 2$ and $\beta 3$ sheets. Thus, this region becomes is the most rigid one in the overall structure. The succinimide ring also favours the formation of the crystal being directly involved in crystal contacts.
3. PSD95-PDZ3 domain presents an equilibrium unfolding intermediate that shows association events, as revealed by CD, DSC, DLS and SEC experiments. Moreover, although experimental results indicate that the degree of association of the intermediate may correspond to a trimer, it could also be a mixture of dimers, trimers and tetramers, which could be summed up statistically to a trimer.
4. This equilibrium intermediate is revealed also as a precursor of annular and fibrillar structures, as observed by TEM and Congo Red binding, ANS and ThT. Additionally, from a kinetically point of view, there is an absence of a lag phase in the formation of these supramacromolecular structures, which is in contrast to the lag phase observed in most of the fibrillation processes observed for some other proteins.
5. The formation of these supramacromolecular structures has a high reversible nature. This reversibility may be due to the greater stability of the native state at room temperature related to the stability of the supramacromolecular structures, in contrast to the majority of examples described so far.
6. Points 4 and 5, together with the reversible formation of annular structures in the absence of membranes or organic solvents, as it has been observed

up to date, make PSD95-PDZ3 domain an interesting system for further studies into aggregation processes coupled to protein folding.

7. The folding intermediate observed in PSD95-PDZ3 domain is not populated at pH values below 3.5, due to its energetic destabilization. At these pH values, fiber formation is faster than at neutral pH and it is irreversible. In addition, no annular structures have been found. Furthermore, the mechanism during fibril growing includes a lag-phase, which is not observed at neutral pH. The irreversibility may be due to the increased stability of the fibers with respect to the native state.
8. The equivalent study of the PSD95-PDZ3 domain in the absence of the $\alpha 3$ helix, where there are some possible residues responsible for the different behaviour observed upon pH changes, has shown that an energetic stabilization of the intermediate is induced by removing this helix. Besides, although the intermediate is able to form fibrils in a faster and irreversible way, it does not form annular structures. Again, at acidic pH values, the intermediate is energetically destabilized, as it occurs in the case of the construction that contains the helix.
9. With all these results a structural organization for the folding intermediate of PSD95-PDZ3 domain can be postulated. Thus, the arrangement of this intermediate would have a secondary structure content similar to the native state, where some of the structural elements, mainly $\beta 3$ and $\beta 5$, could be partly disordered.
10. The results obtained from the Erbin-PDZ domain library selections has confirmed that phage display is a powerful tool to study the binding specificity of PDZ domains. The previously known potential of this methodology has been demonstrated here, being even improved by including a competitive process by which more selective variants can be obtained.

BIBLIOGRAFÍA

- Appleton BA, Zhang Y, Wu P, Yin JP, Hunziker W, Skelton NJ, Sidhu SS, Wiesmann C (2006) Comparative structural analysis of the Erbin PDZ domain and the first PDZ domain of ZO-1. Insights into determinants of PDZ domain specificity. *J Biol Chem* 281: 22312-22320.
- Azuaga AI, Dobson CM, Mateo PL, Conejero-Lara F (2002) Unfolding and aggregation during the thermal denaturation of streptokinase. *Eur J Biochem* 269: 4121-4133.
- Baldwin RL, Muller N (1992) Relation between the convergence temperatures T_h^* and T_s^* in protein unfolding. *Proc Natl Acad Sci U S A* 89: 7110-7113.
- Ballif BA, Carey GR, Sunyaev SR, Gygi SP (2007) Large-Scale Identification and Evolution Indexing of Tyrosine Phosphorylation Sites from Murine Brain. *Journal of Proteome Research* 7: 311-318.
- Barabasi A-L, Albert R (1999) Emergence of Scaling in Random Networks. *Science* 286: 509-512.
- Binger KJ, Pham CLL, Wilson LM, Bailey MF, Lawrence LJ, Schuck P, Howlett GJ (2008) Apolipoprotein C-II Amyloid Fibrils Assemble via a Reversible Pathway that Includes Fibril Breaking and Rejoining. *J Mol Biol* 376: 1116-1129.
- Borg J-P, Marchetto S, Le Bivic A, Ollendorff V, Jaulin-Bastard F, Saito H, Fournier E, Adelaide J, Margolis B, Birnbaum D (2000) ERBIN: a basolateral PDZ protein that interacts with the mammalian ERBB2/HER2 receptor. *Nat Cell Biol* 2: 407-414.
- Brahms S, Brahms J (1980). Determination of protein secondary structure in solution by vacuum ultraviolet circular dichroism. *J Mol Biol* 138(2): 149-178
- Brenman JE, Chao DS, Gee SH, McGee AW, Craven SE, Santillano DR, Wu Z, Huang F, Xia H, Peters MF, Froehner SC, Brecht DS (1996) Interaction of nitric oxide synthase with the postsynaptic density protein PSD-95 and alpha1-syntrophin mediated by PDZ domains. *Cell* 84: 757-767.
- Buchner J, Kiefhaber T (eds) (2005) Protein Folding Handbook. Wiley-VCH.
- Calamai M, Canale C, Relini A, Stefani M, Chiti F, Dobson CM (2005) Reversal of Protein Aggregation Provides Evidence for Multiple Aggregated States. *J Mol Biol* 346: 603-616.
- Calosci N, Chi CN, Richter B, Camilloni C, Engström Ak, Eklund L, Travaglini-Allocatelli C, Gianni S, Vendruscolo M, Jemth P (2008) Comparison of successive transition states for folding reveals alternative early folding pathways of two homologous proteins. *Proc Natl Acad Sci U S A* 105: 19241-19246.
- Cámara-Artigas A, Murciano-Calles J, Gavira JA, Cobos ES, Martínez JC (2010) Novel conformational aspects of the third PDZ domain of the neuronal post-synaptic density-95 protein revealed from two 1.4 Å X-ray structures. *Journal of Structural Biology* 170(3): 565-569.

- Candel AM (2008) Tesis Doctoral: Diseño y caracterización de proteínas quiméricas para el estudio del reconocimiento molecular entre ligandos peptídicos y dominios SH3. Departamento de Química Física, Universidad de Granada
- Carrotta R, Manno M, Bulone D, Martorana V, Biagio PLS (2005) Protofibril Formation of Amyloid β -Protein at Low pH via a Non-cooperative Elongation Mechanism. *J Biol Chem* 280: 30001-30008.
- Caughey B, Ernst D, Race RE (1993) Congo red inhibition of scrapie agent replication. *J Virol* 67: 6270-6272.
- Chen L, Chetkovich DM, Petralia RS, Sweeney NT, Kawasaki Y, Wenthold RJ, Bredt DS, Nicoll RA (2000) Stargazin regulates synaptic targeting of AMPA receptors by two distinct mechanisms. *Nature* 408: 936-943.
- Cheng D, Hoogenraad CC, Rush J, Ramm E, Schlager MA, Duong DM, Xu P, Wijayawardana SR, Hanfelt J, Nakagawa T, Sheng M, Peng J (2006) Relative and Absolute Quantification of Postsynaptic Density Proteome Isolated from Rat Forebrain and Cerebellum. *Molecular & Cellular Proteomics* 5: 1158-1170.
- Chi CN, Gianni S, Calosci N, Travaglini-Allocatelli C, Engstrom K, Jemth P (2007) A conserved folding mechanism for PDZ domains. *FEBS Lett* 581: 1109-1113.
- Cho KO, Hunt CA, Kennedy MB (1992) The rat brain postsynaptic density fraction contains a homolog of the Drosophila discs-large tumor suppressor protein. *Neuron* 9: 929-942.
- Christopherson KS, Sweeney NT, Craven SE, Kang R, El-Husseini AE-D, Bredt DS (2003) Lipid- and protein-mediated multimerization of PSD-95: implications for receptor clustering and assembly of synaptic protein networks. *J Cell Sci* 116: 3213-3219.
- Clackson T, Lowman H (eds) (2004) Phage Display. Oxford University Press, Oxford.
- Clarke S (1987) Propensity for spontaneous succinimide formation from aspartyl and asparaginylyl residues in cellular proteins. *International Journal of Peptide and Protein Research* 30: 808-821.
- Cobos ES (2002) Tesis Doctoral: Estabilidad, plegamiento y estructura de la proteína circular AS-48 y de mutantes del dominio SH3. Departamento de Química Física, Universidad de Granada
- Colledge M, Dean RA, Scott GK, Langeberg LK, Haganir RL, Scott JD (2000) Targeting of PKA to Glutamate Receptors through a MAGUK-AKAP Complex. *Neuron* 27: 107-119.
- Compton LA, Johnson WC (1986) Analysis of protein circular dichroism spectra for secondary structure using a simple matrix multiplication. *Analytical Biochemistry* 155: 155-167.

- Craven SE, Bredt DS (1998) PDZ proteins organize synaptic signaling pathways. *Cell* 93: 495-498.
- Creighton TE (1993) *Proteins: Structures and Molecular Properties*. WH Freeman and Co., New York.
- Crooks GE, Hon G, Chandonia J-M, Brenner SE (2004) WebLogo: A Sequence Logo Generator. *Genome Research* 14: 1188-1190.
- Cusick ME, Klitgord N, Vidal M, Hill DE (2005) Interactome: gateway into systems biology. *Human Molecular Genetics* 14: R171-R181.
- Dakoji S, Tomita S, Karimzadegan S, Nicoll RA, Bredt DS (2003) Interaction of transmembrane AMPA receptor regulatory proteins with multiple membrane associated guanylate kinases. *Neuropharmacology* 45: 849-856.
- Dalboge H, Bayne S, Pedersen J (1990) In vivo processing of N-terminal methionine in E. coli. *FEBS Lett* 266: 1-3.
- Dalva MB, McClelland AC, Kayser MS (2007) Cell adhesion molecules: signalling functions at the synapse. *Nat Rev Neurosci* 8: 206-220.
- Dean C, Scholl FG, Choih J, DeMaria S, Berger J, Isacoff E, Scheiffele P (2003) Neurexin mediates the assembly of presynaptic terminals. *Nat Neurosci* 6: 708-716.
- Dev KK (2004) Making protein interactions druggable: targeting PDZ domains. *Nat Rev Drug Discov* 3: 1047-1056.
- Doyle DA, Lee A, Lewis J, Kim E, Sheng M, MacKinnon R (1996) Crystal structures of a complexed and peptide-free membrane protein-binding domain: molecular basis of peptide recognition by PDZ. *Cell* 85: 1067-1076.
- Dunker AK, Cortese MS, Romero P, Iakoucheva LM, Uversky VN (2005) Flexible nets. *FEBS Journal* 272: 5129-5148.
- Durand D, Vivès C, Cannella D, Pérez J, Pebay-Peyroula E, Vachette P, Fieschi F (2010) NADPH oxidase activator p67phox behaves in solution as a multidomain protein with semi-flexible linkers. *Journal of Structural Biology* 169: 45-53.
- Ernst A, Gfeller D, Kan Z, Seshagiri S, Kim PM, Bader GD, Sidhu SS (2010) Coevolution of PDZ domain-ligand interactions analyzed by high-throughput phage display and deep sequencing. *Molecular BioSystems* 6: 1782-1790.
- Ernst A, Sazinsky SL, Hui S, Currell B, Dharsee M, Seshagiri S, Bader GD, Sidhu SS (2009) Rapid Evolution of Functional Complexity in a Domain Family. *Sci Signal* 2: ra50-.
- Fanning AS, Anderson JM (1996) Protein-protein interactions: PDZ domain networks. *Curr Biol* 6: 1385-1388.

- Feng H, Vu ND, Bai Y (2005) Detection of a hidden folding intermediate of the third domain of PDZ. *J Mol Biol* 346: 345-353.
- Feng W, Zhang M (2009) Organization and dynamics of PDZ-domain-related supramodules in the postsynaptic density. *Nat Rev Neurosci* 10: 87-99.
- Feng W, Wu H, Chan LN, Zhang M (2008) Par-3-mediated Junctional Localization of the Lipid Phosphatase PTEN Is Required for Cell Polarity Establishment. *J Biol Chem* 283(34): 23440-23449.
- Fernandez-Escamilla A-M, Rousseau F, Schymkowitz J, Serrano L (2004) Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nat Biotech* 22: 1302-1306.
- Fersht A (ed) (1999) *Structure and Mechanism in Protein Science*. Freeman.
- Fersht AR, Daggett V (2002) Protein folding and unfolding at atomic resolution. *Cell* 108:573-582.
- Filimonov VV, Prieto J, Martinez JC, Bruix M, Mateo PL, Serrano L (1993) Thermodynamic analysis of the chemotactic protein from Escherichia coli, CheY. *Biochemistry* 32: 12906-12921.
- Filimonov VV, Rogov VV (1996) Reversible Association of the Equilibrium Unfolding Intermediate of [λ] Cro Repressor. *J Mol Biol* 255: 767-777.
- Fu L, Freire E (1992) On the origin of the enthalpy and entropy convergence temperatures in protein folding. *Proc Natl Acad Sci U S A* 89: 9335-9338.
- Gallardo R, Ivarsson Y, Schymkowitz J, Rousseau F, Zimmermann P (2010) Structural Diversity of PDZ-Lipid Interactions. *ChemBioChem* 11(4): 456-467.
- García-Ruiz JM, Charles W. Carter, Jr., Robert MS (2003) Counterdiffusion Methods for Macromolecular Crystallization. In *Methods in Enzymology* pp 130-154. Academic Press.
- Gerber R, Tahiri-Alaoui A, Hore PJ, James W (2008) Conformational pH dependence of intermediate states during oligomerization of the human prion protein. *Protein Science* 17: 537-544.
- Gianni S, Calosci N, Aelen JM, Vuister GW, Brunori M, Travaglini-Allocatelli C (2005) Kinetic folding mechanism of PDZ2 from PTP-BL. *Protein Eng Des Sel* 18: 389-395.
- Gianni S, Engstrom A, Larsson M, Calosci N, Malatesta F, Eklund L, Ngang CC, Travaglini-Allocatelli C, Jemth P (2005) The kinetics of PDZ domain-ligand interactions and implications for the binding mechanism. *J Biol Chem* 280: 34805-34812.
- Gianni S, Walma T, Arcovito A, Calosci N, Bellelli A, Engstrom A, Travaglini-Allocatelli C, Brunori M, Jemth P, Vuister GW (2006) Demonstration of Long-Range Interactions in a PDZ Domain by NMR, Kinetics, and Protein Engineering. *Structure* 14: 1801-1809.

- Gianni S, Geierhaas CD, Calosci N, Jemth P, Vuister GW, Travaglini-Allocatelli C, Vendruscolo M, Brunori M (2007) A PDZ domain recapitulates a unifying mechanism for protein folding. *Proc Natl Acad Sci U S A* 104: 128-133.
- Gianni S, Ivarsson Y, De Simone A, Travaglini-Allocatelli C, Brunori M, Vendruscolo M (2010) Structural characterization of a misfolded intermediate populated during the folding process of a PDZ domain. *Nat Struct Mol Biol* 17: 1431-1437.
- Gill SC, von Hippel PH (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Analytical Biochemistry* 182: 319-326.
- Goh K-I, Oh E, Jeong H, Kahng B, Kim D (2002) Classification of scale-free networks. *Proc Natl Acad Sci U S A* 99: 12583-12588.
- Gokhale RS, Khosla C (2000) Role of linkers in communication between protein modules. *Current Opinion in Chemical Biology* 4: 22-27.
- Grosschedl R, Giese K, Pagel J (1994) HMG domain proteins: architectural elements in the assembly of nucleoprotein structures. *Trends Genet* 10, 94-100.
- Güthe S, Kapinos L, Möglich A, Meier S, Grzesiek S, Kiefhaber T (2004) Very Fast Folding and Association of a Trimerization Domain from Bacteriophage T4 Fibrin. *J Mol Biol* 337: 905-915.
- Han J-DJ, Bertin N, Hao T, Goldberg DS, Berriz GF, Zhang LV, Dupuy D, Walhout AJM, Cusick ME, Roth FP, Vidal M (2004) Evidence for dynamically organized modularity in the yeast protein-protein interaction network. *Nature* 430: 88-93.
- Haq SR, Järgens MC, Chi CN, Koh C-S, Elfström L, Selmer M, Gianni S, Jemth P (2010) The Plastic Energy Landscape of Protein Folding. *J Biol Chem* 285: 18051-18059.
- Harper JD, Lansbury PT (1997) Models Of Amyloid Seeding In Alzheimer's Disease And Scrapie: Mechanistic Truths and Physiological Consequences of the Time-Dependent Solubility of Amyloid Proteins. *Annual Review of Biochemistry* 66: 385-407.
- Harris BZ, Hillier BJ, Lim WA (2001) Energetic determinants of internal motif recognition by PDZ domains. *Biochemistry* 40: 5921-5930.
- Hartwell LH, Hopfield JJ, Leibler S, Murray AW (1999) From molecular to modular Cell biology. *Nature* 402: C47-C52.
- Hasty J, Collins JJ (2001) Protein interactions: Unspinning the web. *Nature* 411: 30-31.
- Hawe A, Sutter M, Jiskoot W (2008) Extrinsic Fluorescent Dyes as Tools for Protein Characterization. *Pharmaceutical Research* 25: 1487-1499.
- Higgins DG, Thompson JD, Gibson TJ, Russell FD (1996) [22] Using CLUSTAL for multiple sequence alignments. In *Methods in Enzymology* pp 383-402. Academic Press.

- Hillier BJ, Christopherson KS, Prehoda KE, Bretz DS, Lim WA (1999) Unexpected modes of PDZ domain scaffolding revealed by structure of nNOS-syntrophin complex. *Science* 284: 812-815.
- Hirel PH, Schmitter MJ, Dessen P, Fayat G, Blanquet S (1989) Extent of N-terminal methionine excision from Escherichia coli proteins is governed by the side-chain length of the penultimate amino acid. *Proc Natl Acad Sci U S A* 86: 8247-8251.
- Hoffmann R, Valencia A (2003) Protein interaction: same network, different hubs. *Trends in Genetics* 19: 681-683.
- Hoyer W, Antony T, Cherny D, Heim G, Jovin TM, Subramaniam V (2002) Dependence of [alpha]-Synuclein Aggregate Morphology on Solution Conditions. *J Mol Biol* 322: 383-393.
- Hsueh YP, Sheng M (1999) Requirement of N-terminal cysteines of PSD-95 for PSD-95 multimerization and ternary complex formation, but not for binding to potassium channel Kv1.4. *J Biol Chem* 274: 532-536.
- Huang YZ, Wang Q, Xiong WC, Mei L (2001) Erbin Is a Protein Concentrated at Postsynaptic Membranes That Interacts with PSD-95. *J Biol Chem* 276: 19318-19326.
- Human Genome Sequencing C (2004) Finishing the euchromatic sequence of the human genome. *Nature* 431: 931-945.
- Im YJ, Park SH, Rho S-H, Lee JH, Kang GB, Sheng M, Kim E, Eom SH (2003) Crystal Structure of GRIP1 PDZ6-Peptide Complex Reveals the Structural Basis for Class II PDZ Target Recognition and PDZ Domain-mediated Multimerization. *J Biol Chem* 278: 8501-8507.
- Irie M, Hata Y, Takeuchi M, Ichtchenko K, Toyoda A, Hirao K, Takai Y, Rosahl TW, Sudhof TC (1997) Binding of neuroligins to PSD-95. *Science* 277: 1511-1515.
- Itoh M, Nagafuchi A, Yonemura S, Kitani-Yasuda T, Tsukita S (1993) The 220-kD protein colocalizing with cadherins in non-epithelial cells is identical to ZO-1, a tight junction-associated protein in epithelial cells: cDNA cloning and immunoelectron microscopy. *The Journal of Cell Biology* 121: 491-502.
- Ivarsson Y, Travaglini-Allocatelli C, Jemth P, Malatesta F, Brunori M, Gianni S (2007) An On-pathway Intermediate in the Folding of a PDZ Domain. *J Biol Chem* 282: 8568-8572.
- Ivarsson Y, Travaglini-Allocatelli C, Morea V, Brunori M, Gianni S (2008) The folding pathway of an engineered circularly permuted PDZ domain. *Protein Eng Des Sel*.
- Ivarsson Y, Travaglini-Allocatelli C, Brunori M, Gianni S (2009) Engineered Symmetric Connectivity of Secondary Structure Elements Highlights Malleability of Protein Folding Pathways. *Journal of the American Chemical Society* 131: 11727-11733.

- Jang H, Zheng J, Lal R, Nussinov R (2008) New structures help the modeling of toxic amyloid β ion channels. *Trends in Biochemical Sciences* 33: 91-100.
- Jelesarov I, Bosshard HR (1999) Isothermal titration calorimetry and differential scanning calorimetry as complementary tools to investigate the energetic of biomolecular recognition. *J Mol Recognit* 12(1): 3-18.
- Jemth P, Gianni S (2007) PDZ Domains: Folding and Binding. *Biochemistry* 46: 8701-8708.
- Jeong H, Mason SP, Barabasi AL, Oltvai ZN (2001) Lethality and centrality in protein networks. *Nature* 411: 41-42.
- Jeong H, Tombor B, Albert R, Oltvai ZN, Barabasi AL (2000) The large-scale organization of metabolic networks. *Nature* 407: 651-654.
- Jiang Y, Li H, Zhu L, Zhou J-M, Perrett S (2004) Amyloid Nucleation and Hierarchical Assembly of Ure2p Fibrils. *J Biol Chem* 279: 3361-3369.
- Johnson WC (1990) Protein secondary structure and circular dichroism: A practical guide. *Proteins: Structure, Function, and Bioinformatics* 7: 205-214.
- Juárez J, López SG, Cambón A, Taboada P, Mosquera Vc (2009) Influence of Electrostatic Interactions on the Fibrillation Process of Human Serum Albumin. *The Journal of Physical Chemistry B* 113: 10521-10529.
- Juárez J, Taboada P, Mosquera V (2009) Existence of Different Structural Intermediates on the Fibrillation Pathway of Human Serum Albumin. *Biophysical Journal* 96: 2353-2370.
- Kad NM, Myers SL, Smith DP, Alastair Smith D, Radford SE, Thomson NH (2003) Hierarchical Assembly of [beta]2-Microglobulin Amyloid In Vitro Revealed by Atomic Force Microscopy. *J Mol Biol* 330: 785-797.
- Kalia LV, Salter MW (2003) Interactions between Src family protein tyrosine kinases and PSD-95. *Neuropharmacology* 45: 720-728.
- Kardos Jz, Yamamoto K, Hasegawa K, Naiki H, Goto Y (2004) Direct Measurement of the Thermodynamic Parameters of Amyloid Formation by Isothermal Titration Calorimetry. *J Biol Chem* 279: 55308-55314.
- Kehoe JW, Kay BK (2005) Filamentous Phage Display in the New Millennium. *Chem Rev* 105: 4056-4072.
- Kelly JW (2000) Mechanisms of amyloidogenesis. *Nat Struct Mol Biol* 7: 824-826.
- Kelly WP, Stumpf MPH (2008) Protein-protein interactions: from global to local analyses. *Current Opinion in Biotechnology* 19: 396-403.
- Kennedy MB (2000) Signal-processing machines at the postsynaptic density. *Science* 290: 750-754.

- Khurana R, Uversky VN, Nielsen L, Fink AL (2001) Is Congo Red an Amyloid-specific Dye? *J Biol Chem* 276: 22715-22721.
- Kim E, Niethammer M, Rothschild A, Jan YN, Sheng M (1995) Clustering of Shaker-type K⁺ channels by interaction with a family of membrane-associated guanylate kinases. *Nature* 378: 85-88.
- Kim E, Sheng M (2004) PDZ domain proteins of synapses. *Nat Rev Neurosci* 5: 771-781.
- Kim SK (1995) Tight junctions, membrane-associated guanylate kinases and Cell signaling. *Current Opinion in Cell Biology* 7: 641-649.
- Klapper LN, Kirschbaum MH, Seta M, Yarden Y, George FVW, George K (1999) Biochemical and Clinical Implications of the ErbB/HER Signaling Network of Growth Factor Receptors. In *Advances in Cancer Research* pp 25-79. Academic Press.
- Klunk WE, Jacob RF, Mason RP, Ronald W (1999) [19] Quantifying amyloid by congo red spectral shift assay. In *Methods in Enzymology* pp 285-305. Academic Press.
- Klunk WE, Pettegrew JW, Abraham DJ (1989) Quantitative evaluation of congo red binding to amyloid-like proteins with a beta-pleated sheet conformation. *Journal of Histochemistry & Cytochemistry* 37: 1273-1281.
- Knowles TPJ, Waudby CA, Devlin GL, Cohen SIA, Aguzzi A, Vendruscolo M, Terentjev EM, Welland ME, Dobson CM (2009) An Analytical Solution to the Kinetics of Breakable Filament Assembly. *Science* 326: 1533-1537.
- Ko J, Kim E (2007) Leucine-rich repeat proteins of synapses. *Journal of NeuroScience Research* 85: 2824-2832.
- Kodali R, Wetzel R (2007) Polymorphism in the intermediates and products of amyloid assembly. *Current Opinion in Structural Biology* 17: 48-57.
- Korkin D, Davis FP, Alber F, Luong T, Shen MY, Lucic V, Kennedy MB, Sali A (2006) Structural modeling of protein interactions by analogy: application to PSD-95. *PLoS Comput Biol* 2: e153.
- Krebs MRH, Devlin GL, Donald AM (2007) Protein Particulates: Another Generic Form of Protein Aggregation? 92: 1336-1342.
- Krahn MP, Klopfenstein DR, Fischer N, Wodarz A (2010) Membrane Targeting of Bazooka/PAR-3 Is Mediated by Direct Binding to Phosphoinositide Lipids. *Current biology* 20(7): 636-642.
- Kunkel TA (1985) Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc Natl Acad Sci U S A* 82: 488-492.
- Lashuel HA, Lansbury PT (2006) Are amyloid diseases caused by protein aggregates that mimic bacterial pore-forming toxins? *Quarterly Reviews of Biophysics* 39: 167-201.

- Lepock JR, Ritchie KP, Kolios MC, Rodahl AM, Heinz KA, Kruuv J (1992) Influence of transition rates and scan rate on kinetic simulations of differential scanning calorimetry profiles of reversible and irreversible protein denaturation. *Biochemistry* 31: 12706-12712.
- LeVine III H, Ronald W (1999) [18] Quantification of [beta]-sheet amyloid fibril structures with thioflavin T. In *Methods in Enzymology* pp 274-284. Academic Press.
- Li Z, Sheng M (2003) Some assembly required: the development of neuronal synapses. *Nat Rev Mol Cell Biol* 4: 833-841.
- Lockless SW, Ranganathan R (1999) Evolutionarily conserved pathways of energetic connectivity in protein families. *Science* 286: 295-299.
- Long J-F, Tochio H, Wang P, Fan J-S, Sala C, Niethammer M, Sheng M, Zhang M (2003) Supramodular structure and synergistic target binding of the N-terminal tandem PDZ domains of PSD-95. *J Mol Biol* 327: 203-214.
- Lorenzo A, Yankner BA (1994) Beta-amyloid neurotoxicity requires fibril formation and is inhibited by congo red. *Proc Natl Acad Sci U S A* 91: 12243-12247.
- Lund M, Jönsson B (2005) On the Charge Regulation of Proteins. *Biochemistry* 44: 5722-5727.
- Luque I, Freire E (2000) Structural stability of binding sites: consequences for binding affinity and allosteric effects. *Proteins Suppl* 4:63-71.
- Luque I, Leavitt SA, Freire E (2002) The linkage between protein folding and functional cooperativity: two sides of the same coin? *Annu Rev Biophys Biomol Struct* 31:235-256.
- Ma JC, Dougherty DA (1997) The Cation- π Interaction. *Chemical Reviews* 97: 1303-1324.
- Makhatadze GI, Privalov PL (1990) Heat capacity of proteins. I. Partial molar heat capacity of individual amino acid residues in aqueous solution: hydration effect. *J Mol Biol* 213: 375-384.
- Malenka RC, Bear MF (2004) LTP and LTD: An Embarrassment of Riches. *Neuron* 44: 5-21.
- Martinez JC, el Harrous M, Filimonov VV, Mateo PL, Fersht AR (1994) A calorimetric study of the thermal stability of barnase and its interaction with 3'GMP. *Biochemistry* 33: 3919-3926.
- Martín-García JM (2009) Tesis Doctoral: Determinantes moleculares de la afinidad y de la especificidad de unión en dominios SH3. Departamento de Química Física, Universidad de Granada
- Mayorga OL, Freire E (1987) Dynamic analysis of differential scanning calorimetry data. *Biophysical Chemistry* 27: 87-96.

- McGee AW, Brecht DS (2003) Assembly and plasticity of the glutamatergic postsynaptic specialization. *Current Opinion in Neurobiology* 13: 111-118.
- McGee AW, Dakoji SR, Olsen O, Brecht DS, Lim WA, Prehoda KE (2001) Structure of the SH3-guanylate kinase module from PSD-95 suggests a mechanism for regulated assembly of MAGUK scaffolding proteins. *Mol Cell* 8: 1291-1301.
- Morais Cabral JH, Petosa C, Sutcliffe MJ, Raza S, Byron O, Poy F, Marfatia SM, Chishti AH, Liddington RC (1996) Crystal structure of a PDZ domain. *Nature* 382: 649-652.
- Morel B, Casares S, Conejero-Lara F (2006) A single mutation induces amyloid aggregation in the alpha-spectrin SH3 domain: analysis of the early stages of fibril formation. *J Mol Biol* 356: 453-468.
- Mullis KB, Faloona FA, Ray W (1987) [21] Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. In *Methods in Enzymology* pp 335-350. Academic Press.
- Murciano-Calles J, Cobos ES, Mateo PL, Camara-Artigas A, Martinez JC (2010) An Oligomeric Equilibrium Intermediate as the Precursory Nucleus of Globular and Fibrillar Supramacromolecular Assemblies in a PDZ Domain. *Biophysical Journal* 99(1): 263-272.
- Nilsson MR (2004) Techniques to study amyloid fibril formation in vitro. *Methods* 34: 151-160.
- Noguchi S, Miyawaki K, Satow Y (1998) Succinimide and isoaspartate residues in the crystal structures of hen egg-white lysozyme complexed with tri-N-acetylchitotriose. *J Mol Biol* 278: 231-238.
- Nourry C, Grant SGN, Borg J-P (2003) PDZ Domain Proteins: Plug and Play! *Sci STKE* 2003: re7-.
- Obenauer JC, Cantley LC, Yaffe MB (2003) Scansite 2.0: proteome-wide prediction of cell signaling interactions using short sequence motifs. *Nucleic Acids Res* 31: 3635-3641
- Okabe S (2007) Molecular anatomy of the postsynaptic density. *Molecular and Cellular Neuroscience* 34: 503-518.
- Oschkinat H (1999) A new type of PDZ domain recognition. *Nat Struct Mol Biol* 6: 408-410.
- Ozaki M, Sasner M, Yano R, Lu HS, Buonanno A (1997) Neuregulin-[beta] induces expression of an NMDA-receptor subunit. *Nature* 390: 691-694.
- Paci E, Vendruscolo M, Dobson CM, Karplus M (2002) Determination of a transition state at atomic resolution from protein engineering data. *J Mol Biol* 324:151-163.
- Palay S (1958) The morphology of synapses in the central nervous system. *Exp Cell Res* 14 (Suppl 5): 275-293.

- Pawson T (1994) Look at a tyrosine kinase. *Nature* 372: 726-727.
- Pawson T (1995) Protein modules and signalling networks. *Nature* 373: 573-580.
- Pawson T, Nash P (2003) Assembly of Cell Regulatory Systems Through Protein Interaction Domains. *Science* 300: 445-452.
- Pecora R (1986) Dynamic Light Scattering; Applications of Photon Correlation Spectroscopy. Plenum Press, New York and London.
- Penkert RR, DiVittorio HM, Prehoda KE (2004) Internal recognition through PDZ domain plasticity in the Par-6-Pals1 complex. *Nat Struct Mol Biol* 11: 1122-1127.
- Petit CM, Zhang J, Sapienza PJ, Fuentes EJ, Lee AL (2009) Hidden dynamic allostery in a PDZ domain. *Proc Natl Acad Sci U S A* 106: 18249-18254.
- Price li WN, Chen Y, Handelman SK, Neely H, Manor P, Karlin R, Nair R, Liu J, Baran M, Everett J, Tong SN, Frouhar F, Swaminathan SS, Acton T, Xiao R, Luft JR, Lauricella A, DeTitta GT, Rost B, Montelione GT, et al. (2009) Understanding the physical properties that control protein crystallization by analysis of large-scale experimental data. *Nat Biotech* 27: 51-57.
- Privalov PL, Dragan AI (2007) Microcalorimetry of biological macromolecules. *Biophysical Chemistry* 126: 16-24.
- Privalov PL, Makhatadze GI (1990) Heat capacity of proteins. II. Partial molar heat capacity of the unfolded polypeptide chain of proteins: protein unfolding effects. *J Mol Biol* 213: 385-391.
- Privalov PL, Tiktopulo EI, Venyaminov SY, Griko YV, Makhatadze GI, Khechinashvili NN (1989) Heat capacity and conformation of proteins in the denatured state. *J Mol Biol* 205: 737-750.
- Puchtler H, Sweat F (1965) CONGO RED AS A STAIN FOR FLUORESCENCE MICROSCOPY OF AMYLOID. *Journal of Histochemistry & Cytochemistry* 13: 693-694.
- Quist A, Doudevski I, Lin H, Azimova R, Ng D, Frangione B, Kagan B, Ghiso J, Lal R (2005) Amyloid ion channels: A common structural link for protein-misfolding disease. *Proc Natl Acad Sci U S A* 102: 10427-10432.
- Ragone R, Colonna G (1994) Enthalpy-entropy balance and convergence temperatures in protein unfolding. *J Biol Chem* 269: 4047-4049.
- Reina J, Lacroix E, Hobson SD, Fernandez-Ballester G, Rybin V, Schwab MS, Serrano L, Gonzalez C (2002) Computer-aided design of a PDZ domain to recognize new target sequences. *Nat Struct Biol* 9: 621-627.

- Rieff HI, Raetzman LT, Sapp DW, Yeh HH, Siegel RE, Corfas G (1999) Neuregulin Induces GABAA Receptor Subunit Expression and Neurite Outgrowth in Cerebellar Granule Cells. *The Journal of Neuroscience* 19: 10757-10766.
- Ruiz-Sanz J, Filimonov VV, Christodoulou E, Vorgias CE, Mateo PL (2004) Thermodynamic analysis of the unfolding and stability of the dimeric DNA-binding protein HU from the hyperthermophilic eubacterium *Thermotoga maritima* and its E34D mutant. *European Journal of Biochemistry* 271: 1497-1507.
- Sadqi M (2000) Tesis Doctoral: Estudio termodinámico de los estados parcialmente plegados del dominio SH3 de alfa espectrina. Departamento de Química Física, Universidad de Granada
- Sainlos M, Tigaret C, Poujol C, Olivier NB, Bard L, Breillat C, Thiolon K, Choquet D, Imperiali B (2011) Biomimetic divalent ligands for the acute disruption of synaptic AMPAR stabilization. *Nat Chem Biol* 7: 81-91.
- Santonico E, Castagnoli L, Cesareni G (2005) Methods to reveal domain networks. *Drug Discovery Today* 10: 1111-1117.
- Saro D, Li T, Rupasinghe C, Paredes A, Caspers N, Spaller MR (2007) A Thermodynamic Ligand Binding Study of the Third PDZ Domain (PDZ3) from the Mammalian Neuronal Protein PSD-95. *Biochemistry* 46: 6340-6352.
- Scannevin RH, Huganir RL (2000) Postsynaptic organisation and regulation of excitatory synapses. *Nat Rev Neurosci* 1: 133-141.
- Schmid EM, McMahon HT (2007) Integrating molecular and network biology to decode endocytosis. *Nature* 448: 883-888.
- Schnell E, Sizemore M, Karimzadegan S, Chen L, Brecht DS, Nicoll RA (2002) Direct interactions between PSD-95 and stargazin control synaptic AMPA receptor number. *Proc Natl Acad Sci U S A* 99: 13902-13907.
- Serio TR, Cashikar AG, Kowal AS, Sawicki GJ, Moslehi JJ, Serpell L, Arnsdorf MF, Lindquist SL (2000) Nucleated Conformational Conversion and the Replication of Conformational Information by a Prion Determinant. *Science* 289: 1317-1321.
- Sheng M, Sala C (2001) PDZ domains and the organization of supramolecular complexes. *Annu Rev Neurosci* 24: 1-29.
- Shimizu T, Matsuoka Y, Shirasawa T (2005) Biological Significance of Isoaspartate and Its Repair System. *Biological & Pharmaceutical Bulletin* 28: 1590-1596.
- Sicorello A, Torrassa S, Soldi G, Gianni S, Travaglini-Allocatelli C, Taddei N, Relini A, Chiti F (2009) Agitation and High Ionic Strength Induce Amyloidogenesis of a Folded PDZ Domain in Native Conditions. *Biophysical Journal* 96: 2289-2298.
- Sidhu SS (ed) (2005) Phage Display in Biotechnology and Drug Discovery. CRC Press.

- Sidhu SS, Lowman HB, Cunningham BC, Wells JA, Jeremy Thorner SDE, John NA (2000) Phage display for selection of novel binding peptides. In *Methods in Enzymology* pp 333-358. Academic Press.
- Skelton NJ, Koehler MF, Zobel K, Wong WL, Yeh S, Pisabarro MT, Yin JP, Lasky LA, Sidhu SS (2003) Origins of PDZ domain ligand specificity. Structure determination and mutagenesis of the Erbin PDZ domain. *J Biol Chem* 278: 7645-7654.
- Smith AM, Jahn TR, Ashcroft AE, Radford SE (2006) Direct Observation of Oligomeric Species formed in the Early Stages of Amyloid Fibril Formation using Electrospray Ionisation Mass Spectrometry. *J Mol Biol* 364: 9-19.
- Smith CA, Kortemme T (2010) Structure-Based Prediction of the Peptide Sequence Space Recognized by Natural and Synthetic PDZ Domains. *J Mol Biol* 402: 460-474.
- Sprinzak E, Sattath S, Margalit H (2003) How Reliable are Experimental Protein-Protein Interaction Data? *J Mol Biol* 327: 919-923.
- Stephenson RC, Clarke S (1989) Succinimide formation from aspartyl and asparaginyl peptides as a model for the spontaneous degradation of proteins. *J Biol Chem* 264: 6164-6170.
- Tanaka T, Kuroda Y, Yokoyama G, Shigeyuki T (2003) Characteristics and prediction of domain linker sequences in multi-domain proteins. *Journal of Structural and Functional Genomics* 4: 79-85.
- Tao Y, Dai P, Liu Y, Marchetto S, Xiong W-C, Borg J-P, Mei L (2009) Erbin regulates NRG1 signaling and myelination. *Proc Natl Acad Sci U S A* 106: 9477-9482.
- Tochio H, Hung F, Li M, Bredt DS, Zhang M (2000) Solution structure and backbone dynamics of the second PDZ domain of postsynaptic density-95. *J Mol Biol* 295: 225-237.
- Tonikian R, Zhang Y, Boone C, Sidhu SS (2007) Identifying specificity profiles for peptide recognition modules from phage-displayed peptide libraries. *Nat Protocols* 2: 1368-1386.
- Tonikian R, Zhang Y, Sazinsky SL, Currell B, Yeh J-H, Reva B, Held HA, Appleton BA, Evangelista M, Wu Y, Xin X, Chan AC, Seshagiri S, Lasky LA, Sander C, Boone C, Bader GD, Sidhu SS (2008) A Specificity Map for the PDZ Domain Family. *PLoS Biology* 6: e239.
- Tsai C-J, Ma B, Nussinov R (2009) Protein-protein interaction networks: how can a hub protein bind so many different partners? *Trends in biochemical sciences*.
- Varela, L (2010) Tesis Doctoral: Kinetic, structural and thermodynamic determinants of amyloid fibril formation by the alpha spectrin SH3 domain. Departamento de Química Física, Universidad de Granada

- Vedadi M, Niesen FH, Allali-Hassani A, Fedorov OY, Finerty PJ, Wasney GA, Yeung R, Arrowsmith C, Ball LJ, Berglund H, Hui R, Marsden BD, Nordlund Pr, Sundstrom M, Weigelt J, Edwards AM (2006) Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proc Natl Acad Sci U S A* 103: 15835-15840.
- Velazquez-Campoy A, Leavitt SA, Freire E (2004) Characterization of protein-protein interactions by isothermal titration calorimetry. *Methods Mol Biol* 261: 35-54.
- Velazquez-Campoy A, Freire E (2006) Isothermal titration calorimetry to determine association constants for high-affinity ligands. *Nat Protoc* 1(1): 186-191.
- Vendruscolo M, Paci E, Dobson CM, Karplus M (2001) Three key residues form a critical contact network in a protein folding transition state. *Nature* 409:641-645.
- Viguera AR, Martinez JC, Filimonov VV, Mateo PL, Serrano L (1994) Thermodynamic and kinetic analysis of the SH3 domain of spectrin shows a two-state folding transition. *Biochemistry* 33: 2142-2150.
- Walsh P, Yau J, Simonetti K, Sharpe S (2009) Morphology and Secondary Structure of Stable β^2 -Oligomers Formed by Amyloid Peptide PrP. *Biochemistry* 48: 5779-5781.
- Watts DJ, Strogatz SH (1998) Collective dynamics of 'small-world' networks. *Nature* 393: 440-442.
- Weast, RC (ed) (1977). CRC Handbook of Chemistry and Physics. CRC Press.
- Weber PC, C.B. Anfinsen (1991) Physical Principles of Protein Crystallization. In *Advances in Protein Chemistry* pp 1-36. Academic Press.
- Weiss GA, Watanabe CK, Zhong A, Goddard A, Sidhu SS (2000) Rapid mapping of protein functional epitopes by combinatorial alanine scanning. *Proc Natl Acad Sci U S A* 97: 8950-8954.
- Wiseman T, Williston S, Brandts JF, Lin LN (1989). Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Analytical Biochemistry* 179(1): 131-137.
- Woods DF, Bryant PJ (1991) The discs-large tumor suppressor gene of Drosophila encodes a guanylate kinase homolog localized at septate junctions. *Cell* 66: 451-464.
- Woody RW (1994) Contributions of tryptophan side chains to the far-ultraviolet circular dichroism of proteins. *European Biophysics Journal* 23: 253-262.
- Wriggers W, Chakravarty S, Jennings PA (2005) Control of protein functional dynamics by peptide linkers. *Peptide Science* 80: 736-746.
- Wu H, Feng W, Chen J, Chan LN, Huang S, Zhang M (2007) PDZ Domains of Par-3 as Potential Phosphoinositide Signaling Integrators. *Molecular Cell* 28(5): 886-898.

- Yang X, Kuo Y, Devay P, Yu C, Role L (1998) A Cysteine-Rich Isoform of Neuregulin Controls the Level of Expression of Neuronal Nicotinic Receptor Channels during Synaptogenesis. *Neuron* 20: 255-270.
- Yeaman C, Grindstaff KK, Nelson WJ (1999) New Perspectives on Mechanisms Involved in Generating Epithelial Cell Polarity. *Physiological Reviews* 79: 73-98.
- Yu JW, Mendrola JM, Audhya A, Singh S, Keleti D, DeWald DB, Murray D, Emr SD, Lemmon MA (2004) Genomewide analysis of membrane targeting by *S. cerevisiae* pleckstrin homology domains. *Mol. Cell* 13, 677-688.
- Zhang Y, Appleton BA, Wiesmann C, Lau T, Costa M, Hannoush RN, Sidhu SS (2009) Inhibition of Wnt signaling by Dishevelled PDZ peptides. *Nat Chem Biol* 5: 217-219.
- Zimmermann P, Meerschaert K, Reekmans G, Leenaerts I, Small JV, Vandekerckhove J, David G, Gettemans J (2002) PIP2-PDZ Domain Binding Controls the Association of Syntenin with the Plasma Membrane. *Molecular Cell* 9: 1215-1225.

APÉNDICE

ARTÍCULOS



Structure Report

Novel conformational aspects of the third PDZ domain of the neuronal post-synaptic density-95 protein revealed from two 1.4 Å X-ray structures

Ana Cámara-Artigas^{a,*}, Javier Murciano-Calles^b, Jose A. Gavira^c, Eva S. Cobos^b, Jose C. Martínez^b

^a Department of Physical Chemistry, Biochemistry and Inorganic Chemistry, University of Almería, 04120 Almería, Spain

^b Department of Physical Chemistry and Institute of Biotechnology, Faculty of Sciences, University of Granada, 18071 Granada, Spain

^c Laboratorio de Estudios Cristalográficos – IACT (CSIC – Universidad de Granada), PT Ciencias de la Salud, Granada 18100, Spain

ARTICLE INFO

Article history:

Received 11 November 2009

Received in revised form 28 February 2010

Accepted 7 March 2010

Available online 19 March 2010

Keywords:

PDZ domains
X-ray structure
Succinimide

ABSTRACT

The crystal structure of the third PDZ domain of the neuronal post-synaptic density-95 protein (PSD95-PDZ3, residues 302–402) has been solved at 1.4 and 1.35 Å from two different crystal forms. These structures lack the cloning artefact present in the carboxyl terminal sequence of the former crystallographic structures and they belong to the space groups P4₃ and P1. The new PDZ structures are identical between the two crystal forms and among the four chains of the P1 crystal form. When we compare the new structures with the previous ones, some important conformational differences in the C-terminal α -helix and in the loop connecting β 2 and β 3 strands have been found. Additionally, the high resolution of the new structures has allowed us to identify a succinimide residue at the position corresponding to Asp332 in the β 2– β 3 loop, which may contribute to the alternate conformation of this loop, and at the same time, to the interaction between residues from this loop and the C-terminal α -helix. Thus, these features would have implications in the recently proposed allosteric role of this third α -helix in the binding of the carboxyl terminal fragments to the PSD95-PDZ3.

© 2010 Elsevier Inc. All rights reserved.

1. The third PDZ domain of the neuronal post-synaptic density-95 protein

PDZ (PSD-95/Discs-large/ZO-1) domains are protein–protein interaction modules approximately 90 amino acids long. These domains interact with C-terminal peptides to act as a major scaffold for other signalling proteins which participate in receptor and channel clustering in neurons (Kim and Sheng, 2004). PDZ domains are often arranged in tandem arrays and/or associated with other interaction domains to organize multidomain-scaffolding proteins. This is the case of the post-synaptic density protein 95, PSD95 (also named as SAP90 or DLG4), a member of the membrane associated guanylate kinase (MAGUK) family, which is composed by three PDZ domains followed by SH3 (Src homology 3) and GK (guanylate kinase-like) domains.

Up-to-date, the third PDZ domain of the PSD95 protein (PSD95-PDZ3) has been one of the best characterized PDZ domains, and its structure was the first one to be solved for a PDZ domain in both the presence and the absence of peptidic ligands (Doyle et al., 1996). In addition to the canonical PDZ organization containing

five or six β -stranded half β -barrel and two α -helices, PSD95-PDZ3 shows a third α -helix at the C-terminus. This third α -helix lies outside the active site and does not make direct contact with the C-terminal peptide ligand, but it has been recently reported to play an important role in the binding of carboxyl terminal sequences to the PSD95-PDZ3 domain (Petit et al., 2009). So far, all the crystallographic PSD95-PDZ3 structures solved until now belong to the same space group (P4₃2) with comparable unit cell parameters. On the same grounds, all of them show an additional antiparallel β -hairpin at the C-terminal region organized by the residues 403–415, which are not present in the original PSD95 sequence (Doyle et al., 1996). Thus, the presence of this cloning artefact might preclude some important features related to dynamic and functional processes of the protein. Our interest in these subjects has led us to solve the structure of the protein lacking the additional sequence (residues 302–402), and to compare it with the previous ones.

2. Cloning, expression and purification of PSD95-PDZ3

pQE30 plasmid coding for the PDZ3 domain was provided by Drs. José Reina and Luis Serrano (EMBL-CRG, Barcelona). The construct, including residues from 302 to 402 of the sequence (Doyle et al., 1996), was subcloned into pBAT4 vector (EMBL) by using

Abbreviations: PDB, protein data bank.

* Corresponding author. Fax: +34 950 015008.

E-mail address: acamara@ual.es (A. Cámara-Artigas).

NcoI and *HindIII* restriction sites. A methionine codon present in the *NcoI* site was used as translation initiator and a glycine residue was introduced after the first methionine to accommodate the *NcoI* restriction site in the DNA linker. PSD95-PDZ3 domain was over-expressed in *Escherichia coli* BL21 (DE3) strain. Incubation of the cells took place in LB media at 37 °C and the induction was achieved by adding IPTG at a final concentration of 0.2 mM when the OD reached a value of 0.6. After centrifugation, collected cells were lysed in 50 mM Tris at pH 7.5 in a French press. The cell lysate was acidified at pH 3 using 0.1 M HCl, and the PDZ3 domain was recovered from the supernatant by precipitation in ammonium sulfate at 75%. The precipitated protein was solubilised in 50 mM Tris buffer (pH 7.5) and 400 mM NaCl, and purified by a gel filtration step on a Hi-Load Superdex 75 column (Pharmacia). A purity of ~95% was checked by SDS-PAGE electrophoresis.

Mass spectrometry (CIC services at the University of Granada) and N-terminal sequencing (CSIC sequencing service, Madrid, Spain) of the sample revealed that the first methionine is naturally cleaved. Sample concentration was determined spectrophotometrically by using an extinction coefficient of 2985 M⁻¹ cm⁻¹ (Gill and von Hippel, 1989).

3. Crystallization, data collection, phasing and structure determination of PSD95-PDZ3

Crystallization conditions were identified using the Hampton Research Crystal Screen I and the hanging-drop vapour-diffusion method at 15 °C in 48-well plates. In brief, 2 µL of protein was mixed with 2 µL of the reservoir solution and equilibrated against 200 µL of the reservoir solution. Good-quality diffracting crystals were obtained at protein concentration of 30 mg/mL in 10 mM Tris pH 8 and the precipitant solution containing 0.2 M ammonium sulphate, 0.1 M ammonium acetate trihydrate pH 4.6, 25% w/v PEG 4000 (condition #20), and 0.1 M HEPES sodium pH 7.5, 0.8 M sodium phosphate monobasic monohydrate, 0.8 M potassium phosphate monobasic (condition #35).

The crystals grown under condition #20 were obtained using the contradiffusion technique (García-Ruiz, 2003) as described in detail in Cámara-Artigas et al. (2007) and measured at room temperature inside the capillary at the beamline BM-16 of the European Synchrotron Radiation Facility (ESRF) using a 0.97 Å wavelength in a ADSC Quantum 4R detector. Data were indexed, integrated and scaled with HKL2000 suite (Otwinowski and Minor, 1997).

The crystals grown under condition #35 were obtained in a hanging-drop setup, and transferred to a mother liquor solution containing 10% glycerol as a cryoprotectant. The looped crystals were then placed in a cold nitrogen stream (Kryoflex) maintained at 110 K. Oscillation images were collected with a Bruker Microfocus (Montel Optics) Microstar rotating anode X-ray generator operated at 45 kV and 60 mA with kappa configuration and a Proteum X8 CCD detector. Integrated intensity information for each reflection was obtained with the program SAINT and scaled with the SADABS module of the PROTEUM suite. The crystallographic parameters and statistics of the data collection are listed in Table 1.

The initial phasing of the tetragonal structure was obtained using Molrep (Vagin and Teplyakov, 1997) taking as reference the coordinates for PSD95-PDZ3 domain (Doyle et al., 1996) without the cloning artefact. Several cycles of restrained positional and temperature factor refinement using Refmac (Murshudov et al., 1997) were alternated with manual building using the resulting σA -weighted ($2F_o - F_c$) and ($F_o - F_c$) electron density maps and the COOT program (Emsley and Cowtan, 2004).

Table 1
X-ray data collection and refinement statistics.

Space group	P4 ₃	P1
Unit cell dimensions a, b, c	42.98, 42.98, 46.89	43.32, 42.15, 47.60
		90.00, 90.00, 92.08
Resolution range (Å)	20–1.4	15–1.35
Number of observations	16934(3138)	67847 (6858)
Redundancy	5.88 (2.72)	2.4 (2.3)
Data completeness (%)	100 (100)	92.0 (92.9)
R _{merge} (%) ^a	5.48 (30.8)	6.20 (41.1)
I/σ (I)	17.05 (3.10)	16.07 (1.84)
Matthews coefficient/Solvent content	1 mol/AU 1.93/36%	4 mol/AU 1.93/36%
<i>Refinement</i>		
Protein residues	98	392
Solvent	97	250
R _{work} (%)	17.1 (24.8)	18.9 (28.2)
R _{free} (%)	21.1 (26.3)	21.9 (29.1)
<i>RMS deviations from ideal geometry</i>		
Bonds (Å)	0.024	0.022
Angles (°)	2.093	2.172
Mean B (protein) (Å ²)	17.35	17.97
Residues in allowed regions of the Ramachandran plot (%) ^b	100	100

The values in parentheses are for the highest resolution bin.

^a $R_{merge} = \left[\frac{\sum_i \sum_j |I_{ij} - \langle I_{ij} \rangle|}{\sum_i \sum_j I_{ij}} \right] \times 100\%$, where $\langle I_{ij} \rangle$ is the mean intensity of i reflections with intensity I_{ij} .

^b From program PROCHECK statistics.

The triclinic structure was solved by following the same procedure, but in this case the coordinates of the tetragonal structure were used as the search model. Refinement statistics for both structures are collected in Table 1.

When we compare the single chain of the tetragonal structure with the four chains present at the asymmetric unit of the triclinic crystal form, all the chains show main chain rmsd values lower than 0.4 Å, allowing us to affirm that they are almost identical among them. The most important differences between the chains of both structures are found in the residues 350–352 belonging to the loop connecting the helix $\alpha 1$ and the strand $\beta 4$ (maximum rmsd ~1.5 Å) and in the residue 384 at the beginning of the strand $\beta 5$ (maximum rmsd ~3 Å). It is worth noticing that these differences are within the range of the ones found when the four chains present in the triclinic structure are compared between them. Besides, these four chains are arranged in a tetramer that is almost identical to the one generated by the symmetry operators of the higher symmetry space group P4₃ with a slight displacement of the molecules, which would explain the loss of the symmetry (Fig. 1). Additionally, the composition of each precipitant solution used for the crystallization resulted in different solvent molecules placed at the binding site and crystal contacts. Thus, the only remarkable difference between both structures seems to be the distribution of solvent molecules present at the asymmetric unit of each crystal form.

4. Comparison with previous PSD95-PDZ3 X-ray structures

It is worth remarking that the structures reported in this work represent the first structural characterization of the PSD95-PDZ3 domain in the absence of the additional N- and C-terminal sequences coming from the cloning vector, and that some important differences arise when we compare our high-resolution structures with those previously reported. The rmsd plot corresponding to the main-chain atom positions comprised average values between 1.3

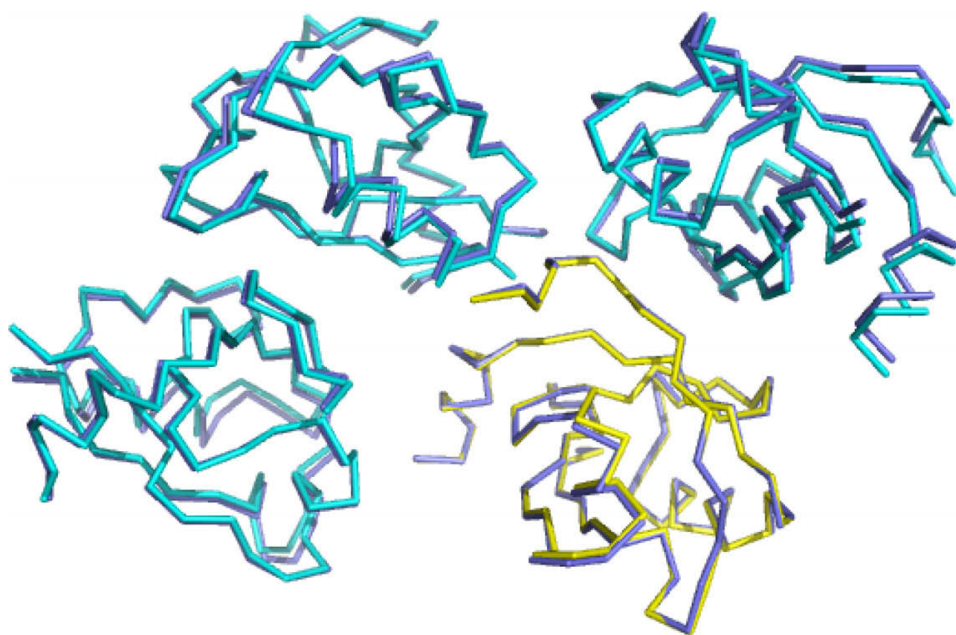


Fig. 1. Superposition of the main chain of the tetragonal structure of the PSD95-PDZ3 (yellow) and the chains of the tetramer obtained by symmetry (cyan) with those of the triclinic structure (blue).

and 1.5 Å and, as can be expected, the major differences in rmsd values match up with the terminal regions of the protein, but also with the loops connecting the $\beta 1$ – $\beta 2$ and $\beta 2$ – $\beta 3$ strands (Fig. 2A).

The additional C-terminal amino acids present in the former PSD95-PDZ3 structures appear well ordered and folded in a β -hairpin packed against the third α -helix and, at the same time, they participate in the biggest interface present in the crystal (ASA buried ~ 700 Å²) through various intra- and inter-chain hydrogen bonds and salt bridges contacts. In the new structures the lack of this additional sequence results in a changed conformation of the third α -helix (residues 393–402, Fig. 2B), arising from differences in the intra- and inter-chain hydrogen bonds and salt bridges formed by the side chains of the residues. One of the most noticeable changes is found in the residue Tyr397, which in the previous

structures is placed at hydrogen bond distance of the side chain of Glu401, but this interaction is not present in the new structures. This tyrosine residue has been reported to be phosphorylated (Ballif et al., 2008) and it has been equally suggested to play an important role in the PSD95-PDZ3 regulation *in vivo* (Petit et al., 2009). Therefore, the changes in the hydrogen bond network around this residue could have a deep impact in the dynamics of the protein. There are several residues near to Tyr397 that could stabilize the phosphorylation site by means of saline bridges and hydrogen bonds: Lys355, Ser339 and Glu401. When we compare the position of these residues in the former structures and the new ones, the most striking difference is found in the orientation of the side chain of Glu401 (Fig. 2B). It, instead of interacting with Tyr397, moves to form a new interaction with Glu334. As a consequence, the dis-

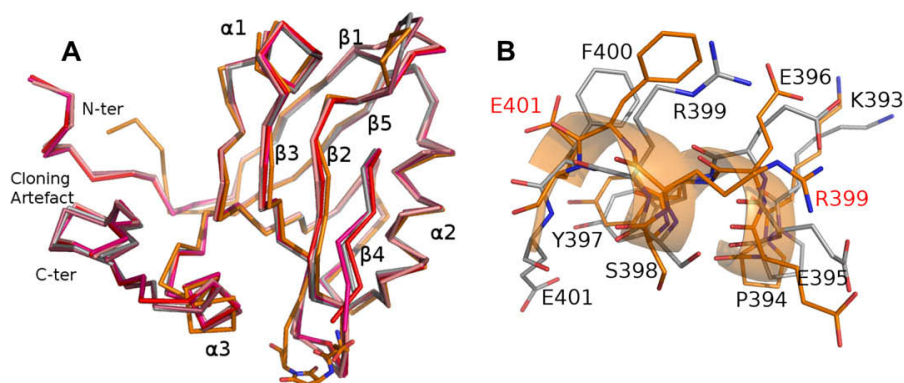


Fig. 2. (A) Superposition of the tetragonal structure of the PSD95-PDZ3 from this work (orange) and previous structures of the bound and unbound PSD95-PDZ3 domain: 1BE9, pink; 1BEF, gray; 1TP3, magenta; 1TP5, red; 1TQ3 raspberry. Each peptide is shown in the same colour as the backbone of the structure. (B) Comparison of the third α -helix of the unbound structure 1BEF (gray) and of the structures of this work (orange).

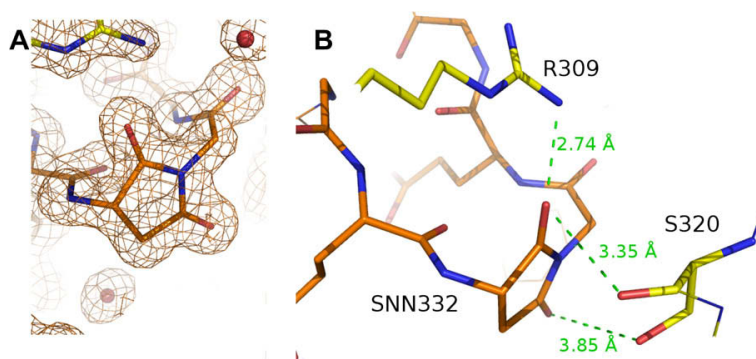


Fig. 3. (A) $2F_o - F_c$ map of the loop connecting $\beta 2$ and $\beta 3$ strands. (B) Residues Arg309 and Ser320 of two symmetry related molecules are shown in yellow sticks and their respective hydrogen bonds formed with the carbonyl of the succinimide are shown in green.

tance between the backbone of the third α -helix and the loop formed by the residues Gly329–Gly335, which connect the $\beta 2$ – $\beta 3$ strands, is shortened in 3 Å. These features may have some important consequences since the C-terminal α -helix of PSD95–PDZ3 strongly influences ligand binding (Petit et al., 2009).

5. Aspartate 332 is a succinimide in the 1.4 Å PSD95–PDZ3 structures

As it has been indicated above, the loop connecting $\beta 2$ and $\beta 3$ strands, formed by the residues Gly329–Gly335 (GGEDGEG), is the second region of PDZ3 with higher-than-average rmsd values. Besides, this loop shows different orientations in the five previous structures and the side chain of the residue Asp332 has been modelled just for the one with PDB code 1TQ3. Nevertheless, the higher resolution of our data (1.35 and 1.4 Å) has allowed us to identify the presence of a succinimide residue at the position of Asp332 in both structures (Fig. 3). Succinimide ring formation is a common intermediate step in the processes of aspartate isomerization and asparagine deamidation in peptides and proteins (Clarke, 1987).

From a conformational point of view, the presence of this succinimide residue induces the decrease of the loop flexibility at the same time that it results in the loss of a negative charge from a carboxyl group on the protein surface, and both processes could play some decisive roles in both crystal nucleation and growth. Accordingly, the residues in this loop, together with the amino terminal residues, show the highest *B*-factors in all the previous crystal structures, while in our succinimide-containing structures this loop shows *B*-factor values below the observed average value. In addition, the lack of a carboxylate group results in a decrease of the solubility of the molecules containing the succinimide residue, which favours their crystallization (Noguchi et al., 1998). Furthermore, the direct participation of the succinimide residue in the crystal contacts implies that in our case this residue must be present in the protein solution before the crystal formation takes place in order to participate in the crystalline contact (Fig. 3B). Besides, the hydrogen bonds formed by the residues implied in this contact might avoid the nucleophilic attack of water molecules that initiates the hydrolysis of the succinimide, thus preventing its conversion into isoaspartyl or aspartil residues.

The recent experimental evidence of an allosteric network that allows the C-terminal α -helix to modulate binding affinities of PSD95–PDZ3 (Petit et al., 2009) might thus be influenced by succinimide conversion since, as we stated above, the interactions established between polar residues located at the C-termi-

nal α -helix and the $\beta 2$ – $\beta 3$ loop may also contribute to the positioning of the helix and, as a consequence, to modulate binding energetics.

The importance at physiological level of the appearance of the succinimide residues is still unknown. However, the ubiquitous presence of the enzyme D-aspartyl/L-isoaspartyl methyltransferase (PIMT), whose role is to recover isoaspartyl residues to aspartate, indicates that these modified residues would have some functional role. Up to now we have not found any evidence of the presence of a succinimide or isoaspartate residue in the PSD95 *in vivo*, and we do not know how the presence of the succinimide residue, or its product isoaspartate, would affect to its function. It is our conviction that, although a deeper characterization of the biological role of succinimides is obviously needed, structural information will be essential for the future studies.

Atomic coordinates

The atomic coordinates and structure factors have been deposited at the Research Collaboratory for Structural Bioinformatics with PDB ID 3K82 (1.4 Å, P4₃) and 3I4W (1.35 Å, P1).

Acknowledgments

This research was funded by Grants BIO2006-15517-C02-01/02 and BIO2009-13261-C02-01/02 and P09-CVI-5063. We thank to the project OptiCryst of the VI Framework program from UE and the “Factoría Española de Cristalización” Ingenio/Consolider 2010. J.M.C. is recipient of a FPU research contract from the Spanish Ministry of Education and Sciences. E.S.C was funded by a contract provided by the University of Granada.

References

- Ballif, B.A., Carey, G.R., et al., 2008. Large-scale identification and evolution indexing of tyrosine phosphorylation sites from murine brain. *J. Proteome Res.* 7 (1), 311–318.
- Cámara-Artigas, A., Palencia, A., et al., 2007. Crystallization by capillary counter-diffusion and structure determination of the N114A mutant of the SH3 domain of Abl tyrosine kinase complexed with a high-affinity peptide ligand. *Acta Crystallogr. D Biol. Crystallogr.* 63 (Pt. 5), 646–652.
- Clarke, S., 1987. Propensity for spontaneous succinimide formation from aspartyl and asparaginyl residues in cellular proteins. *Int. J. Pept. Protein Res.* 30 (6), 808–821.
- Doyle, D.A., Lee, A., et al., 1996. Crystal structures of a complexed and peptide-free membrane protein-binding domain: molecular basis of peptide recognition by PDZ. *Cell* 85 (7), 1067–1076.
- Emsley, P., Cowtan, K., 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* 60, 2126–2132.

- García-Ruiz, J.M., 2003. Counterdiffusion methods for macromolecular crystallization. *Methods Enzymol.* 368, 130–154.
- Gill, S.C., von Hippel, P.H., 1989. Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* 182 (2), 319–326.
- Kim, E., Sheng, M., 2004. PDZ domain proteins of synapses. *Nat. Rev. Neurosci.* 5 (10), 771–781.
- Murshudov, G.N., Vagin, A.A., et al., 1997. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* 53 (Pt. 3), 240–255.
- Noguchi, S., Miyawaki, K., et al., 1998. Succinimide and isoaspartate residues in the crystal structures of hen egg-white lysozyme complexed with tri-*N*-acetylchitotriose. *J. Mol. Biol.* 278 (1), 231–238.
- Otwinowski, Z., Minor, W., 1997. *Processing of X-ray Diffraction Data Collected in Oscillation Mode*. Academic Press.
- Petit, C.M., Zhang, J., et al., 2009. Hidden dynamic allostery in a PDZ domain. *Proc. Natl. Acad. Sci. USA* 106 (43), 18249–18254.
- Vagin, A., Teplyakov, A., 1997. MOLREP: an automated program for molecular replacement. *J. Appl. Crystallogr.* 30, 1022–1025.

An Oligomeric Equilibrium Intermediate as the Precursory Nucleus of Globular and Fibrillar Supramacromolecular Assemblies in a PDZ Domain

Javier Murciano-Calles,[†] Eva S. Cobos,[†] Pedro L. Mateo,[†] Ana Camara-Artigas,[‡] and Jose C. Martinez^{†*}

[†]Department of Physical Chemistry and Institute of Biotechnology, Faculty of Sciences, University of Granada, Granada, Spain;

and [‡]Department of Physical Chemistry, Biochemistry and Inorganic Chemistry, Faculty of Experimental Sciences, University of Almería, Almería, Spain

ABSTRACT The equilibrium unfolding at neutral pH of the third PDZ domain of PSD95, as followed by DSC, is characterized by the presence of an equilibrium intermediate with clear signs of oligomerization. DLS and SEC measurements indicate that at 60–70°C small oligomers populate, showing a typical β -sheet far-UV CD spectrum. These intermediate species lead to the formation of rodlike particulates of ~12 nm, which remain in solution after 2 weeks incubation and grow until they adopt annular/spherical shapes of ~50 nm and protofibrils, which are subsequently fully transformed into fibrils. The fibrils can also disaggregate after the addition of 1:1 buffer dilution followed by cooling to room temperature, thus returning to the initial monomeric state. Growth kinetics, as shown by ThT and ANS fluorescence, show that the organization of the different supramacromolecular structures comes from a common nucleation unit, the small oligomers, which organize themselves before reaching the incubation temperature of 60°C. Our experiments point toward the existence of a well-defined reversible, stepwise, and downhill organization of the processes involved in the association-dissociation of the intermediate. We estimate the enthalpy change accompanying the association-dissociation equilibria to be $130 \text{ kJ} \times \text{mol}^{-1}$. Furthermore, the coalescence under essentially reversible conditions of different kinds of supramacromolecular assemblies renders this protein system highly interesting for biophysical studies aimed at our further understanding of amyloid pathological conditions.

INTRODUCTION

PDZ domains constitute the most abundant protein interaction modules in mammals, where they have been identified as the main structural component of a variety of the so-called hub proteins. These highly connected protein-nodes may have hundreds of links, either simultaneously or sequentially, suggesting that hubs organize the proteome by acting as molecular engines that connect biological processes. As far as the PSD95 belonging to the MAGUK family, is concerned, its three PDZ domains, together with a SH3 domain and a guanylate kinase module, are involved in the complex process of synapse scaffolding (1). It is surprising that only five functional domains might be responsible for the organization of the dynamic trafficking of a variety of synaptic proteins at the inner face of the neuron membrane. The explanation may well be that the intrinsic structural plasticity of hub proteins derives from the corresponding plasticity of their domains and linker sequences, which are able to

explore the conformational space to discover the different strategies necessary to develop their various functions. In this way, the correlation between folding (stability) and binding (peptide dissociation rate) parameters shown recently for the PTPBL-PDZ2 domain strongly suggests that these domains may depend on their stability and dynamics to develop their function (2).

A computational and mutational study has shown that an evolutionarily conserved network of interacting residues exists in the PSD95-PDZ3 domain, which may work as a communication pathway between distal parts of the molecule (3). Nevertheless, there is little evidence, if any, of a conformational transition on binding that might be associated with a functional allosteric network in PDZ3, the process being a pure lock-key mechanism (4,5). In addition, the few equilibrium-unfolding experiments carried out with PDZ domains (including PDZ3) using spectroscopic probes have shown a single transition that can be described by the two-state model (2,6). Given their presumable functional flexibility and their relatively greater size (>100 amino acids) compared to the small, two-state, globular proteins studied to date, it would be unlikely that these domains do not contain any marginally stable conformations under equilibrium conditions.

Thus we decided to study the folding behavior of the PSD95-PDZ3 domain at neutral pH in far greater depth by using calorimetric and spectroscopic information. Contrary to previous evidence, our DSC experiments show two unfolding transitions, which undoubtedly indicate the existence

Submitted November 30, 2009, and accepted for publication April 1, 2010.

*Correspondence: jcmh@ugr.es

Abbreviations used: ANS, 1-anilino-naphtalene-8-sulfonic acid; CD, circular dichroism; DLS, dynamic light scattering; DSC, differential scanning calorimetry; MAGUK, membrane-associated guanylate kinase; PDZ, postsynaptic density-95 protein, disk-large tumor suppressor protein, zonula occludens-1; PSD95, postsynaptic density-95 protein; PSD95-PDZ3 or PDZ3, third PDZ domain of PSD95 protein; PTPBL-PDZ2, second PDZ domain of protein tyrosine phosphatase-BL; SEC, size-exclusion chromatography; SH3, Src-homology 3 domain; TEM, transmission electron microscopy; ThT, thioflavine T; UV, ultraviolet.

Editor: Jane Clarke.

© 2010 by the Biophysical Society
0006-3495/10/07/0263/10 \$2.00

doi: 10.1016/j.bpj.2010.04.003

of at least three well-defined macroscopic states on unfolding. Moreover, the endotherms tend to separate along the T axis concomitantly with an increase in protein concentration, suggesting some degree of association of the intermediate state. The characterization of the size distribution of these oligomerization phenomena by DLS, SEC, and TEM shows that the monomeric native state partially unfolds on heating, thus generating small oligomeric species that lead firstly to disklike assemblies and protofibrils, and then to amyloid fibrils after incubation on various days at 60°C. A further conformational analysis of the intermediate state shows a high content of β -sheet structure as well as the absence of any lag-phase in time-dependent ThT and ANS fluorescence measurements.

The occurrence of globular and fibrillar arrangements in PDZ3 (simultaneous at some stages) may help to answer some crucial questions within this field of inquiry, which is of key interest toward our understanding of amyloid-generating pathologies and the development of effective therapies against them. So far, increasing evidence suggests that soluble oligomeric intermediates, rather than fully formed fibrils, are likely to be the primary toxic species, fibrillar aggregates being merely a storage or even protective mechanism (7). Within this context, cell degeneration and pathophysiology may derive from the organization of pores in membranes by such oligomers, most probably by specific ionic transport via ion channels, which would destabilize ionic homeostasis *in vivo* (8–10).

RESULTS

PDS95-PDZ3 thermal unfolding shows the population of an oligomeric equilibrium intermediate at neutral pH

The few equilibrium unfolding studies carried out to date using spectroscopic probes (fluorescence and CD) show that PDZ domains display a single sigmoid transition on both thermal and chemical denaturation. This behavior has been analyzed by using a two-state folding mechanism (6,11–13). We have measured the CD thermal unfolding of PSD95-PDZ3 in 50 mM K-phosphate buffer at pH 7.5 and been able to corroborate this behavior (data not shown). Nevertheless, nothing can be concluded at this juncture, as will be shown below.

Among the variety of biophysical techniques available, DSC would seem to be the best choice to detect any equilibrium conformation(s) that might be populated during unfolding. In this thermodynamic study, in what we believe to be the first reported for a PDZ domain to date, we have explored the unfolding of PSD95-PDZ3 in 50 mM K-phosphate buffer at pH 7.5 (Fig. 1). Protein preparation and experimental details are described in the Materials and Methods section in the [Supporting Material](#). The thermal unfolding profiles were partly reversible during the timescale of the DSC exper-

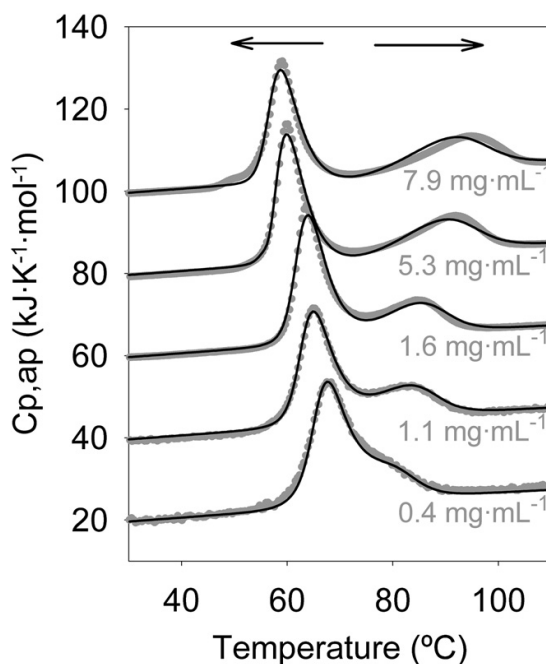


FIGURE 1 Thermal unfolding profiles of PSD95-PDZ3 monitored by DSC as a function of protein concentration. Experimental conditions were 50 mM K-phosphate, pH 7.5. The best fitting to the $N \rightleftharpoons I_n \rightleftharpoons U$ model is shown as black solid lines through the experimental data (gray symbols).

iments (at least 60%) and the scan rate did not affect their shape in any appreciable way (data not shown). The reversibility of some heat-denatured solutions was further checked by far-UV CD spectrum analysis, which showed that the complete N -state spectrum can be arrived at after a few hours. These features indicate clearly the slow relaxation times of these association phenomena but do not contradict their equilibrium nature at neutral pH.

Despite the evidence of the spectroscopic probes, however, the two unfolding transitions observed in Fig. 1 for all protein concentrations undoubtedly indicate the existence of not less than three well-defined macroscopic states during unfolding. Furthermore, both endotherms distance themselves from each other along the T axis concomitantly with an increase in protein concentration, suggesting some degree of association of the intermediate state. Thus, the movement of the first transition to lower temperatures (corresponding to the $N \rightleftharpoons I$ equilibrium) indicates an increase in the association stage of the final I -state during unfolding, whereas the opposite is true for the second transition (corresponding to the $I \rightleftharpoons U$ equilibrium). The effect of protein concentration on DSC traces has been described elsewhere in some other examples and is a consequence of the increase in population of the associated species concomitant with protein concentration (14,15). This characteristic

behavior may in fact explain the apparent single-transition two-state behavior observed previously because at the low PDZ3 concentrations used in spectroscopic thermal unfolding experiments the associated intermediate does not populate significantly enough to be observed. Thus it can be seen in Fig. 1 that the behavior of the DSC trace at $0.4 \text{ mg} \times \text{mL}^{-1}$ is more that of a single-transition than those obtained at higher protein concentrations, where both endothermic effects are well distinguished. In short, PDZ3 unfolding can be summarized by the equilibrium scheme



This model fully describes the experimental unfolding behavior of PDZ3, as will be explained in the following sections.

The oligomeric intermediate can lead to the reversible organization of annular/spherical and fibrillar aggregates at neutral pH

To characterize the postulated association-dissociation equilibria we undertook DLS experiments as a function of temperature (Fig. S1). A PDZ3 solution at a concentration of $8 \text{ mg} \times \text{mL}^{-1}$ in 50 mM K-phosphate pH 7.5 was heated initially from 20 to 60°C, which is the maximum operating temperature of the instrument and high enough to ensure that the I_n species should be most populated according to DSC analysis (see next section). The temperature was then kept at 60°C, at which moment the mass evolution of the species (raw relative intensity) as a function of incubation time was recorded (Fig. S1). Finally, the protein sample was cooled down inside the instrument to 20°C. A summary of the spectra obtained at different times after reaching this temperature is set out in Fig. S1.

As can be seen in Fig. S1, at low temperature the size distribution approaches that of a monomer ($R_h = 1.8 \text{ nm}$).

Nevertheless, at 60°C the size increases until reaching that corresponding to trimeric species ($R_h = 2.6 \text{ nm}$) when analyzed according to the globular protein model provided by the software of the instrument. Some other heavier particles ($R_h = 12 \text{ nm}$) also appear and become the major components of the solution after ~2 h incubation (Fig. S1). These particles, each containing ~100 monomers, increase in size when incubated for longer periods. Finally, on cooling the protein to 20°C, the 12 nm species appears once more in solution, as do the monomeric state and the far-UV CD spectra of the N -state after a few hours (Fig. S1). All these features support the reversible nature of these association reactions.

We incubated PDZ3 samples for longer periods of time at 60°C, during which the formation of supramacromolecular structures was followed by TEM imaging. In Fig. 2 we show a history of the evolution of such assemblies over time. In accordance with the DLS evidence, a background of rodlike particles of ~10–20 nm can be seen at the beginning (Fig. 2 A) and remain in the incubated solution for various days, growing to ~50 nm (Fig. 2, C and D). It is difficult to discern from TEM images what the true tridimensional shape of the globular aggregates might be; they may be considered to be either roughly annular, as can be deduced from the darkening of the circles toward the center (Fig. 2 D, inset), or possibly spheres, although in this case the darkening of the particles should be in the opposite direction. These globular particles coexist with other assemblies such as protofibrils (Fig. 2 B) and, eventually, a few large, thin fibrils can be made out (Fig. 2, A and E). The images taken over periods >15 days show only fibrillar structures (Fig. 2 G).

The final arrangement of the fibrils also seems to coexist reversibly with circular formations, because they disappear from the solutions incubated for 1 month on diluting 1:1 with buffer and cooling to room temperature, which

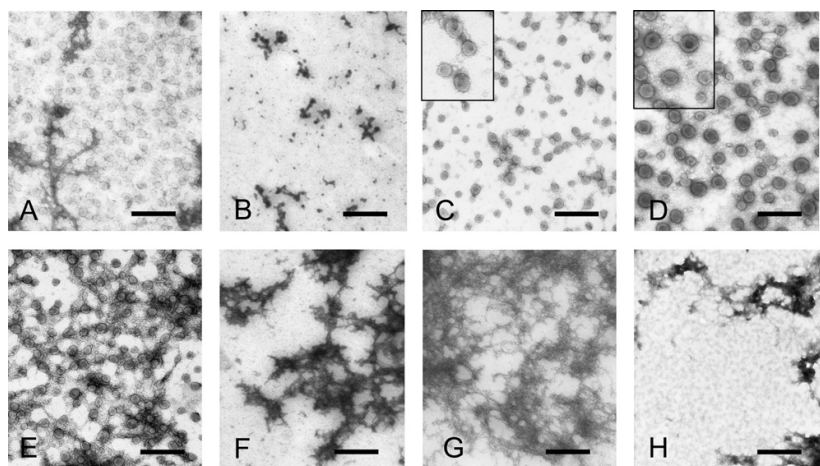


FIGURE 2 Transmission microscope analysis of PSD95-PDZ3 supramacromolecular assemblies. From left to right we show the globular particles obtained after (A) 3 and (B) 5 days incubation of the PDZ3 sample ($8 \text{ mg} \times \text{mL}^{-1}$) at 60°C and pH 7.5, the globular species and thin fibrils after (C) 8 and (D–F) 15 days, and (G) the fibrils that appear after 1 month. (H) Microaggregates seen after ~15 h reincubation of a 1:1 buffer-dilution on a 1-month-old sample. The scale bar corresponds to a size of 200 nm. The insets in the C and D are an enlargement of some of the globular motifs.

accelerates the disaggregation of the fibrils to within a few minutes. In addition, the background of 12 nm rodlike structures appeared again when the diluted samples were reincubated at 60°C for ~15 h (Fig. 2 H). Whatever the case, the essentially reversible character of this process does not seem to be 100%, because some aggregates remain in the solution after dilution (Fig. 2 H). It may be that irreversibility increases concomitantly with the size of the fibrils and ageing, as has been reported elsewhere for other examples (16).

In summary, it seems that the oligomeric equilibrium intermediate organizes itself into a set of globular particles and protofibrils at the early stages of the incubation reaction and that these persist for 2 weeks incubation, after which they all become fibrillar. The most interesting feature that emerges from DLS and TEM imaging is definitely the coexistence in a reversible fashion of both fibrils and annular/spherical arrangements in the same PDZ3-incubated solution at neutral pH. The images taken after a few days incubation at 60–70°C show that some thin protofibrils seem to appear (Fig. 2 B), in some cases as though they were emerging from the surfaces of the globular species (Fig. 2 E), perhaps growing at the expense of the latter because we could not find any trace of circular-shaped particles within the oldest incubated samples. A further imaging analysis using more powerful instruments is needed to settle these matters.

Insights toward energetic and conformational aspects of the oligomeric intermediate through the equilibrium analysis of PSD95-PDZ3 unfolding

We carried out some fitting sessions of DSC experiments using the three-state association-dissociation model (Scheme 1), as described in Materials and Methods in the Supporting Material. The DLS results (Fig. S1) indicate that the degree of association of the I_n -state can be globally assumed as $n = 3$ because during the initial 20 min of incubation at 60°C most of the species had an average hydrodynamic radius of 2.6 nm, which is close to the corresponding molecular size of trimeric species (33 kDa). This period roughly reflects the time that the protein sample remains at these high temperatures inside the DSC cell because the instrument is usually operated at scan rates ranging between 1–2°/min. As can be seen in Fig. 1, the quality of the nonlinear curve fitting of the calorimetric traces leads us to conclude that the model fully describes the experimental unfolding behavior of PSD95-PDZ3.

DSC experiments have also been fitted to the model, keeping the value of n free to float. The results of this fitting session give a value of $n = 3.8 \pm 0.3$ and also the best R - and R^2 -values (0.996; Fig. S2). Correspondingly, as shown in Fig. S2, the fitting with a fixed value of $n = 4$ meets the convergence criteria somewhat better than the one with $n = 3$. There are some small deviations in the second endotherms of traces at high protein concentrations, but the fitting with $n = 3$ reproduces the shape of the trace at $0.4 \text{ mg} \times \text{mL}^{-1}$ more satisfactorily. The use of other values for n , from 2 to 6, resulted in worse fittings (Fig. S2). The explanation of these features might be that as protein concentration increases the larger species become more populated concomitantly with the stabilization of the intermediate state and thus their influence in DSC experiments is somewhat more noticeable. The thermodynamic parameters in Table 1 have been calculated from the fitting sessions using $n = 3$, where the error intervals have been estimated by comparing the parameters obtained at both $n = 3$ and $n = 4$ values. These errors are slightly higher and possibly more realistic values than fitting standard errors.

The dependence on concentration of DSC traces has been studied extensively in the past by different research groups, including ours (14,15). These studies have shown clearly that the relative displacement of traces obtained under the same solvent conditions can be exclusively attributed to protein concentration differences and, thus, the global stoichiometry of the association-dissociation process can be precisely obtained from the evaluation of such displacement through global fitting analysis of the whole set of traces (as we have done in this study), as long as the protein concentrations have been correctly measured for the set of experiments. We obtained additional evidence of these association phenomena from DLS and SEC experimental approaches (Fig. S1). As well as the DSC evidence, the DLS experiments carried out at low protein concentration ($0.4 \text{ mg} \times \text{mL}^{-1}$) do not show any significant population of aggregated species at any time of incubation at 60°C, which is further supported by the absence of any fluorescence enhancement of solutions incubated with ANS or ThT (data not shown).

The hypothetical trimeric nature of the oligomeric species is shared by other examples (8,17) and can be structurally understood here from a computational analysis of PDZ3 sequence propensity to β -aggregation. The calculations using the Tango algorithm developed by Fernandez-Escamilla et al. (18) show that propensities are negligible over

TABLE 1 Thermodynamic parameters of the thermal unfolding of the PSD95-PDZ3 domain in 50 mM K-phosphate pH 7.5 obtained from the analysis of DSC experiments

T_{N-U} (°C)	$\Delta H_{N-U}(T_{N-U})$ (kJ \times mol $^{-1}$)	$\Delta G_{N-U}(298)$ (kJ \times mol $^{-1}$)	T_{U-I_n} (°C)	$\Delta H_{U-I_n}(T_{U-I_n})$ (kJ \times mol $^{-1}$)	$\Delta G_{U-I_n}(343)$ (kJ \times mol $^{-1}$)
70.4 ± 0.5	335 ± 20	39 ± 6	79.2 ± 1.2	-130 ± 20	25 ± 5

The error intervals were calculated as described in the text. The values of thermodynamic magnitudes for the $U \rightleftharpoons I_n$ equilibrium were estimated for a $P_{\text{ref}} = 100 \mu\text{M}$. The heat capacity functions obtained from the fitting were $C_{\text{pN}} = -9.23 + (0.095^*T)$; $C_{\text{pU}} = 2.91 + (0.064^*T)$; $C_{\text{pin}} = -99.0 + (0.347^*T)$ in $\text{kJ} \times \text{mol}^{-1}$.

the whole domain with the exception of residues ranging from 340 to 350 (pertaining to the $\beta 3$ and $\alpha 1$ motifs) and for the 392–397 ones (located at $\beta 5$ and $\alpha 3$ secondary structures). The tridimensional arrangement of these residues in the high-resolution x-ray structures, recently solved by us (PDB codes: 3K82 and 3I4W (19)), shows two main interfaces roughly located at opposite faces of the globular structure, which may allow each monomer to interact with two additional ones (Fig. S3). Thus, it is possible that the increase in solvent exposure of these aggregation-prone regions during unfolding may in PDZ3 monomers develop some tendency to self-associate.

In any case, we cannot rule out the idea that the experimentally observed oligomeric arrangement would result from a mixture of dimers, trimers, tetramers, or other oligomers, where the fraction of the lower-sized species decreases concomitantly with a decrease in protein concentration, which might be more probable from a mechanistic point of view. In fact, the coexistence of different association stages for these nucleation species has been reported for some other amyloid-forming proteins (20). The existence of such additional equilibria might also constitute an alternative explanation for the small DSC deviations observed in the fitting sessions and may contribute to some extent to the greater sharpness and asymmetry of the first endotherm (Fig. 1). To investigate these issues, we conducted SEC experiments at pH 7.5 and 60°C with 8 mg \times mL⁻¹ PDZ3 samples incubated for different periods of time (Fig. S1). A broad peak appears, probably due to the presence of a mixture of sizes, albeit the maximums observed correspond to monomers, high aggregates (at the exclusion volume) and, once again, roughly trimeric species.

To achieve a better understanding of the energetics of the intermediate state we dissected the two DSC endotherms (Fig. 1) into their respective heat contributions. To this end, we developed the equations corresponding to the model scheme described by Eq. 2 of the Supporting Material. This alternative scheme, mathematically equivalent to the one in Eq. 1 of the Supporting Material, allows us to obtain separately the contributions corresponding to the unfolding and association-dissociation processes, which may help to understand the unfolding process phenomenologically. Thus, the deconvolution analysis shown in Fig. 3 shows that the evolution of heat during the unfolding process occurs exclusively during the first endotherm ($N \rightleftharpoons I_n$), which seems to be the result of the combination of two processes of opposite signature: the positive contribution of the whole heat of the protein unfolding process and the exothermic component deriving from the association phenomena of the intermediate state. Nevertheless, the second evolution can be attributed merely to the heat released by the disruption of interdomain interactions during the dissociation of I_n oligomers into monomeric molecules (Fig. 3 B), which does not necessarily include any other heat contribution deriving from the further unfolding of the intermediate. Consequently,

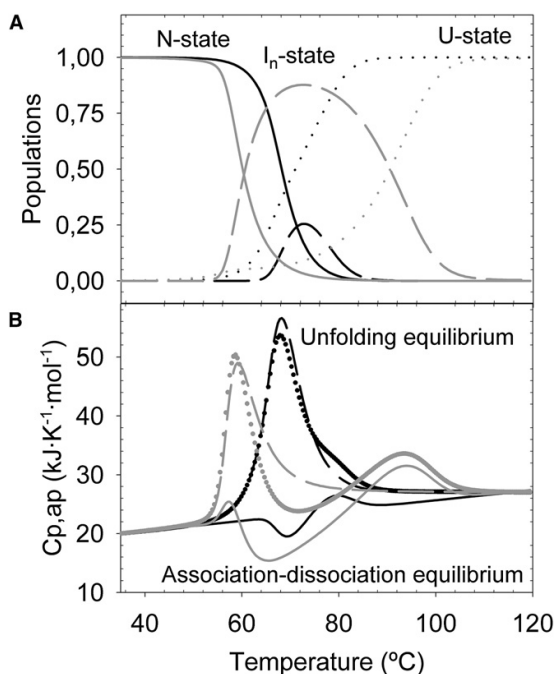


FIGURE 3 Deconvolution analysis of the different equilibria for PSD95-PDZ3 thermal unfolding. (A) Distribution of populations of the different PDZ3 states for the two extreme protein concentrations assayed, 0.4 mg \times mL⁻¹ (black) and 8 mg \times mL⁻¹ (gray). An increase in I_n -state population can be seen on protein concentration. (B) Dashed lines show the respective thermal evolutions of the unfolding equilibrium ($N \rightleftharpoons U$) and the solid lines the association-dissociation term ($U \rightleftharpoons I_n$). Symbols are used for experimental DSC data as a function of protein concentration. As can be seen, the latter equilibrium is symmetrical in shape because the intermediate associates first (exothermic half) and dissociates afterward (endothermic half) to complete unfolding.

these findings indicate that from an energetic point of view the I_n -state may resemble a partly unfolded state in which most of the tertiary cooperative interactions are relaxed, rather than a highly structured, cooperative, natively-like arrangement. The nonnative character of this partly folded state is also revealed by its incapability of recognizing the high-affinity PDZ3 ligand KKETAV, which in no way influences the history set out in Fig. 2.

Whatever the case, the high quality of the fittings (Fig. 1 and Fig. S2) cannot in itself be considered conclusive with regard to the completely noncooperative character of the unfolding of this intermediate and so we cannot definitely rule out the existence of more-or-less well ordered regions in this state. The oligomers may possibly maintain some local structural arrangements or perhaps contain new structural elements not present in the native state, and therefore this state may be considered as being partly folded even though it is not fully homogeneous. The SEC experiments (Fig. S1), for example, indicate that at 60°C monomeric

species of this intermediate state can be discerned, which supports the idea that it displays some thermodynamic stability by itself. The smallness of the structural arrangements will lead to a relatively small and rather wide heat effect.

Spectroscopic and kinetic analysis of the early conformational stages of the PDZ3 misfolding pathway

We investigated the structural ordering of the intermediate by far-UV CD spectroscopy. Fig. S4 includes the PDZ3 spectra at 25°C and 70°C, the two key temperatures for the *N*-state and the *I_n*-state respectively (Fig. 3 A). The spectrum at 70°C shows a high secondary-structure content for the associated states, which, according to the minimum observed at 217 nm, would be mainly β -sheet structures. These β -sheet organizations remained unaltered after several hours at this temperature. The absence of Trp residues, together with the low aromatic content within the hydrophobic core of PDZ3, prevented any further CD research into tertiary arrangements at the near-UV region.

Congo red and ThT dyes have been used traditionally as histochemical stains for amyloids (21). Both gave positive results when assayed with the mature amyloid fibrils of PDZ3, the former leading to a blue shift in the UV-visible spectrum (data not shown) and the latter to a noticeable enhancement of fluorescence signal (Fig. S5). Assays with the ANS fluorescent probe were also positive. This probe has been reported to bind solvent-exposed hydrophobic surfaces in proteins, producing an increase in fluorescence signal together with a blue shift in the maximum wavelength emission (Fig. S5), which indicate clearly the accumulation of partially structured species with solvent-exposed nonpolar surfaces.

To achieve some insight into the sequence of events that develop during the formation of PDZ3 fibrils we carried out a real-time continuous monitoring of the fluorescence signal from ThT and ANS (Fig. 4). The $8 \text{ mg} \times \text{mL}^{-1}$ PDZ3 sample was heated inside the fluorimeter to 60°C (Fig. 4 A) and, once at this temperature, the time-course of the fluorescence signal was monitored during the first few hours of the aggregation reaction at 60°C (Fig. 4 B). Finally, some fluorescence spectra were obtained after cooling the sample down to room temperature (Fig. S5). Thus, these time-course experiments paralleled the DLS ones described previously. When the PDZ3 sample was heated to >50°C some fluorescence signal was obtained for both dyes (Fig. 4 A). After setting the collection mode of the fluorimeter to time-dependent measurements at 60°C we found that at the very beginning both the ThT and ANS signals rose rapidly, reaching a clear maximum after 100 min with ThT and 400 min for ANS, and then decreasing. It might be surmised that the main reason for this decrease was the increase in light-scattering that parallels the greater viscosity

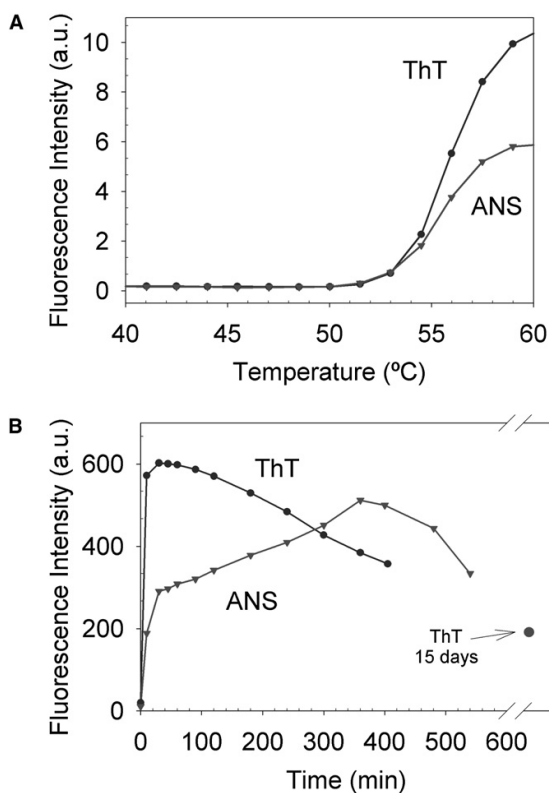


FIGURE 4 Growth kinetics of ThT and ANS fluorescence emission. A $8 \text{ mg} \times \text{mL}^{-1}$ PDZ3 sample was incubated in the presence of $12 \mu\text{M}$ of ThT or $20 \mu\text{M}$ ANS at pH 7.5. (A) Temperature dependence of fluorescence emission from both dyes. (B) Time dependence of the fluorescence signal versus incubation time at 60°C is plotted.

and turbidity of samples brought about by the insolubility of the supramacromolecular structures organized within this time (21–23). The protein samples became turbid and viscous and when the solutions were cooled to room temperature a hyper-chromic effect appeared as a consequence of the disaggregation of the assemblies into more soluble and lower molecular weight species (Fig. S5).

An interesting feature that emerges from Fig. 4 is the absence of the typical sigmoidal shape of the ThT and ANS curves obtained for the majority of examples studied previously (20,22,24). DLS measurements indicate that >50°C the monomers increase in size, to become practically 100% oligomeric at 60°C (Fig. S1). Accordingly, in Fig. 4 A it can be seen that the fluorescence signal of both probes begins to become temperature dependent at >50°C, which indicates clearly that the amyloid β -pleated structure identified by ThT fluorescence and the hydrophobic surface areas monitored by ANS fluorescence arose concurrently, as has been reported elsewhere for other examples (25). Thus, the formation of nuclei, held to be responsible for

the lag-phase of S-shaped curves, was complete before the sample's reaching the starting temperature. This would seem to indicate that under these experimental conditions nucleation is not a rate-limiting step for PDZ3 amyloid formation. Aggregation in these cases has been reported to proceed via a downhill process that does not require a highly organized or stable nucleus (20,26,27).

DISCUSSION

The PSD95-PDZ3 domain represents a typical medium-sized protein that unfolds via a multistate scheme

The thermodynamic parameters set out in Table 1 show that PDZ3 is a stable domain with an unfolding Gibbs energy of $40 \text{ kJ} \times \text{mol}^{-1}$, which is comparable to the stability of other proteins of a similar size, such as barnase (28). This value does not compare well with the previous stability data obtained for PTPBL-PDZ2 or for PDZ3 itself, which ranges from 14 to $30 \text{ kJ} \times \text{mol}^{-1}$ (2,12). It is worth bearing in mind, however, that these previous results were obtained from a two-state analysis of spectroscopic unfolding data.

The oligomeric equilibrium intermediate detected in our experiments can be considered as a novel species not reported previously in other PDZ folding studies. Thus, the folding mechanism of several PDZ domains, including PSD95-PDZ3 itself, has been studied by folding kinetics. It can be described as having two transition states separated by a high-energy obligatory (on-pathway) intermediate. Both transition states are located at roughly the same positions along the reaction coordinate, indicating a conserved folding mechanism for PDZ domains. Nevertheless, despite the obligatory character of the intermediate, it never accumulates under equilibrium (2,13,29). An interesting NMR study has also been published on native-state hydrogen exchange using a longer construct of PDZ3, which identifies an equilibrium intermediate after the rate-limiting folding step (12). The structural ensemble of the intermediate closely resembles that of the native PDZ3 fold, with the exception of the first residues of the N-terminal tail, the third α -helix and a further β -hairpin organized by the C-terminal residues. Nevertheless, the PDZ3 sequence analyzed in our work, corresponding to residues 302–402 of the PSD95 protein, does not contain the β -hairpin sequence deriving from the cloning vector used in the work of Feng et al. and in other previous structural analyses (12). In any case, the additional evidence about the low cooperative folding character, the essentially β -sheet content and the oligomeric nature of the intermediate detected by DSC using the short PDZ3 construct makes this equilibrium intermediate substantially different from the one with the longer PDZ3 sequence, which is essentially nativelike (12).

We can only find one previous example in the DSC literature of a protein of a similar size to PDZ3 that also displays a monomeric α - β topology and has been studied according to

the same equilibrium model, and that is the 14 kDa protein CheY (30). CheY and PDZ3 intermediates, as detected by DSC, share some common features, such as their tendency to associate themselves and their low unfolding cooperativity. Nevertheless, the organization of supramacromolecular structures in CheY has not yet been described. In any case, both PDZ3 and CheY firmly support the idea that medium-sized monomeric proteins ($>10 \text{ kDa}$) usually unfold in a more complex way than smaller ones and present at least one equilibrium intermediate.

Thermodynamic and conformational aspects of fibril organization on PSD95-PDZ3 unfolding

The formation of fibrils in PDZ domains has been reported very recently in the PTPBL-PDZ2 domain, as a result of a combination of mechanical stirring and high-ionic-strength solvents under native conditions at neutral pH. The combination of both forces, mechanical and electrostatic, leads to the appearance of amyloid fibrils in $<24 \text{ h}$ (6). Nevertheless, despite the native conditions of the solution, the fibrils show the typical CD and infrared spectra of pure β -sheet structures, as in the case of the PDZ3 unfolding intermediate.

Our studies with PDZ3 show some interesting features concerning the physical-chemical nature of processes involved in fibril organization. First, seeding may well occur through a β -sheet oligomeric nucleus, which suggests strongly a stepwise model for the organization of such large structures, as has been reported elsewhere for other examples (22,31,32). This nucleation phase is completed before temperature equilibration of PDZ3 samples, which results in a noncooperative (nonsigmoid) behavior of fibril growth (Fig. 4), in which the ensemble of critical oligomers adds to the different supramacromolecular assemblies in a downhill process (26,27,31). Second, the fibrils disaggregate after 1:1 dilution of the sample with a buffer solution at pH 7.5 and cooling to room temperature, and reform on heating again to 60°C . This evidence points to the reversible nature of processes involved in fibril formation despite their long relaxation times because they can disintegrate and rejoin spontaneously in a reversible manner. There are only a couple of other amyloid-forming proteins in which this phenomenon has been described: acylphosphatase (16) and apolipoprotein C-II (33). Therefore, PDZ3 fibrils offer an interesting and rare opportunity to study the disaggregation of these structural ensembles, which may help in the detection of new key species to explain the molecular nature of fibril formation and in the further development of treatments for diseases associated with protein aggregation.

From an energetic point of view, our work has also provided an unusual example of the thermodynamics of fibril formation. We have observed that the growth of oligomers does not seem to be accompanied by any conformational change in the intermediate state because the last DSC transition does not include any appreciable heat effects other than

that caused by the disaggregation of oligomers (Fig. 3 B). Accordingly, the far-UV CD spectra acquired as a function of time at 70°C do not show any apparent changes in shape (Fig. S4). A similar feature has been found in the previous calorimetric characterization of amyloid fibril formation in β_2 -microglobulin (23) and in SH3 domains (24), which also display other similarities with PDZ3 that are worth mentioning. Thus, the enthalpy change of the dissociation of PDZ3 oligomers ($-130 \pm 20 \text{ kJ} \times \text{mol}^{-1}$) coincides fairly closely with that corresponding to β_2 -microglobulin (-125 kJ/mol^{-1}) and SH3 ($-100 \text{ kJ} \times \text{mol}^{-1}$) fibrils. These values are also in the range of those found for the dissociation of dimeric-equilibrium intermediates in CheY protein (30), which shares a common DSC-thermal unfolding behavior with PDZ3, as has been explained above. The quite small specific enthalpy value estimated for PDZ3 ($11.8 \text{ J} \times \text{g}^{-1}$) as well as for the other three examples indicates a relatively low density of side-chain packing interactions inside the fibrils. Whatever the case, all these enthalpic similarities argue in favor of entropic factors as being the driving force and point toward a similar energetic pathway for the organization of β -sheet stacking, as has also been suggested by the kinetic experiments (23,24).

In addition, the Gibbs energy at 70°C (the temperature at which these associated species populate according to Fig. 3) was $25 \pm 5 \text{ kJ} \times \text{mol}^{-1}$, whereas it was almost negligible at 25°C, where the stability of the native state has been estimated to be $40 \pm 6 \text{ kJ} \times \text{mol}^{-1}$. The greater stability of the native state than that of amyloid fibrils in PDZ3 may explain the reversibility observed between them because the native state is much preferred to the associated structures at room temperature. Accordingly, the contrary effect has been reported for the irreversible fibrils in β_2 -microglobulin, where a higher stability value has been found for the fibrillar assemblies than for the N-state at any temperature (23). The stability of PDZ3 assemblies is clearly lower than that corresponding to β_2 -microglobulin ones ($44 \text{ kJ} \times \text{mol}^{-1}$), which may indicate possible structural differences between them of entropic origin.

The PDZ3 unfolding scheme as a paradigm for the biophysical study of interconversion between supramacromolecular assemblies

The capacity to form amyloid fibrils is regarded today as a general property of all polypeptide chains. Less evidence exists concerning the formation of annular and spherical arrangements, but some recent and interesting work carried out by Quist et al. (8) has also shown that a set of unrelated amyloidogenic proteins and peptides display quite similar annular formations when assayed in the presence of reconstructed membranes, and are also able to act as ion transporters through them, i.e., as ion channels. This evidence also points clearly to a parallel ubiquitous nature for the formation of these species. The diameter of the channels

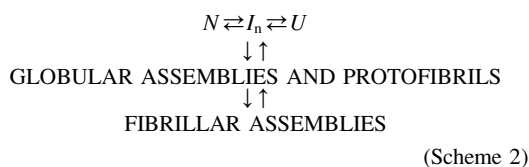
ranges from 8 to 12 nm, similar to those reported here for the PDZ3 rodlike species of 12 nm (Fig. S1). A comparison of the different PDZ3 particulates also shows a relative homogeneity in their size (Fig. 2), which suggests that the organization of these arrangements is well-defined (8). Correspondingly, spherical aggregates ranging from 5 nm to 200 μm have been described, also pointing to their being a general property of polypeptide chains (34) and behaving as pore-forming species in lipid environments (35).

The relationship between annular/spherical species and amyloid fibrils has not been clearly determined. The question of whether there is a common pathway for the organization of these supramacromolecular structures or whether they are products of alternative misfolding pathways has remained unresolved until now, given the limitations of experimental systems (10). The majority of examples described shows a clear tendency to associate themselves in solution into small oligomeric species (generally from dimers to hexamers), although they can lead to disklike assemblies only in the presence of membranes (8). Alternatively, spheres usually organize under conditions close to the isoelectric point of the protein (34) and/or by adding organic solvents (35).

Interestingly, the PDZ3 globular arrangements can organize themselves in a membrane-free context, coexisting with the fibrils in the same buffered solution at neutral pH, lacking cosolvents or organic reagents. This rare phenomenon, together with the persistence of these globular assemblies in solution for at least 2 weeks, supports the view that they may be on-pathway, as has been proposed for human serum albumin (27), Sup35 (36), and Ure2p (32) proteins. In fact, as the TEM images show (Fig. 2), the fibrils grow at the expense of the globular arrangements, which disappear from the solution during the last stages of the incubation experiments. On the other hand, the concurrence of protofibrils during the early phases of the experiments, together with the fact that we could detect no globular assemblies in the $8 \text{ mg} \times \text{mL}^{-1}$ PDZ3 samples incubated at 37°C, only protofibrils and a few thin fibrils, does not allow us to arrive at any definitive conclusions about these matters (the TEM images obtained after 1 month incubation at 37°C were identical to those shown in Fig. 2 B).

The time experiments followed by DLS and ThT/ANS fluorescence provide additional information about these mechanistic matters because the absence of a nucleation phase (lag-phase) $>50^\circ\text{C}$ and the coincidence of the beginnings of the time-courses monitored by the different probes implies that a certain quantity of oligomeric ensembles displaying hydrophobic patches and β -sheet amyloidlike structures form concurrently after a common nucleation event. Something similar has been observed in the α -synuclein amyloidogenic protein. Consequently, we can also conclude that, if separate pathways exist (off-pathway mechanism), the branchpoint must be located after nucleation, i.e., after the I_n species in the PDZ3 folding scheme (37).

In summary, the experimental information can be summarized in the following scheme, intended to explain the unfolding behavior of the PSD95-PDZ3 domain in 50 mM K-phosphate at pH 7.5:



where the double arrows indicate the reversible nature of the different processes and not necessarily any true equilibrium situation, at least for the ones pertaining to the misfolding pathway. This folding behavior displayed by PDZ3 opens the possibility of entering more deeply into structural studies with the different supramacromolecular assemblies. CD experiments, together with ThT and ANS fluorescence emission, show that these structures have a high β -sheet content organized into amyloidlike structures that also expose a relatively high amount of hydrophobic surface to the solvent. In addition, the I_n -state cannot be considered as native-like because it appears after a highly cooperative unfolding process from the native state, as shown by DSC, and neither is it able to recognize the high-affinity peptide KKETAV because its presence in the PDZ3-incubated solutions does not alter the temporal sequence of events given in Fig. 2.

Nevertheless, some other questions remain to be solved. It remains unclear, for example, how such oligomers can manage to organize themselves into defined ion-channel-like structures (10). Recent molecular dynamics simulations (38) have shown that in a membrane-free context double-layered annular structures exhibit high structural stability due to strong hydrophobic interactions and geometrical constraints induced by the closed-ring shape. These studies also reveal that annular structures can be postulated as an intermediate step in the aggregation pathway, leading to fibrils at later stages, which completely agrees with our PDZ3 experimental results. The quaternary structures contain two layers of β -sheet molecules stacked either parallel or antiparallel to the fibril axis, giving rise to disklike arrangements of 8–12 nm in diameter, with inner cavity pores of around 2 nm, which agrees with previous evidence (8) and our own experimental results.

The most recent studies indicate that fibrils could simply be a storage mechanism and that these smaller globular arrangements are sufficient to induce cell degradation by allowing ion exchange across the plasma membrane and thus disrupting homeostasis. This phenomenon has given rise to promising strategies for the treatment of amyloid-derived diseases by the selective blocking of such ion channels (8–10). Therefore, the unfolding framework of PDZ3 described by Scheme 2 is extremely informative when it comes to obtaining experimental insight into the still poorly understood mechanism of amyloid formation and its resultant toxicity.

The reversible coalescence under roughly physiological conditions of different kinds of well arranged supramacromolecular organizations during PDZ3 unfolding, together with the membrane-free organization of spherical/annular particulates, makes this protein system especially interesting for further studies into aggregation processes coupled to folding.

SUPPORTING MATERIAL

Methods and five figures are available at [http://www.biophysj.org/biophysj/supplemental/S0006-3495\(10\)00434-0](http://www.biophysj.org/biophysj/supplemental/S0006-3495(10)00434-0).

We thank Dr. J. Ruiz-Sanz and Dr. L. Bejarano for their helpful discussions and Dr. J. Trout for revising and correcting our English text. J.M.C. is a Spanish Government PhD fellow.

This work was supported by the Spanish Ministry of Science and Education (BIO2006-15517-C02-01/02, BIO2009-13261-C02-01/02), the Andalusian Regional Government (grant P09-CVI-5063 to A.C.A.), and the University of Granada (contract to E.S.C.).

REFERENCES

- Feng, W., and M. Zhang. 2009. Organization and dynamics of PDZ-domain-related supramodules in the postsynaptic density. *Nat. Rev. Neurosci.* 10:87–99.
- Jemth, P., and S. Gianni. 2007. PDZ domains: folding and binding. *Biochemistry*. 46:8701–8708.
- Lockless, S. W., and R. Ranganathan. 1999. Evolutionarily conserved pathways of energetic connectivity in protein families. *Science*. 286:295–299.
- Gianni, S., T. Walma, ..., G. W. Vuister. 2006. Demonstration of long-range interactions in a PDZ domain by NMR, kinetics, and protein engineering. *Structure*. 14:1801–1809.
- Petit, C. M., J. Zhang, ..., A. L. Lee. 2009. Hidden dynamic allostery in a PDZ domain. *Proc. Natl. Acad. Sci. USA*. 106:18249–18254.
- Sicorello, A., S. Torrasa, ..., F. Chiti. 2009. Agitation and high ionic strength induce amyloidogenesis of a folded PDZ domain in native conditions. *Biophys. J.* 96:2289–2298.
- Bucciantini, M., E. Giannoni, ..., M. Stefani. 2002. Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature*. 416:507–511.
- Quist, A., I. Doudevski, ..., R. Lal. 2005. Amyloid ion channels: a common structural link for protein-misfolding disease. *Proc. Natl. Acad. Sci. USA*. 102:10427–10432.
- Shirwany, N. A., D. Payette, ..., Q. Guo. 2007. The amyloid beta ion channel hypothesis of Alzheimer's disease. *Neuropsychiatr. Dis. Treat.* 3:597–612.
- Jang, H., J. Zheng, ..., R. Nussinov. 2008. New structures help the modeling of toxic amyloidbeta ion channels. *Trends Biochem. Sci.* 33:91–100.
- Gianni, S., N. Calosci, ..., C. Travaglini-Allocatelli. 2005. Kinetic folding mechanism of PDZ2 from PTP-BL. *Protein Eng. Des. Sel.* 18:389–395.
- Feng, H., N. D. Vu, and Y. Bai. 2005. Detection of a hidden folding intermediate of the third domain of PDZ. *J. Mol. Biol.* 346:345–353.
- Chi, C. N., S. Gianni, ..., P. Jemth. 2007. A conserved folding mechanism for PDZ domains. *FEBS Lett.* 581:1109–1113.
- Filimonov, V. V., and V. V. Rogov. 1996. Reversible association of the equilibrium unfolding intermediate of lambda Cro repressor. *J. Mol. Biol.* 255:767–777.

15. Ruiz-Sanz, J., V. V. Filimonov, ..., P. L. Mateo. 2004. Thermodynamic analysis of the unfolding and stability of the dimeric DNA-binding protein HU from the hyperthermophilic eubacterium *Thermotoga maritima* and its E34D mutant. *Eur. J. Biochem.* 271:1497–1507.
16. Calamai, M., C. Canale, ..., C. M. Dobson. 2005. Reversal of protein aggregation provides evidence for multiple aggregated states. *J. Mol. Biol.* 346:603–616.
17. Güthe, S., L. Kapinos, ..., T. Kiefhaber. 2004. Very fast folding and association of a trimerization domain from bacteriophage T4 fibritin. *J. Mol. Biol.* 337:905–915.
18. Fernandez-Escamilla, A. M., F. Rousseau, ..., L. Serrano. 2004. Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nat. Biotechnol.* 22:1302–1306.
19. Cámara-Artigas, A., J. Murciano-Calles, ..., J. C. Martínez. 2010. Novel conformational aspects of the third PDZ domain of the neuronal post-synaptic density-95 protein revealed from two 1.4Å X-ray structures. *J. Struct. Biol.* 170:565–569.
20. Smith, A. M., T. R. Jahn, ..., S. E. Radford. 2006. Direct observation of oligomeric species formed in the early stages of amyloid fibril formation using electrospray ionization mass spectrometry. *J. Mol. Biol.* 364:9–19.
21. Hawe, A., M. Sutter, and W. Jiskoot. 2008. Extrinsic fluorescent dyes as tools for protein characterization. *Pharm. Res.* 25:1487–1499.
22. Kad, N. M., S. L. Myers, ..., N. H. Thomson. 2003. Hierarchical assembly of beta2-microglobulin amyloid in vitro revealed by atomic force microscopy. *J. Mol. Biol.* 330:785–797.
23. Kardos, J., K. Yamamoto, ..., Y. Goto. 2004. Direct measurement of the thermodynamic parameters of amyloid formation by isothermal titration calorimetry. *J. Biol. Chem.* 279:55308–55314.
24. Morel, B., S. Casares, and F. Conejero-Lara. 2006. A single mutation induces amyloid aggregation in the alpha-spectrin SH3 domain: analysis of the early stages of fibril formation. *J. Mol. Biol.* 356:453–468.
25. Apetri, M. M., N. C. Maiti, ..., V. E. Anderson. 2006. Secondary structure of alpha-synuclein oligomers: characterization by Raman and atomic force microscopy. *J. Mol. Biol.* 355:63–71.
26. Carrotta, R., M. Manno, ..., P. L. San Biagio. 2005. Protofibril formation of amyloid beta-protein at low pH via a non-cooperative elongation mechanism. *J. Biol. Chem.* 280:30001–30008.
27. Juárez, J., P. Taboada, and V. Mosquera. 2009. Existence of different structural intermediates on the fibrillation pathway of human serum albumin. *Biophys. J.* 96:2353–2370.
28. Martínez, J. C., M. el Harrous, ..., A. R. Fersht. 1994. A calorimetric study of the thermal stability of barnase and its interaction with 3' GMP. *Biochemistry.* 33:3919–3926.
29. Calosci, N., C. N. Chi, ..., P. Jemth. 2008. Comparison of successive transition states for folding reveals alternative early folding pathways of two homologous proteins. *Proc. Natl. Acad. Sci. USA.* 105:19241–19246.
30. Filimonov, V. V., J. Prieto, ..., L. Serrano. 1993. Thermodynamic analysis of the chemotactic protein from *Escherichia coli*, CheY. *Biochemistry.* 32:12906–12921.
31. Kodali, R., and R. Wetzel. 2007. Polymorphism in the intermediates and products of amyloid assembly. *Curr. Opin. Struct. Biol.* 17:48–57.
32. Jiang, Y., H. Li, ..., S. Perrett. 2004. Amyloid nucleation and hierarchical assembly of Ure2p fibrils. Role of asparagine/glutamine repeat and nonrepeat regions of the prion domains. *J. Biol. Chem.* 279:3361–3369.
33. Binger, K. J., C. L. Pham, ..., G. J. Howlett. 2008. Apolipoprotein C-II amyloid fibrils assemble via a reversible pathway that includes fibril breaking and rejoining. *J. Mol. Biol.* 376:1116–1129.
34. Krebs, M. R., G. L. Devlin, and A. M. Donald. 2007. Protein particulates: another generic form of protein aggregation? *Biophys. J.* 92:1336–1342.
35. Walsh, P., J. Yau, ..., S. Sharpe. 2009. Morphology and secondary structure of stable beta-oligomers formed by amyloid peptide PrP(106-126). *Biochemistry.* 48:5779–5781.
36. Serio, T. R., A. G. Cashikar, ..., S. L. Lindquist. 2000. Nucleated conformational conversion and the replication of conformational information by a prion determinant. *Science.* 289:1317–1321.
37. Hoyer, W., T. Antony, ..., V. Subramaniam. 2002. Dependence of alpha-synuclein aggregate morphology on solution conditions. *J. Mol. Biol.* 322:383–393.
38. Zheng, J., H. Jang, ..., R. Nussinov. 2008. Annular structures as intermediates in fibril formation of Alzheimer Abeta17-42. *J. Phys. Chem. B.* 112:6856–6865.

Biophysical Journal, Volume 99

Supporting Material

An oligomeric equilibrium intermediate as the precursory nucleus of globular and fibrillar supra-macromolecular assemblies in a PDZ domain

Javier Murciano-Calles, Eva S. Cobos, Pedro L. Mateo, Ana Camara-Artigas, and Jose C. Martinez

Supporting Material

An oligomeric equilibrium intermediate as the precursory nucleus of globular and fibrillar supra-macromolecular assemblies in a PDZ domain

Javier Murciano-Calles¹, Eva S. Cobos¹, Pedro L. Mateo¹, Ana Camara-Artigas² and Jose C. Martinez^{1,}.*

¹ Department of Physical Chemistry and Institute of Biotechnology, Faculty of Sciences, University of Granada, 18071 Granada, Spain.

² Department of Physical Chemistry, Biochemistry and Inorganic Chemistry, Faculty of Experimental Sciences, University of Almeria, 04510 Almeria, Spain.

MATERIALS AND METHODS

Protein samples

The original PDZ3 sequence cloned into pQE30 plasmid was given generously by Dr. Jose Reina (Institute of Biomedical Research, IRB, Barcelona, Spain) (1). Residues 302 to 402 (in PSD95 numbering) were sub-cloned into pBAT4 (EMBL Core Purification Facility, Heidelberg, Germany) and expressed in *Escherichia coli* BL21/DE3 cells (Novagen, Madison, WI). The cells were lysed and ultracentrifuged to obtain the soluble protein-rich fraction, after which PDZ3 was purified by 75 % ammonium sulfate precipitation and further Superdex-75 chromatography (GE-Healthcare, Fairfield, CT) in 50 mM Tris/HCl, 400 mM KCl at pH 7.5. A ratio of 15 mg of highly pure PDZ3 per liter of LB culture was obtained. The PDZ3 sequence from pQE30 plasmid was also purified using the same protocol.

The extinction coefficient was $2985 \text{ cm}^{-1} \cdot \text{M}^{-1}$, determined as described elsewhere (2). A molecular mass of 11,004 Da was obtained by MALDI-TOF experiments carried out at the Scientific Instrumentation Services (CIC) of the University of Granada. This analysis revealed that the first Met residue had been naturally cleaved, as described elsewhere in other examples (3,4). This point was confirmed by N-terminal sequencing analysis at the CIB instrumentation services of the CSIC, Madrid, Spain. Experimental samples were always prepared by extensive dialysis against 50 mM K-phosphate buffer at pH 7.5 and 4 °C.

General description of the experimental procedures

Differential scanning calorimetry (DSC) experiments were carried out indistinctly in two VP-DSC instruments from Microcal INC (Northampton, MA), as described previously (5). Dynamic light scattering (DLS) and circular dichroism (CD) experimental details can also be obtained from previous references, as well as the two-state analysis conducted in the CD-thermal denaturation experiments (6-8). The molecular weight of oligomers was calculated from their respective hydrodynamic radii by using the “globular protein” model included in the software of the DLS instrument (DynaPro, DYNAMICS V6, Wyatt Technology Corporation, Santa Barbara, CA). Size-exclusion chromatography (SEC) measurements were carried out using a Superdex 75 HR10/30 column (GE-Healthcare, Fairfield, CT). The column was calibrated by means of different markers of known molecular weight: blue dextran (to evaluate the void volume; 2000 kDa), cytochrome C (12.4 kDa), bovine serum albumin (66 kDa), myoglobin (17 kDa), aprotinin (6.5 kDa) and pepsin (35 kDa). Transmission electron microscopy (TEM) images were taken at the CIC services of the University of Granada as described in a previous work (7). Staining was done by using 1 % uranyl acetate solution. Growth kinetics by ThT and ANS were calculated according to the procedure previously described in (9). Briefly, protein aliquots in 50 mM K-phosphate (pH 7.5) were mixed at room temperature with ThT or ANS phosphate-buffered solutions to obtain final protein concentrations of 0.4 or 8 $\text{mg} \cdot \text{mL}^{-1}$. Dye concentrations were 12 μM for ThT or 25 μM ANS. Fluorimetry measurements were made as described elsewhere (7).

Evaluation of thermodynamic magnitudes by model analysis of DSC traces

The established three-state mechanism including an oligomeric intermediate



can be mathematically defined as it is in Eq. 1 or by any other alternative order of equilibrium, such as



Both models are completely equivalent from a thermodynamic point of view and give rise to absolutely equal fitting results, but refer to different equilibrium aspects. We used the scheme in Eq. 2 for curve fitting analysis to obtain the contributions of the unfolding (first equilibrium) and of the association-dissociation processes (second equilibrium) separately. This is particularly convenient for reaching a better understanding of the deconvolution analysis shown in this work, as well as obtaining the unfolding parameters of PDZ3 directly from the N-U equilibrium, including stability at 25 °C. A similar case was first analyzed by Filimonov & Rogov using a similar approach (10). In any case, we can define the equilibrium constants that characterize both processes as

$$K_{N-I_n} = \frac{[I_n]^{1/n}}{[N]} \quad K_{I_n-U} = \frac{[U]}{[I_n]^{1/n}} \quad K_{N-U} = \frac{[U]}{[N]} \quad K_{U-I_n} = \frac{[I_n]^{1/n}}{[U]} \quad (3)$$

in which it is obvious that $K_{N-U} = K_{N-I_n} \cdot K_{I_n-U}$ and $K_{U-I_n} = 1/K_{I_n-U}$, and therefore $\Delta H_{N-U} = \Delta H_{N-I_n} + \Delta H_{I_n-U}$ and $\Delta H_{U-I_n} = -\Delta H_{I_n-U}$. This means that the thermodynamic parameters of any arrangement of equilibria in the model can easily be obtained from these simple mathematical relationships.

According to the reaction scheme in Eq. 2, the populations or fractions of the different species can be defined as

$$f_N = \frac{[N]}{P} \quad f_U = \frac{[U]}{P} \quad f_{I_n} = \frac{n[I_n]}{P} \quad (4)$$

where P is the total protein concentration in the experiment and $f_N + f_U + f_{I_n} = 1$ at any temperature. By replacing the concentration values into Eq 3 we get

$$K_{N-U} = \frac{f_U}{f_N} \quad K_{U-I_n} = \frac{1 - f_N - f_U}{n \cdot f_U \cdot P^{n-1}} \quad (5)$$

and can combine both equations to solve the fraction of protein in the unfolded state. We thus have

$$\left(n \cdot K_{U-I_n}^n \cdot K_{N-U} \cdot P^{n-1} \right) f_U^n + (1 + K_{N-U}) f_U - K_{N-U} = 0 \quad (6)$$

A numerical approach to solve Eq. 6 could be Newton's method, which is designed to solve potential equations of the general formula $g(x_n) = 0$. Bearing in mind the fact that at any temperature the fraction of unfolded protein should be a value between 0 and 1, we can develop an iterative method where the initial estimate (x_n) progressively approaches the correct value according to

$$x_{n+1} = x_n - \frac{g(x_n)}{g'(x_n)} \quad (7)$$

Thus, x_{n+1} is the resulting value of the tangent line to the $g(x_n)$ function at x_n . We have checked that 10 calculations are enough to define precisely the value of f_U , using the value obtained from the immediately previous calculation as the input value.

This leaves the temperature dependencies of thermodynamic functions to be estimated. The heat-capacity functions of different equilibrium states have been defined by linear functions of temperature

$$C_{pN} = a + (b \cdot T) \quad C_{pU} = c + (d \cdot T) \quad C_{pIn} = e + (f \cdot T) \quad (8)$$

The enthalpy functions for both the equilibria in Eq. 2 were obtained by integrating Kirchoff's equations as

$$\Delta H_{N-U} = \Delta H_{N-U}(T_{N-U}) + [(c - a)(T - T_{N-U})] + \left[\frac{(d - b)(T^2 - T_{N-U}^2)}{2} \right] \quad (9)$$

$$\Delta H_{U-In} = \Delta H_{U-In}(T_{U-In}) + [(e - c)(T - T_{U-In})] + \left[\frac{(f - d)(T^2 - T_{U-In}^2)}{2} \right] \quad (10)$$

where T_{N-U} and T_{U-In} are the midpoint temperatures of the unfolding and association-dissociation processes respectively (corresponding to f_U and f_{In} values of 0.5), whereas $\Delta H_{N-U}(T_{N-U})$ and $\Delta H_{U-In}(T_{U-In})$ are the corresponding enthalpy values. In a similar way, the entropy functions can be calculated as

$$\Delta S_{N-U} = \Delta S_{N-U}(T_{N-U}) + \left[(c - a) \ln \left(\frac{T}{T_{N-U}} \right) \right] + [(d - b)(T - T_{N-U})] \quad (11)$$

$$\Delta S_{U-In} = \Delta S_{U-In}(T_{U-In}) + \left[(e - c) \ln \left(\frac{T}{T_{U-In}} \right) \right] + [(f - d)(T - T_{U-In})] \quad (12)$$

where

$$\Delta S_{N-U}(T_{N-U}) = \frac{\Delta H_{N-U}(T_{N-U})}{T_{N-U}}; \quad \Delta S_{U-In}(T_{U-In}) = \frac{\Delta H_{U-In}(T_{U-In})}{T_{U-In}} + R \cdot \ln \left[\frac{2 \cdot (1/2)^{1/n} \cdot P_{ref}^{1/n}}{n^{1/n} \cdot P_{ref}} \right] \quad (13)$$

The second term of the association-dissociation entropy function is responsible for the protein concentration dependencies shown in Figure 1, giving rise to a unique set of thermodynamic parameters compatible with all thermal transitions at different protein concentrations. Thus the parameters shown in Table 1 correspond to a hypothetical experiment made at P_{ref} concentration. This value has been established as 100 μ M (11). Finally, the Gibbs energies can be easily calculated from the well known thermodynamic relations $\Delta G = \Delta H - T \cdot \Delta S$ and $\Delta G = -RT \cdot \ln K$.

The temperature function of apparent enthalpy describing the thermal unfolding of PDZ3 can be expressed as

$$H_{ap} = H_N + (\Delta H_{N-U} \cdot f_U) + [(\Delta H_{N-U} + \Delta H_{U-In}) \cdot f_{In}] \quad (14)$$

and the corresponding heat-capacity function, which will be fitted to the experimental traces, can easily be obtained by numerical methods, since

$$C_{p,ap} = \left(\frac{\partial H_{ap}}{\partial T} \right)_p \quad (15)$$

For the analysis of the PDZ3 traces shown in Figure 1 and obtained in 50 mM K-phosphate buffer at pH 7.5 and different protein concentrations we took it, as is assumed by the model, that a single set of thermodynamic magnitudes should be compatible with all the experimental traces, their differences being a consequence of the changes in equilibrium induced by variations in protein concentration. Thus we undertook an overall analysis to obtain the same heat-capacity $C_{pN}(T)$, $C_{pU}(T)$ and $C_{pIn}(T)$ functions (Eq. 8) for the five experiments, as well as the T_{U-In} and $\Delta H_{U-In}(T_{U-In})$ parameters related to the dissociation process. Both the T_{N-U} and $\Delta H_{N-U}(T_{N-U})$ parameters were fitted individually for each experiment in an attempt to assess the standard error associated to this value from the average. Table 1 shows the values, including the errors, for all the thermodynamic magnitudes. Calculations were made using Sigma Plot 2000 (Systat Software Inc.).

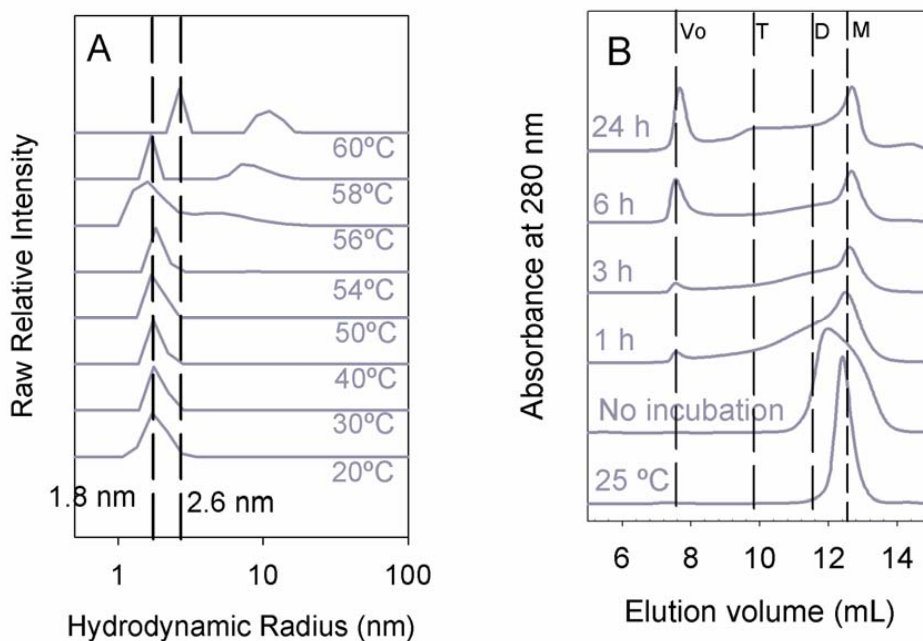
References

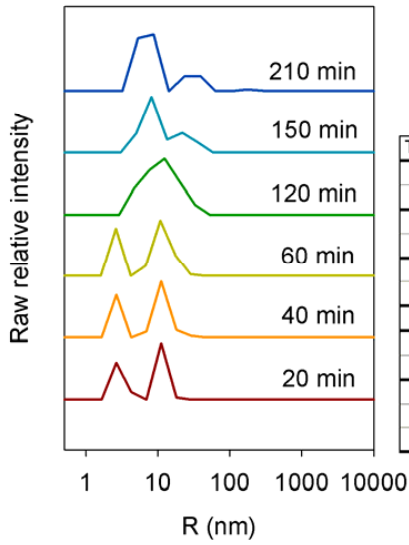
1. Reina, J., E. Lacroix, S. D. Hobson, G. Fernandez-Ballester, V. Rybin, M. S. Schwab, L. Serrano, and C. Gonzalez. 2002. Computer-aided design of a PDZ domain to recognize new target sequences. *Nat Struct Biol* 9:621-627.
2. Gill, S. C. and P. H. von Hippel. 1989. Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem* 182:319-326.
3. Hirel, P. H., M. J. Schmitter, P. Dessen, G. Fayat, and S. Blanquet. 1989. Extent of N-terminal methionine excision from *Escherichia coli* proteins is governed by the side-chain length of the penultimate amino acid. *Proc Natl Acad Sci U S A* 86:8247-8251.
4. Dalboge, H., S. Bayne, and J. Pedersen. 1990. In vivo processing of N-terminal methionine in *E. coli*. *FEBS Lett* 266:1-3.
5. Viguera, A. R., J. C. Martinez, V. V. Filimonov, P. L. Mateo, and L. Serrano. 1994. Thermodynamic and kinetic analysis of the SH3 domain of spectrin shows a two-state folding transition. *Biochemistry* 33:2142-2150.
6. Cobos, E. S., M. T. Pisabarro, M. C. Vega, E. Lacroix, L. Serrano, J. Ruiz-Sanz, and J. C. Martinez. 2004. A miniprotein scaffold used to assemble the polyproline II binding epitope recognized by SH3 domains. *J Mol Biol* 342:355-365.
7. Morel, B., S. Casares, and F. Conejero-Lara. 2006. A single mutation induces amyloid aggregation in the alpha-spectrin SH3 domain: analysis of the early stages of fibril formation. *J Mol Biol* 356:453-468.
8. Cobos, E. S., M. Iglesias-Bexiga, J. Ruiz-Sanz, P. L. Mateo, I. Luque, and J. C. Martinez. 2009. Thermodynamic characterization of the folding equilibrium of the human Nedd4-WW4 domain: at the frontiers of cooperative folding. *Biochemistry* 48:8712-8720.
9. Kad, N. M., S. L. Myers, D. P. Smith, D. A. Smith, S. E. Radford, and N. H. Thomson. 2003. Hierarchical assembly of beta2-microglobulin amyloid in vitro revealed by atomic force microscopy. *J Mol Biol* 330:785-797.
10. Filimonov, V. V. and V. V. Rogov. 1996. Reversible association of the equilibrium unfolding intermediate of lambda Cro repressor. *J Mol Biol* 255:767-777.
11. Ruiz-Sanz, J., V. V. Filimonov, E. Christodoulou, C. E. Vorgias, and P. L. Mateo. 2004. Thermodynamic analysis of the unfolding and stability of the dimeric DNA-binding protein HU from the hyperthermophilic eubacterium *Thermotoga maritima* and its E34D mutant. *Eur J Biochem* 271:1497-1507.

Figure S1. The full set of DLS and SEC experiments carried out at pH 7.5 with the PDZ3 domain of PSD95.

This page: Panel A. Protein samples at $8 \text{ mg}\cdot\text{mL}^{-1}$ were heated from 20 to $60 \text{ }^\circ\text{C}$ at a scan rate of $1 \text{ deg}\cdot\text{min}^{-1}$. Experimental conditions were 50 mM K-phosphate, pH 7.5 . **Panel B.** PDZ3 samples at $8 \text{ mg}\cdot\text{mL}^{-1}$ and pH 7.5 were incubated at $60 \text{ }^\circ\text{C}$ for different periods of time (shown in the panel) and then loaded into a Superdex 75 HR $10/30$ column also equilibrated at $60 \text{ }^\circ\text{C}$. The lowest run corresponded to a $25 \text{ }^\circ\text{C}$ incubated sample. Vertical lines represent the elution volume of hypothetical monomeric (M), dimeric (D) and trimeric (T) PDZ3 species, V_0 being the exclusion volume of the column.

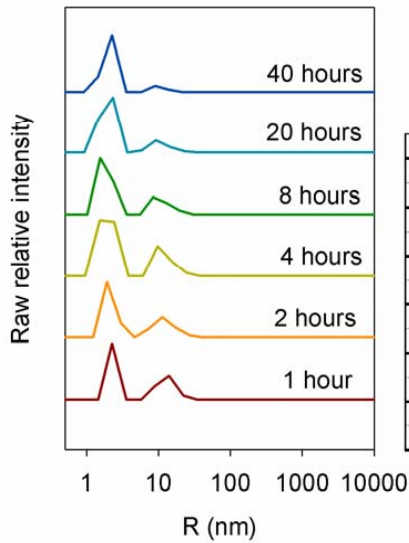
Next page: A protein solution at a concentration of $7.9 \text{ mg}\cdot\text{mL}^{-1}$ was heated initially from 20 to $60 \text{ }^\circ\text{C}$ (the maximum operating temperature of the instrument), where the I_n species should be most populated according to DSC analysis (Fig. 3, panel A). It was then kept at $60 \text{ }^\circ\text{C}$, at which point the mass evolution of the species as a function of incubation time was recorded (upper panel). Finally, the protein sample was cooled down inside the instrument to $20 \text{ }^\circ\text{C}$. A summary of the spectra obtained at different times after reaching this temperature is shown in the lower panel. The tables on the right summarize the mass distribution of the respective curves shown in the accompanying figures.





60 °C

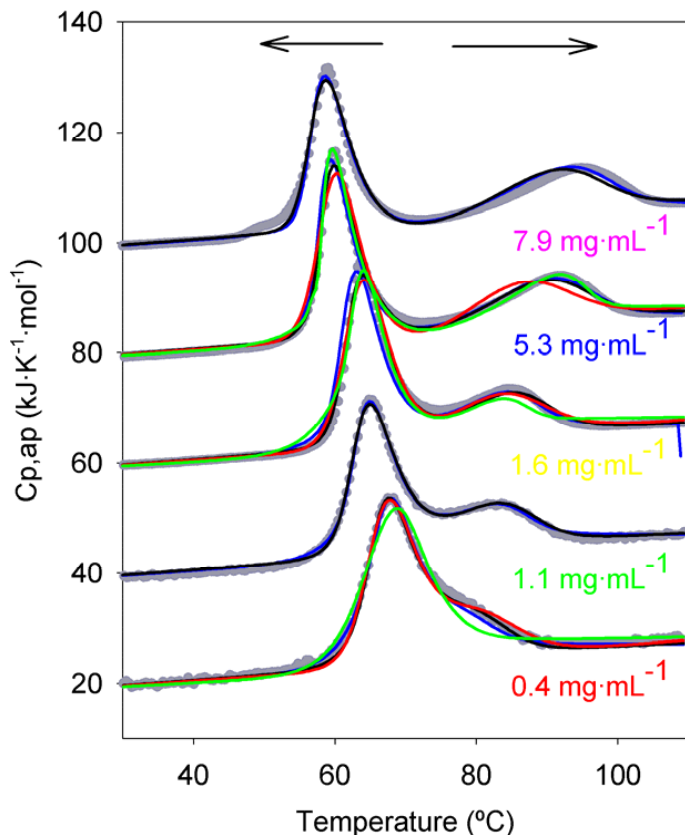
Time (min)	R (nm)	% Mass	MW (kDa)
20	2,6	43	32
	11,4	57	1004
40	2,7	37	33
	12	63	1133
60	2,6	35	32
	12,4	65	1219
120	12,7	100	1280
150	8,3	75	473
	26,9	25	7486
210	7	77	323
	31,5	21	10612
	199,4	2	816146



20 °C

Time (hours)	R (nm)	% Mass	MW (kDa)
1	2,2	57	23
	13,2	43	1408
2	2,1	63	20
	12,7	37	1279
4	1,9	69	14
	12,5	31	1253
8	1,8	72	12
	11,6	28	1054
20	1,9	81	14
	10,9	19	900
40	1,8	89	12
	8,9	11	560

Figure S2. DSC thermal unfolding curves of PDZ3 at pH 7.5 analyzed according to model $nN \rightleftharpoons nU \rightleftharpoons I_n$ taking into account different stoichiometries. The experimental traces (gray symbols) have been displaced along the Y-axis to give a better perspective. The respective fittings for $n=2$ (red), $n=3$ (black), $n=4$ (blue) and $n=6$ (green) are represented by solid lines. The protein concentrations are labeled in the figure. In the extreme cases of $n=2$ and $n=6$ we could only fit three of the five curves, which supports the clear low quality of these fittings compared to the other.



The thermodynamic parameters of the thermal unfolding of the PSD95-PDZ3 domain in 50 mM K-phosphate pH 7.5 obtained from the analysis of DSC experiments using different n values*

n	T_{N-U} (°C)	$\Delta H_{N-U}(T_{N-U})$ (kJ·mol ⁻¹)	T_{U-I_n} (°C)	$\Delta H_{U-I_n}(T_{U-I_n})$ (kJ·mol ⁻¹)	R_{Square}
2 [†]	70.8 ± 0.3	380 ± 10	82.6 ± 0.3	-150 ± 5	0.982
3	70.4 ± 0.1	335 ± 2	79.2 ± 0.1	-130 ± 2	0.991
4	69.7 ± 0.1	315 ± 1	86.2 ± 0.1	-120 ± 1	0.993
6 [†]	66.4 ± 0.2	280 ± 5	87.5 ± 0.3	-115 ± 4	0.977

* The error intervals have been taken to be the pure fitting errors. The thermodynamic magnitudes for the $U \rightleftharpoons I_n$ equilibrium were estimated for $P_{\text{ref}} = 100 \mu\text{M}$.

[†] The fittings at the extreme values $n=2$ and $n=6$ were done with only three of the five curves shown in Figure S2, since it was impossible to fit the whole set of curves due to the lower ($n=2$) or higher ($n=6$) relative distance predicted for both endotherms respectively.

Figure S3. The Tango analysis of β -aggregation-prone regions in PSD95-PDZ3. The input sequence was that corresponding to residues 302 to 402 from PSD95. The surface representation of the PDZ3 structure (PDB code: 3K82 and 3I4W) is represented in green, where the residues with some propensity to aggregate are shown in red. The three images are 90° views of the PDZ3 structure.

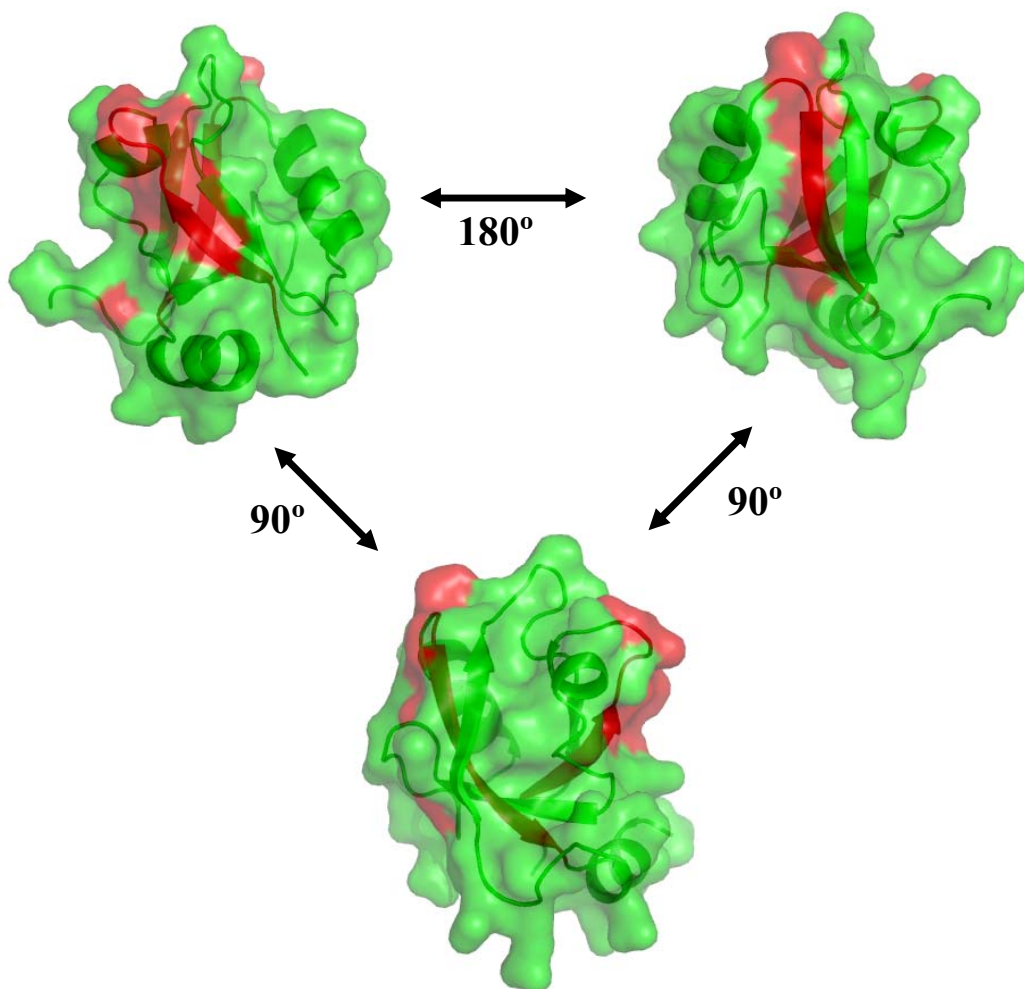


Figure S4. far-UV CD spectral analysis of the native and intermediate states of PSD95-PDZ3. The PDZ3 spectra in 50 mM K-phosphate buffer at pH 7.5 and 70 °C, (in which the intermediate state predominates) as a function of time passed after temperature equilibration: from 9 min to 4 h (colored lines), where a small increase in ellipticity is observed after this time (pink line). The protein concentration was $7.9 \text{ mg}\cdot\text{mL}^{-1}$, high enough to ensure that at 70 °C the intermediate was the main species in solution (Fig. 3, panel A). We also compare the far-UV CD-spectrum of PDZ3 under native conditions at pH 7.5 and 25 °C (black line).

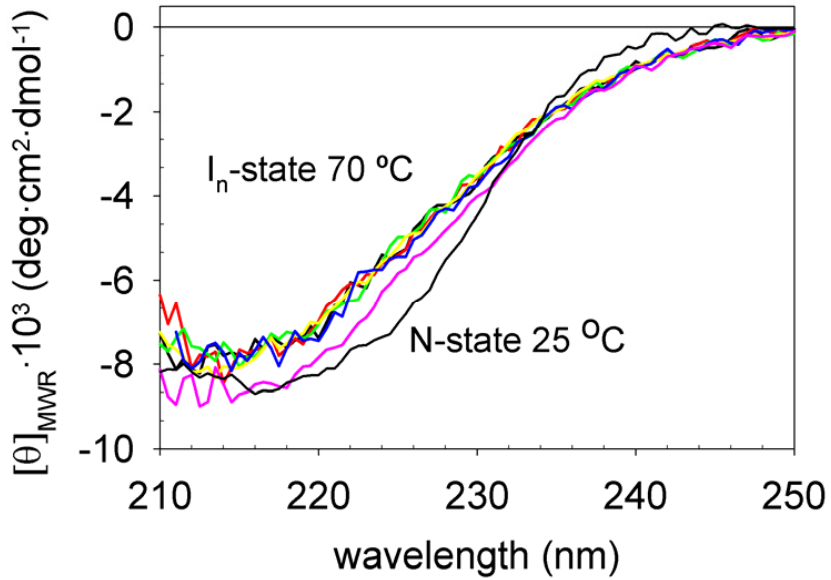
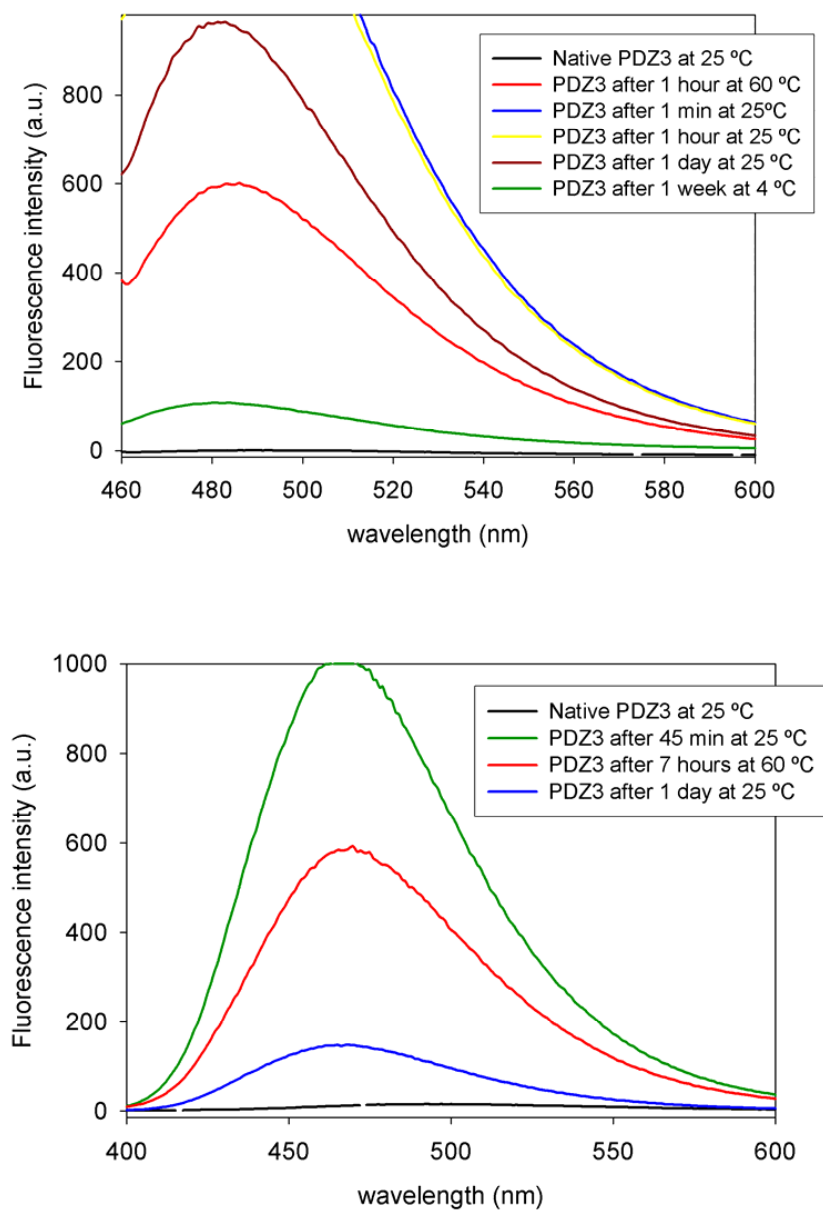


Figure S5. Fluorescence emission spectra of ThT (upper panel) and ANS (lower panel) dyes. The PDZ3 samples at $7.9 \text{ mg}\cdot\text{mL}^{-1}$ and pH 7.5 were cooled to room temperature after heating them to $60 \text{ }^\circ\text{C}$ for some few hours. Each spectrum represents a different stage of the cooled sample at $25 \text{ }^\circ\text{C}$.



Elsevier Editorial System(tm) for Biophysical Chemistry
Manuscript Draft

Manuscript Number:

Title: Electrostatic determinants for the pH-dependent behavior of the self-organization of supramacromolecular structures in a PDZ domain

Article Type: Research Paper

Keywords: folding intermediates; protein thermodynamics and stability; PDZ domains; oligomerization processes; annular/spherical aggregates; amyloid fibrils.

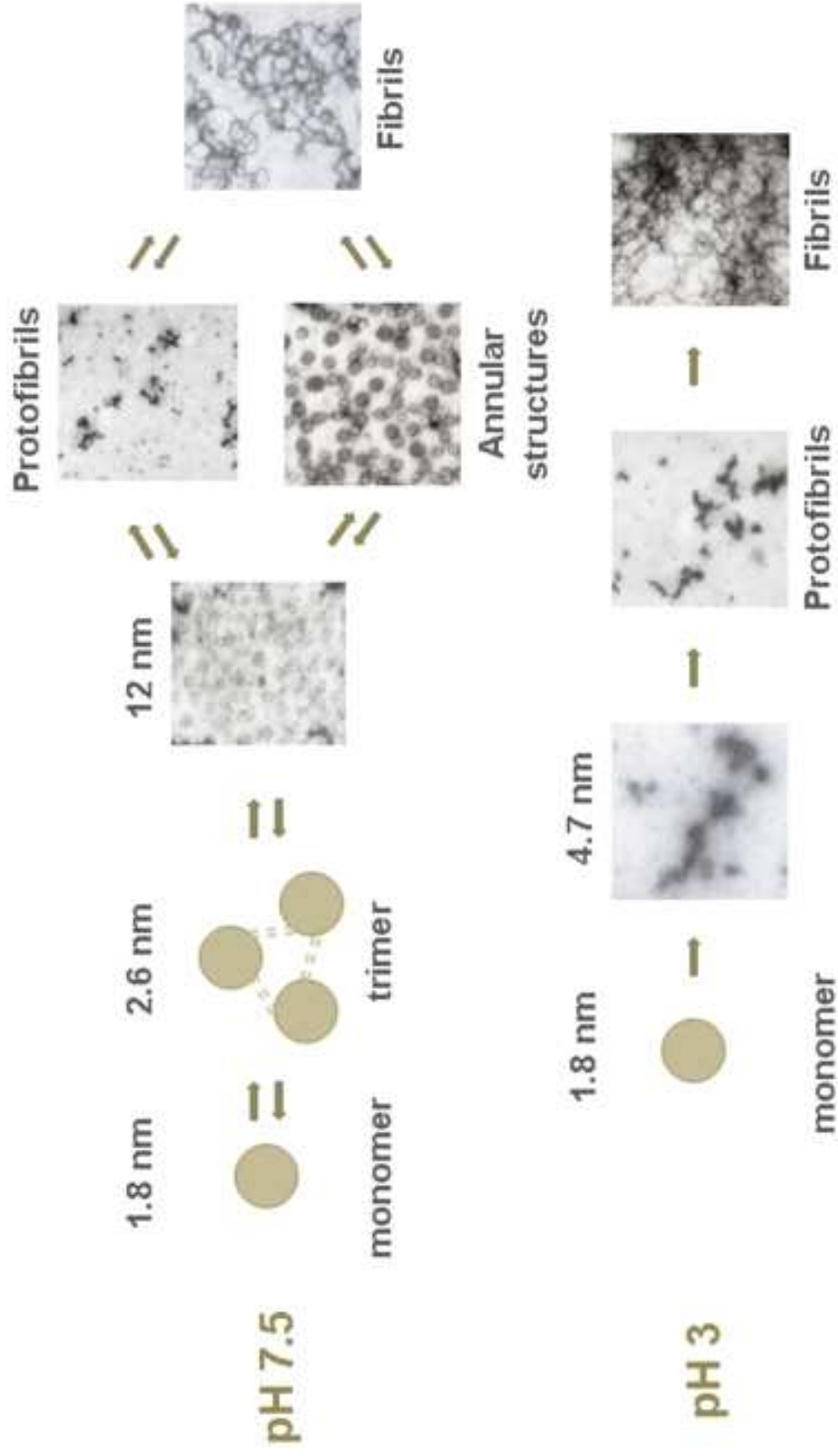
Corresponding Author: Dr. Jose C. Martínez,

Corresponding Author's Institution: Facultad de Ciencias, Universidad de Granada

First Author: Javier Murciano-Calles, Degree in Chemistry

Order of Authors: Javier Murciano-Calles, Degree in Chemistry; Eva S Cobos, Doctor in Chemistry; Pedro L Mateo, Professor in Chemistry; Ana Camara-Artigas, Doctor in Chemistry; Jose C. Martínez

Abstract: Equilibrium unfolding at neutral pH of the third PDZ domain of PSD95 is well described by the presence of a partly unfolded intermediate that presents association phenomena. After some days' incubation annular and fibrillar structures form from the oligomers. At pH values below 3, however, differential scanning calorimetry shows that PDZ3 seems to unfold under a two-state scheme. Kinetic measurements followed by dynamic light scattering, ThT and ANS fluorescence reveal that the misfolding pathway still exists despite the absence of any populated intermediates and shows an irreversible assembling of the supramacromolecular structures as well as an appreciable lag-phase, contrary to what is found in similar experiments at neutral pH. Moreover, as shown by transmission-electron-microscopy images, the annular structures seen at neutral pH completely disappear from incubated solutions. According to the structural information, this titration behavior appears to be the consequence of a conformational equilibrium that depends on the protonation of some Glu residues located at the C-terminal alpha-3 helix and at the hairpin formed by strands beta-2 and beta-3. Our calculations suggest that the enthalpic contribution of these interactions may well be as much as 40 kJ*mol⁻¹. The possible regulatory role of this equilibrium upon PDZ3 functionality and amyloid formation is briefly discussed.



Unfolding at $\text{pH} > 3.5$ of PSD95-PDZ3 reveals the presence of an equilibrium oligomeric intermediate state that self-associates into annular and fibrillar structures reversibly

At $\text{pH} < 3$ PDZ3 seems to unfold under a two-state scheme, although the misfolding pathway still exists

The supramacromolecular structures organize irreversibly and show a lag-phase at acidic pH

This titration behavior is due to a conformational equilibrium that depends on the protonation of some Glu residues

The enthalpic contribution of these interactions may well be as much as $40 \text{ kJ}\cdot\text{mol}^{-1}$.

Electrostatic determinants for the pH-dependent behavior of the self-organization of supramacromolecular structures in a PDZ domain

Javier Murciano-Calles[†], Eva S. Cobos[†], Pedro L Mateo[†], Ana Camara-Artigas[‡] and Jose C. Martinez^{†*}

[†] Department of Physical Chemistry and Institute of Biotechnology, Faculty of Sciences, University of Granada, 18071 Granada, Spain.

[‡] Department of Physical Chemistry, Biochemistry and Inorganic Chemistry, Faculty of Experimental Sciences, 04120 Almeria, Spain.

*Corresponding author:

Jose C. Martinez

Department of Physical Chemistry

Faculty of Sciences

University of Granada

18071- Granada, Spain

Tel: +34 958 242370

Fax: +34 958 272879

E-mail: jcmh@ugr.es

Abstract

Equilibrium unfolding at neutral pH of the third PDZ domain of PSD95 is well described by the presence of a partly unfolded intermediate that presents association phenomena. After some days' incubation annular and fibrillar structures form from the oligomers. At pH values below 3, however, differential scanning calorimetry shows that PDZ3 seems to unfold under a two-state scheme. Kinetic measurements followed by dynamic light scattering, ThT and ANS fluorescence reveal that the misfolding pathway still exists despite the absence of any populated intermediates and shows an irreversible assembling of the supramacromolecular structures as well as an appreciable lag-phase, contrary to what is found in similar experiments at neutral pH. Moreover, as shown by transmission-electron-microscopy images, the annular structures seen at neutral pH completely disappear from incubated solutions. According to the structural information, this titration behavior appears to be the consequence of a conformational equilibrium that depends on the protonation of some Glu residues located at the C-terminal $\alpha 3$ helix and at the hairpin formed by strands $\beta 2$ and $\beta 3$. Our calculations suggest that the enthalpic contribution of these interactions may well be as much as $40 \text{ kJ}\cdot\text{mol}^{-1}$. The possible regulatory role of this equilibrium upon PDZ3 functionality and amyloid formation is briefly discussed.

Keywords

folding intermediates; protein thermodynamics and stability; PDZ domains; oligomerization processes; annular/spherical aggregates; amyloid fibrils.

Abbreviations

DSC, differential scanning calorimetry; DLS, dynamic light scattering; CD, circular dichroism; TEM, transmission electron microscopy; PSD95, post-synaptic density-95 protein; PDZ, post-synaptic density-95 protein, disk-large tumor suppressor protein, zonula occludens-1; PSD95-PDZ3 or PDZ3, third PDZ domain of PSD95 protein; SH3, Src-homology 3 domain; MAGUK, membrane-associated guanylate kinase; ThT, thioflavine T; ANS, 1-anilino-naphthalene-8-sulfonic acid.

1. Introduction

Protein aggregation is considered today to be a generic property of polypeptides and is thus a considerable problem for living organisms, which must develop response strategies to avoid its harmful effects. These effects span from arthritis to serious neurodegenerative diseases. At the molecular and energetic level, the self-organization mechanism of proteins relies on many properties related to both the protein system (hydrophobicity, β -sheet propensities etc.) and the solvent conditions (pH, ionic strength, organic reagents etc.). These properties can modulate the misfolding pathway in ways ranging from downhill to nucleation-cooperative aggregation mechanisms.

In spite of extensive reports in the literature, our knowledge of the basic processes by which polypeptide chains can give rise to well ordered supramacromolecular structures is still poorly understood [1]. It has been suggested that the concepts underlying protein aggregation might be similar to those describing the organization of synthetic polymers, which may explain some features of the growth-kinetic processes [2]. A recent study has put forward a general unifying mechanism based on the formulation of master equations that describe the kinetics of fibrillar self-assembling as resulting from three basic processes: i) a nucleation-dependent polymerization reaction (lag-phase), usually slower than the ii) elongation of pre-existing nuclei, and finally iii) a secondary nucleation event deriving from the fragmentation of fibrils [3].

The almost universal irreversibility of amyloid-growth processes has precluded any deeper understanding of them than their purely kinetic features, and so little is known about the dynamics and nature of the ensemble of states, the structural and energetic aspects of the organization of their supramacromolecular assemblies or even of the inter-conversion between their various types. In a previous work [4] we explored the misfolding pathway of the third PDZ domain of the PSD95 neuronal protein (PDZ3), comprising residues 302 to 402 of PSD95. At neutral pH at 60-70 °C in plain water-buffered solutions the PDZ3 domain populates a trimeric β -sheet-rich intermediate state that leads to a stepwise and reversible formation of annular and protofibrillar structures, which remain in solution for various days and finally give rise to fibrils. Some authors have suggested that it is these annular and protofibrillar aggregates that are the toxic agents responsible for amyloid diseases, most probably through membrane interactions by pore formation [5,6].

To arrive at some insight into the molecular aspects of the misfolding pathway of PDZ3 we undertook DSC experiments under different pH conditions ranging from acidic to neutral. Surprisingly, the domain unfolded under a simple two-state scheme at pH values below 3, whereas the previously described three-state regime appeared above this value. Kinetic measurements followed by DLS, ThT and ANS fluorescence revealed that the misfolding pathway still existed despite the absence of any populated intermediates and showed an irreversible assembling of the supramacromolecular structures as well as an appreciable lag-phase at $\text{pH} \leq 3$, contrary to what was found in respective experiments at neutral pH. Moreover, as shown by TEM images, the annular structures seen at neutral pH completely disappeared from incubated solutions.

This titration behavior might be attributed to the protonation equilibria of Glu and/or Asp residues, the pKa values of which are the only ones within the range of pH 3-4. An analysis of two 1.4 Å X-ray structures of PDZ3 reveals that some Glu residues appear to be principally responsible for the attachment of the $\alpha 3$ helix to the whole PDZ3 structure through the β -hairpin formed by strands $\beta 2$ and $\beta 3$. In addition, we have already described the cyclation of the acidic residue D322 into a succinimide ring at the β -turn organizing this β -hairpin [7]. Thus, when these acidic residues titrate above pH 3 the proton release would generate some degree of electrostatic repulsion between their negative charges in the $\beta 2$ - $\beta 3$ hairpin as well as between the Glu side-chains and the π -electron clouds of some aromatic

rings of the C-terminal $\alpha 3$ helix, which might contribute to a weakening of the attachment. It has been proposed that this $\alpha 3$ helix may be a regulatory element of the binding properties of PDZ3 [7,8].

2. Materials and Methods

The PDZ3 sequence cloned into pQE30 plasmid was generously donated by Drs. Jose Reina (Institute of Biomedical Research, IRB, Barcelona, Spain) and Luis Serrano (EMBL-CRG, Barcelona, Spain). Residues 302 to 402 (in PSD95 numbering) were subcloned into pBAT4 (EMBL Core Purification Facility) and expressed in *E. coli* BL21/DE3 cells. PDZ3 can be obtained at a ratio of 15 mg per liter of LB culture and purified by Superdex-75 chromatography (GE-Pharmacia) in 50 mM K-phosphate, 150 mM KCl, at pH 7.5. The molecular mass was 11004 Da as determined by MALDI-TOF experiments carried out at the Centro de Instrumentacion Cientifica (CIC) services of the University of Granada. The extinction coefficient used was $2985 \text{ cm}^{-1} \cdot \text{M}^{-1}$, determined as described elsewhere [9].

DSC experiments were conducted in two VP-DSC instruments from Microcal INC., as described elsewhere [10]. PDZ3 experimental conditions were 50 mM buffer (either phosphate at pH 7.5, acetate at pH 4.0 or glycine/HCl at pH 2.0-3.5). DLS, CD and ThT and ANS fluorescence experimental details can be obtained from previous references [4,11]. TEM images of amyloid fibrils were made at the CIC services of the University of Granada. 1 % uranyl acetate was used for staining.

The two-state non-linear curve fittings of DSC traces at acidic pH values were made on the basis of the equations reported elsewhere [10]. The established three-state model including an oligomeric intermediate used to fit DSC traces above pH 3 ($nN = I_n = nU$) has also been described previously by us [4]. Linear functions for the heat-capacity functions $C_{pN}(T)$, $C_{pI}(T)$ and $C_{pU}(T)$ were always used. The unfolding (and dissociation in the case of the three-state model) temperatures, enthalpies and heat capacity functions for every experiment were obtained from the respective fitting analyses. The error intervals were taken to be three times the fitting standard errors (99 % confidence).

3. Results

3.1. PDZ3 unfolds under an apparent two-state scheme at acidic pH

As we showed in a previous work [4], the thermal unfolding of PDZ3 at neutral pH reveals the presence of an oligomeric equilibrium intermediate that populates maximally at around 60-70 °C. The DSC traces comprise two well separated unfolding transitions, which can be fully described by a three-state association-dissociation equilibrium model ($nN = I_n = nU$). We observed a similar behavior when we extended our experiments to buffered solutions at pH values >3.5 (Fig. 1). Nevertheless, at pH values ≤ 3 the biphasic behavior disappeared and showed just a single endotherm, which can be described by a two-state model ($N = U$; Fig. 1). The quality of the fittings is surprisingly good, as revealed by Figure 1 and by their R-square parameters (Table 1). In addition, the reversibility of the unfolding process increased dramatically from 60 % for the biphasic traces to 90 % for those below pH 3.

Nevertheless, this simple unfolding model does not seem to describe the whole unfolding behavior of PDZ3 at pH ≤ 3 since an increase in protein concentration above $5 \text{ mg} \cdot \text{mL}^{-1}$ produced a post-transition shoulder similar to that observed at pH 3.5 (Fig. 1). This feature should not appear in a pure two-state unfolding system, which has been described as being concentration independent [10,12]. Therefore, we can conclude that the intermediate state, although still being inherent to the unfolding pathway, is not populated at acidic pH values and low protein concentrations due to its energetic destabilization.

3.2. Evaluation of electrostatic contributions of acidic residues to the thermodynamic parameters of unfolding

The titration behavior observed in DSC experiments is certainly due to the protonation equilibria of Glu and/or Asp residues, the pKa values of which are the only ones within the range of pH 3-4 (see Discussion for further details). To evaluate the whole energetic contribution of the electrostatic balance of Glu and Asp residues we have represented the unfolding enthalpies as a function of their respective unfolding temperatures obtained under the various pH conditions (Fig. 2). The values deriving from two-state analysis ($\text{pH} \leq 3$) correlate well with a linear dependency. Furthermore, they coincide well with the hypothetical value of $54 \text{ J}\cdot\text{g}^{-1}$ at $110 \text{ }^\circ\text{C}$, which, as postulated by some authors [12-16], represents a convergence value of the enthalpy functions of small globular proteins. The slope of this regression provides a value for the unfolding heat-capacity change of $5.6 \text{ kJ}\cdot\text{K}^{-1}\cdot\text{mol}^{-1}$, which is comparable to the values obtained for other proteins of similar size [12,15].

The enthalpy values obtained from three-state analysis ($\text{pH} > 3$) do not, however, fit any such linear dependency (Fig. 2). Thus, considering that at these pH values the most relevant energetic feature appears to be that related to the titration of Glu and Asp residues, we can conclude that the roughly $40 \text{ kJ}\cdot\text{mol}^{-1}$ discrepancy between the values above pH 3 and the linear dependency of the enthalpy values under most acidic pH conditions may well be related to the heat contribution of some repulsive electrostatic forces developed upon the proton release of Glu/Asp residues above pH 3. Our estimation will only include the net contribution of such electrostatic balancing, since the far-UV CD spectra do not show any appreciable differences between pH 3 and pH 7.5 (Fig. S1), indicating a similar structural arrangement for PDZ3 under the solvent conditions studied.

3.3. A different mechanism of PDZ3 fibril organization at acidic pH

As we have shown in a previous work [4], at pH 7.5 and $60\text{-}70 \text{ }^\circ\text{C}$ the intermediate species of PDZ3 can develop a misfolding pathway that begins with the formation of trimers, which appear above $50 \text{ }^\circ\text{C}$. These trimers self-associated into 12 nm particles that grew within a few days into protofibrillar and annular assemblies and then into fibrillar aggregates. The sequence of equilibrium events is summarized in Figure 3.

DLS experiments carried out with PDZ3 at acidic pH revealed a stepwise misfolding reaction, differing in some important aspects from that described above. Thus, the heating from $25 \text{ }^\circ\text{C}$ to $60 \text{ }^\circ\text{C}$ of a $8 \text{ mg}\cdot\text{mL}^{-1}$ protein sample in 50 mM glycine/HCl buffer at pH 3 indicates that monomeric species remained in solution until the final temperature, coexisting with other particles of 4.7 nm , which appeared above $50 \text{ }^\circ\text{C}$ and comprised as much as 10% of the total mass (not shown). When the sample was incubated at $60 \text{ }^\circ\text{C}$ as a function of time the picture summarized in Figure 4 was obtained. Basically, both species remained within the solution after 4h although the mass percentage of the 4.7 nm species increased to 40% . It was after this time when monomers definitely disappeared from the protein solution and the 4.7 nm particles populated to a maximum. A further incubation of the samples for longer times revealed that the 4.7 nm particles increased in size. It should be noted that at pH 7.5 the monomers disappeared and fully trimerized before $60 \text{ }^\circ\text{C}$ was reached [4]. These trimers never appeared at pH 3 however (Fig. 3).

Long-term incubated samples of PDZ3 at $60 \text{ }^\circ\text{C}$ and pH 3 were analyzed by TEM (Fig. 5). Some particles of a few nm appeared after 4 hours' incubation, which may be the 4.7 nm particles seen by DLS. After 1 day only fibrillar structures appeared and these became longer, more populated and thicker with time. No annular structures were found at all in any of the different preparations. The fibrils had a similar curly shape to those seen at neutral pH, although they grew faster when compared to those at pH 7.5. Competition with annular structures within the misfolding pathway at neutral pH prevented the fibrils from maturing

due to the early formation of these structures, whereas at acidic pH fibrils grew exclusively at the expense of the precursory oligomers. In addition, the dilution and cooling down of the samples to 25 °C did not destroy the fibrils at pH 3, as it did at pH 7.5, indicating that the misfolding processes were essentially irreversible. DLS experiments also confirmed the irreversibility of misfolding, since the distribution of the particles remained the same when incubation at 60 °C was interrupted.

3.4. ANS and ThT fluorescence experiments reveal a lag-phase during fibril growth and confirm the irreversibility of misfolding at pH 3

Heating a sample of a 8 mg·mL⁻¹ PDZ3 inside the fluorimeter from 25 to 60 °C at pH 3 did not lead to fluorescence when ThT 12 µM was added, although there was a small amount in the presence of 20 µM ANS (Fig. S2). When both ThT and ANS samples were incubated at 60 °C, the fluorescence emission recorded as a function of time displayed a sigmoid kinetic profile with a lag-phase of around 100 min, where the fluorescence remained minimal, although a substantial increase in the signal occurred afterwards (Fig. 6). These curves displayed no sigmoidal shape at neutral pH, at which we observed a substantial and practically concomitant increase in the fluorescence of ThT and ANS samples at the beginning of their incubation at 60 °C, indicating an essentially non-cooperative growth [4]. The lag period of these curves is well accepted as being a nucleation stage during which non-fluorescent oligomeric species are formed prior to the fibrillar structures that will grow under a subsequent rapid, cooperative regime. This behavior is that which has most frequently been observed among the many amyloidogenic proteins studied to date [3,17-19].

The considerable difference in shape of the ThT and ANS profiles also argues in favor of a different mechanism for the misfolding pathway between neutral and acidic pH conditions for PDZ3. Thus, their increase in fluorescence after roughly 2 h indicates that the non-fluorescent species formed during this lag period (basically monomers and 4.7 nm particles) did not expose hydrophobic patches and did not show typical β -aggregation, as occurred at neutral pH with the respective trimeric and 12 nm oligomers. Therefore, the molecular nature of the fibril nuclei formed under either pH condition is substantially different.

Surprisingly, at pH 3 the DLS experiments showed only the presence of monomers and 4.7 nm particles after 4 hours' incubation, which does not coincide with the 2-hour lag-phase observed in the fluorescence experiments carried out in the presence of ThT and ANS. The question that arises is, "What might the species responsible for the increase in fluorescence after the lag-phase be?" The explanation could be an internal structural arrangement of PDZ3 monomers and/or oligomers. This conformational change may give rise to the exposure of hydrophobic surfaces and to β -sheet-rich species that act as precursors to the fibrillar structures. This conformational change was not detected by far-UV CD, all the spectra collected at different incubation times being practically identical in shape and displaying a minimum at 217 nm (indicative of β -strand arrangements), which increased in intensity concomitantly with incubation time (Fig. S1).

Finally, in addition to DLS and TEM evidence, the essentially irreversible formation of the supramacromolecular assemblies at acidic pH was proven by spectral analysis of samples cooled down to 25 °C after various hours' incubation at 60 °C. The fluorescence spectra remained unaltered in all the solutions (Fig. S2), whereas at neutral pH a hyperchromic effect appeared, due to the disaggregation of these assemblies into their constituent protein monomers [4].

4. Discussion

4.1. *pH-dependent differential features upon folding and misfolding of PDZ3*

In a previous paper we described the folding pathway of the PDZ3 domain from PSD95 protein at neutral pH [4]. It consists of a well organized, stepwise, reversible process in which a trimeric intermediate state, which maximally populates at 60-70 °C, may be the precursory ensemble of a misfolding pathway. During misfolding this oligomeric intermediate self-associates into 12 nm particles that, after some days, lead to annular and fibrillar supramacromolecular assemblies (Fig. 3).

In this work we have extended our experiments on PDZ3 folding to acidic pH conditions and found some significant differences. Thus, DSC experiments revealed that below pH 3 the PDZ3 domain apparently unfolds under a two-state regime as a consequence of an energetic destabilization of the intermediate state, which does not populate at any temperature. Nevertheless, the misfolding pathway still remains under these acidic pH values, since fibrillar assemblies appeared after 1-2 days incubation at 60 °C.

The DLS results also showed considerable differences during the early stages of self-assembling processes under acidic and neutral pH conditions. Thus, particles of different sizes formed in PDZ3 prior to the growth of fibrils (Fig. 3). In addition, TEM imaging only revealed fibrillar assemblies at pH 3 and the ThT and ANS fluorescence kinetic experiments were sigmoidal (lag-phase) compared to the exponential shape at neutral pH (no lag-phase). Furthermore, the whole set of kinetic experiments clearly indicated the completely irreversible nature of the misfolding species at acidic pH, which also runs contrary to the reversibility of the misfolding pathway found at neutral pH.

4.2. *A thermodynamic explanation of the irreversibility of PDZ3 misfolding at acidic pH*

As has been previously demonstrated in some examples [4,20], the energetic character of self-assembling processes can be satisfactorily explained by thermodynamic criteria. The stability of PDZ3 at $\text{pH} \leq 3$ was within the range of 15-25 $\text{kJ}\cdot\text{mol}^{-1}$ (Table 1), whereas we had previously estimated the stability of the PDZ3 misfolding assemblies to be 25 $\text{kJ}\cdot\text{mol}^{-1}$ [4]. Thus, the higher stability of the N-state (around 40 $\text{kJ}\cdot\text{mol}^{-1}$; Table 1) than that of the aggregates may well explain the reversibility of misfolding at neutral pH, while the misfolded structures are preferred under acidic pH conditions, where the fibrils are slightly more stable than the N-state itself. The fact that the aggregates assembled more quickly at pH 3 than under neutral pH conditions can be explained by the competition between fibrils and annular structures within the misfolding pathway at neutral pH, whereas at acidic pH the fibrils grew exclusively at the expense of the precursory oligomers.

4.3. *Electrostatic forces are responsible for the differential folding behavior of PDZ3 at neutral and acidic pH*

The relevant role of electrostatic interactions in the organization of amyloid fibrils has been described elsewhere [21-23]. Variations in pH can modulate the electrostatic balance between intra- and inter-molecular forces among protein molecules. The different behavior of fibril growth found in the kinetic experiments under neutral (non-sigmoidal) and acidic (sigmoidal) conditions must be a direct consequence of the different stabilities of the precursory intermediate state shown by DSC. Thus, the presence of a lag-phase at acidic pH suggests that the formation of oligomeric assemblies needs to reach a critical number/size prior to promoting further aggregation. At neutral pH this process is energetically favoured by the stable precursory intermediate state that already oligomerizes during heating from room temperature to 60 °C [4]. A similar pH-dependent behavior has been observed in human-serum albumin fibrillation processes [23].

Kinetic experiments also showed that the lag-phase at pH 3 continued for approximately two hours, but DLS did not reveal the appearance during this time of any new particle to which we might attribute the increase in fluorescence observed. Thus, under such conditions, monomeric and 4.7 nm particles remained exclusively for roughly four hours (Fig. 4). In addition, the lack of ThT and ANS fluorescence of these particles suggests that they neither exposed hydrophobic patches (ANS) nor showed typical β -aggregation (ThT), as happened at neutral pH with the corresponding trimers and 12 nm particles (Fig. 3). These observations clearly suggest that the oligomeric species will display different structural arrangements at the two pH values. Following this reasoning, the increase in fluorescence after the lag-phase at pH 3 would be the consequence of a conformational change in the protein molecules prior to the assembling of the fibrils, but we could not detect this by CD spectrum analysis.

4.4. The 1.4 Å X-ray structures of PDZ3 reveal a cluster of Glu/Asp residues as being responsible for the energetic destabilization of the unfolding intermediate at acidic pH

According to our experimental evidence, PDZ3 retains the three-state regime described at neutral pH, characterized by the presence of an energetically stable intermediate state, down to pH values of 4.0-3.5 (Fig. 1), whereas at $\text{pH} \leq 3$ the intermediate does not stabilize at any temperature. This titration behavior, as observed in DSC experiments, can be attributed to the protonation equilibria of Glu and/or Asp residues, the pKa values of which are the only ones within the range of pH 3-4.

The high-resolution X-ray structures of PDZ3 [PDB: 3I4W and 3K82 [7]] show that with only a few exceptions these residues are highly exposed to the solvent. Thus, some Glu residues appear to be mainly responsible for the attachment of the $\alpha 3$ helix to the whole PDZ3 structure. These residues, mainly E334 from strand $\beta 3$, and E396 and E401 from helix $\alpha 3$, can interact by means of the aliphatic side-chains between them, these interactions being influenced by the aromatic rings of residues F337 from the $\beta 3$ strand, and Y397 and F400 from the $\alpha 3$ helix (Fig. 7). Nevertheless, when Glu residues titrate above pH 3 ($\text{pKa} = 3.0\text{-}3.5$) the proton release should generate some degree of electrostatic repulsion between the negative charges of the Glu side-chains, where the π -electron clouds of the aromatic rings might also contribute to this energy balance to some degree [24-28]. This phenomenon gives rise to an opening of the C-terminal $\alpha 3$ helix and of the $\beta 2$ - $\beta 3$ region, as has also been shown by NMR [29]. The $\alpha 3$ helix constitutes an additional structural element compared to other typical PDZ domains, which are characterized by six β -strands and just two α -helices and which has been described elsewhere as an allosteric regulatory element that may modulate the binding affinity of this PDZ3 domain [8].

Furthermore, the X-ray structures of the loop connecting the $\beta 2$ and $\beta 3$ strands, formed by residues G329-G335 (GGEDGEG), identify the presence of a succinimide residue at the position of D332 (Fig. 7), which is considered to be an intermediate step in the process of Asp isomerization and Asn deamidation in proteins. From a conformational point of view, this circular arrangement decreases loop flexibility and at the same time results in the loss of a negative charge from the Asp carboxyl group [7]. Apart from these observations, the only relevant feature of the other acidic residues consists of a salt-bridge formed by residues R312-D357, which probably does not enter into the energetic modulation between the different pH conditions since it is buried from the solvent.

4.5. A molecular explanation for the titration behavior of PDZ3 folding and misfolding

The biophysical analysis described here has led us to conclude that the intermediate species developed by PDZ3 upon unfolding are substantially different under neutral and acidic pH (≤ 3) conditions. Thus, the equilibrium experiments showed a net energetic destabilization of the trimeric species seen at neutral pH, which persisted as monomers during

misfolding at $\text{pH} \leq 3$. In addition, the non-fluorescent character when using ThT and ANS probes suggests that the intermediate structure at acidic pH is more compact than that at neutral pH and displays a lesser hydrophobic content to the solvent.

A β -aggregation propensity analysis of the PDZ3 X-ray structures by means of the Tango algorithm [30] did not reveal any substantial differences between pH 7.5 and pH 3, residues 340-350 (pertaining to $\beta 3$ and $\alpha 1$ secondary structures) and 392-397 (deriving from $\beta 5$ and $\alpha 3$) being the most prone to β -aggregation (Fig. S3). An inspection of PDZ3 X-ray structures shows that residues E334 from $\beta 3$, and E396 and E401 from $\alpha 3$ may generate some electrostatic repulsion upon proton release at $\text{pH} \geq 3.5$ [quantified as $40 \text{ kJ}\cdot\text{mol}^{-1}$ according to DSC unfolding experiments (Fig. 2)] and thus cause the neighbouring residues from the $\beta 3$ - $\alpha 1$ and $\beta 5$ - $\alpha 3$ regions to protrude from the protein leaving them susceptible to β -aggregation (Figs 7 and S3). These partly unfolded intermediates stabilized themselves through oligomerization in the trimeric species seen at neutral pH. Nevertheless, at $\text{pH} \leq 3$ the protonated Glu residues did not develop such repellent electrostatic forces and the regions prone to aggregation did not protrude from PDZ3, which remained as a monomer in the solution. The 4.7 nm particles might arise from a non-specific aggregation of monomers, as has been shown in other examples [19]. These features may also explain the absence of fluorescence in the acidic particles and the highly fluorescent nature of those at neutral pH.

In support of these arguments it should be pointed out that the flipping of the $\alpha 3$ helix has been observed by other authors in hydrogen-exchange NMR experiments [29] and that we ourselves have described the structural interconnection between the $\alpha 3$ helix and the loop connecting the $\beta 2$ and $\beta 3$ strands [7]. Furthermore, as mentioned above, the $\alpha 3$ structural element has been proven to be a regulatory element of PDZ3 binding capacity [8] and is also an extra structural element of PDZ3, as compared to other PDZ topologies. Thus we may conclude that the C-terminal $\alpha 3$ helix also appears to be a regulatory element in the folding and assembling of the PDZ3 domain. In support of these conclusions we can point to the essentially different physico-chemical nature of the residues involved in the stacking of the $\alpha 3$ helix (mainly the above-mentioned Glu and aromatic residues) from those of the rest of the secondary structural elements, in which a collection of methyl-methyl mediated hydrophobic interactions, involving mainly aliphatic side-chains, can be found [31].

It must be pointed out that these electrostatic interactions between the $\alpha 3$ helix and the $\beta 2$ - $\beta 3$ loop residues are almost certainly ultimately responsible for the reversible formation of the annular structures seen at pH 7.5 (Fig. 3) [4]. Further research is needed into this point to clarify whether such annular arrangements can cause toxicity *in vivo*, as has been shown in some other examples [6,32,33]. Less is known about the molecular mechanism by which they are formed from PDZ3 trimers or about the hypothetical biological relevance of these supramacromolecular assemblies within the context of the “hub” protein PSD95, a member of the MAGUK family, the function of which has been described as a recruitment centre for proteins at the inner face of the neuron membrane [34,35]. Within this context, the electrostatics of the PDZ3 module might be modulated by the membrane micro-environment [24-28] to control, for instance, the spatial- and time-dependent release (as hub proteins do) of protein ligands via a conformational change in the $\alpha 3$ helix [8]. A similar distribution of charged and aromatic groups can be seen at the $\alpha 3$ helix in the PDZ3 module of the homologous DLG-2 protein from *drosophila* (PDB: 1PDR). Structural modelling has also begun to point to some possibility of a functional role for the linker sequence connecting PDZ1 and PDZ2 in PSD95, as well as a regulatory role for other linker segments in PSD95 [36-38].

Acknowledgements

This work was financed by grants CVI-05915 and CVI-05063 from the Andalusian Regional Government and BIO2009-13261-C02-01/02 from the Spanish Ministry of Science and Education. J.M.C. is a Spanish Government PhD fellow. E.S.C. was funded by a contract provided by the University of Granada. We thank our colleague Dr. Jon Trout for revising and correcting our English text.

Figure Captions

Figure 1. Calorimetric traces of PDZ3. The experiments were carried out at a protein concentration of $1.5 \text{ mg}\cdot\text{mL}^{-1}$ and obtained at different pH values (gray symbols). Their best fittings to the respective equilibrium models are also shown (solid black lines).

Figure 2. Analysis of the electrostatic contribution of Glu and Asp residues to the unfolding enthalpy of PDZ3. The linear correlation among enthalpies derived from two-state analyses (circles) is represented by a solid line. Dashed gray lines represent the 95 % confidence intervals of this linear regression. Gray squares with error bars are the enthalpy values for three-state analyses. The diamond represents the convergence value of $54 \text{ J}\cdot\text{g}^{-1}$ at $110 \text{ }^\circ\text{C}$ for small globular proteins.

Figure 3. A representation of the most salient features of PDZ3 misfolding under neutral and acidic pH conditions.

Figure 4. The DLS distribution of hydrodynamic species of PDZ3 at pH 3. A protein solution at a concentration of $8 \text{ mg}\cdot\text{mL}^{-1}$ was heated to $60 \text{ }^\circ\text{C}$, at which point the mass evolution of the species as a function of incubation time (in gray) was recorded.

Figure 5. TEM analysis of PDZ3 supramacromolecular assemblies at pH 3. The left panel shows the 4.7 nm particles, seen also by DLS, after 4 hours' incubation of an $8 \text{ mg}\cdot\text{mL}^{-1}$ sample at $60 \text{ }^\circ\text{C}$. The fibrils organized after 1 and 7 days' incubation are shown in the centre and right panels respectively. The scale bar corresponds to 200 nm .

Figure 6. Growth kinetics of ThT and ANS fluorescence emission. An $8 \text{ mg}\cdot\text{mL}^{-1}$ PDZ3 sample was incubated in the presence of $12 \text{ }\mu\text{M}$ of ThT or $20 \text{ }\mu\text{M}$ ANS at pH 3. Time dependence of the fluorescence signal versus incubation time at $60 \text{ }^\circ\text{C}$ for ThT (upper panel) and ANS (lower panel).

Figure 7. Electrostatic surface of the PDZ3 domain calculated with APBS [39] implemented in PyMOL (DeLano Scientific Palo Alto, CA; www.pymol.org). A detail of the interface of the $\alpha 3$ helix at the C-terminus of PDZ3 with the $\beta 2$ - $\beta 3$ loop showing the spatial arrangement of Phe, Tyr and Glu residues is shown. The green dashed lines represent relative distances between residues.

Table 1. Thermodynamic parameters for the unfolding of the PDZ3 domain of PSD95 obtained from fittings to $N \rightleftharpoons U$ and $3N \rightleftharpoons 3U \rightleftharpoons I_3$ models *

pH	T_{N-U} (°C)	$\Delta H_{N-U}(T_{N-U})$ (kJ·mol⁻¹)	T_{U-I_n} (°C)	$\Delta H_{U-I_n}(T_{U-I_n})$ (kJ·mol⁻¹)	$\Delta G_{N-U}(298)$ (kJ·mol⁻¹)	R^2
$3N \rightleftharpoons 3U \rightleftharpoons I_3$						
7.5	70.4 ± 0.5	335 ± 20	79.2 ± 1.2	-130 ± 10	39 ± 6	0.996
4.0	70.0 ± 0.3	335 ± 10	89.3 ± 1.4	-145 ± 30	33 ± 8	0.983
3.5	67.4 ± 0.2	340 ± 10	77.0 ± 1.7	-110 ± 20	37 ± 7	0.999
$N \rightleftharpoons U$						
3.0	59.0 ± 0.1	325 ± 5	-	-	26 ± 2	0.999
2.5	51.6 ± 0.1	280 ± 5	-	-	17 ± 2	0.998
2.0	47.5 ± 0.1	265 ± 5	-	-	14 ± 2	0.998

* The unfolding values at pH 7.5 were obtained from a previous work [4]. The error intervals for the three-state parameters were estimated by comparing the values obtained when $n = 3$ and 4. The errors in two-state analyses were taken to be three times the fitting standard errors (99 % confidence).

References

- [1] Frieden, C. Protein aggregation processes: In search of the mechanism. *Protein Sci* 16 (2007) 2334-2344.
- [2] Pappu, R. V., X. Wang, A. Vitalis, and S. L. Crick A polymer physics perspective on driving forces and mechanisms for protein aggregation. *Arch Biochem Biophys* 469 (2008) 132-141.
- [3] Knowles, T. P., C. A. Waudby, G. L. Devlin, S. I. Cohen, A. Aguzzi, M. Vendruscolo, E. M. Terentjev, M. E. Welland, and C. M. Dobson An analytical solution to the kinetics of breakable filament assembly. *Science* 326 (2009) 1533-1537.
- [4] Murciano-Calles, J., E. S. Cobos, P. L. Mateo, A. Camara-Artigas, and J. C. Martinez An oligomeric equilibrium intermediate as the precursory nucleus of globular and fibrillar supramacromolecular assemblies in a PDZ domain. *Biophys J* 99 (2010) 263-272.
- [5] Lansbury, P. T. and H. A. Lashuel A century-old debate on protein aggregation and neurodegeneration enters the clinic. *Nature* 443 (2006) 774-779.
- [6] Jang, H., J. Zheng, R. Lal, and R. Nussinov New structures help the modeling of toxic amyloidbeta ion channels. *Trends Biochem Sci* 33 (2008) 91-100.
- [7] Camara-Artigas, A., J. Murciano-Calles, J. A. Gavira, E. S. Cobos, and J. C. Martinez Novel conformational aspects of the third PDZ domain of the neuronal post-synaptic density-95 protein revealed from two 1.4A X-ray structures. *J Struct Biol* 170 (2010) 565-569.
- [8] Petit, C. M., J. Zhang, P. J. Sapienza, E. J. Fuentes, and A. L. Lee Hidden dynamic allostery in a PDZ domain. *Proc Natl Acad Sci U S A* 106 (2009) 18249-18254.
- [9] Gill, S. C. and P. H. von Hippel Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem* 182 (1989) 319-326.
- [10] Viguera, A. R., J. C. Martinez, V. V. Filimonov, P. L. Mateo, and L. Serrano Thermodynamic and kinetic analysis of the SH3 domain of spectrin shows a two-state folding transition. *Biochemistry* 33 (1994) 2142-2150.
- [11] Cobos, E. S., M. T. Pisabarro, M. C. Vega, E. Lacroix, L. Serrano, J. Ruiz-Sanz, and J. C. Martinez A miniprotein scaffold used to assemble the polyproline II binding epitope recognized by SH3 domains. *J Mol Biol* 342 (2004) 355-365.
- [12] Martinez, J. C., M. el Harrou, V. V. Filimonov, P. L. Mateo, and A. R. Fersht A calorimetric study of the thermal stability of barnase and its interaction with 3'GMP. *Biochemistry* 33 (1994) 3919-3926.
- [13] Fu, L. and E. Freire On the origin of the enthalpy and entropy convergence temperatures in protein folding. *Proc Natl Acad Sci U S A* 89 (1992) 9335-9338.

- [14] Baldwin, R. L. and N. Muller Relation between the convergence temperatures T_h^* and T_s^* in protein unfolding. *Proc Natl Acad Sci U S A* 89 (1992) 7110-7113.
- [15] Privalov, P. L. Stability of proteins. Proteins which do not present a single cooperative system. *Adv Protein Chem* 35 (1982) 1-104.
- [16] Ragone, R. and G. Colonna Enthalpy-entropy balance and convergence temperatures in protein unfolding. *J Biol Chem* 269 (1994) 4047-4049.
- [17] Harper, J. D. and P. T. Lansbury, Jr. Models of amyloid seeding in Alzheimer's disease and scrapie: mechanistic truths and physiological consequences of the time-dependent solubility of amyloid proteins. *Annu Rev Biochem* 66 (1997) 385-407.
- [18] Kad, N. M., S. L. Myers, D. P. Smith, D. A. Smith, S. E. Radford, and N. H. Thomson Hierarchical assembly of beta2-microglobulin amyloid in vitro revealed by atomic force microscopy. *J Mol Biol* 330 (2003) 785-797.
- [19] Smith, A. M., T. R. Jahn, A. E. Ashcroft, and S. E. Radford Direct observation of oligomeric species formed in the early stages of amyloid fibril formation using electrospray ionisation mass spectrometry. *J Mol Biol* 364 (2006) 9-19.
- [20] Kardos, J., K. Yamamoto, K. Hasegawa, H. Naiki, and Y. Goto Direct measurement of the thermodynamic parameters of amyloid formation by isothermal titration calorimetry. *J Biol Chem* 279 (2004) 55308-55314.
- [21] Jeppesen, M. D., P. Westh, and D. E. Otzen The role of protonation in protein fibrillation. *FEBS Lett* 584 (2010) 780-784.
- [22] Gerber, R., A. Tahiri-Alaoui, P. J. Hore, and W. James Conformational pH dependence of intermediate states during oligomerization of the human prion protein. *Protein Sci* 17 (2008) 537-544.
- [23] Juarez, J., S. G. Lopez, A. Cambon, P. Taboada, and V. Mosquera Influence of electrostatic interactions on the fibrillation process of human serum albumin. *J Phys Chem B* 113 (2009) 10521-10529.
- [24] Lockhart, D. J. and P. S. Kim Electrostatic screening of charge and dipole interactions with the helix backbone. *Science* 260 (1993) 198-202.
- [25] Ma, J. C. and D. A. Dougherty The Cation-minus signpi Interaction. *Chem Rev* 97 (1997) 1303-1324.
- [26] Lund, M. and B. Jonsson On the charge regulation of proteins. *Biochemistry* 44 (2005) 5722-5727.
- [27] Berry, B. W., M. M. Elvekrog, and C. Tommos Environmental modulation of protein cation-pi interactions. *J Am Chem Soc* 129 (2007) 5308-5309.

- [28] Wu, R. and T. B. McMahon Investigation of cation- π interactions in biological systems. *J Am Chem Soc* 130 (2008) 12554-12555.
- [29] Feng, H., N. D. Vu, and Y. Bai Detection of a hidden folding intermediate of the third domain of PDZ. *J Mol Biol* 346 (2005) 345-353.
- [30] Fernandez-Escamilla, A. M., F. Rousseau, J. Schymkowitz, and L. Serrano Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nat Biotechnol* 22 (2004) 1302-1306.
- [31] Gianni, S., T. Walma, A. Arcovito, N. Calosci, A. Bellelli, A. Engstrom, C. Travaglini-Allocatelli, M. Brunori, P. Jemth, and G. W. Vuister Demonstration of long-range interactions in a PDZ domain by NMR, kinetics, and protein engineering. *Structure* 14 (2006) 1801-1809.
- [32] Shirwany, N. A., D. Payette, J. Xie, and Q. Guo The amyloid beta ion channel hypothesis of Alzheimer's disease. *Neuropsychiatr Dis Treat* 3 (2007) 597-612.
- [33] Lashuel, H. A. and P. T. Lansbury, Jr. Are amyloid diseases caused by protein aggregates that mimic bacterial pore-forming toxins? *Q Rev Biophys* 39 (2006) 167-201.
- [34] Kim, E. and M. Sheng PDZ domain proteins of synapses. *Nat Rev Neurosci* 5 (2004) 771-781.
- [35] Feng, W. and M. Zhang Organization and dynamics of PDZ-domain-related supramodules in the postsynaptic density. *Nat Rev Neurosci* 10 (2009) 87-99.
- [36] Long, J. F., H. Tochio, P. Wang, J. S. Fan, C. Sala, M. Niethammer, M. Sheng, and M. Zhang Supramodular structure and synergistic target binding of the N-terminal tandem PDZ domains of PSD-95. *J Mol Biol* 327 (2003) 203-214.
- [37] Korkin, D., F. P. Davis, F. Alber, T. Luong, M. Y. Shen, V. Lucic, M. B. Kennedy, and A. Sali Structural modeling of protein interactions by analogy: application to PSD-95. *PLoS Comput Biol* 2 (2006) e153.
- [38] Wang, W., J. Weng, X. Zhang, M. Liu, and M. Zhang Creating conformational entropy by increasing interdomain mobility in ligand binding regulation: a revisit to N-terminal tandem PDZ domains of PSD-95. *J Am Chem Soc* 131 (2009) 787-796.
- [39] Dolinsky, T. J., P. Czodrowski, H. Li, J. E. Nielsen, J. H. Jensen, G. Klebe, and N. A. Baker PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res* 35 (2007) W522-525.

Figure 1
[Click here to download high resolution image](#)

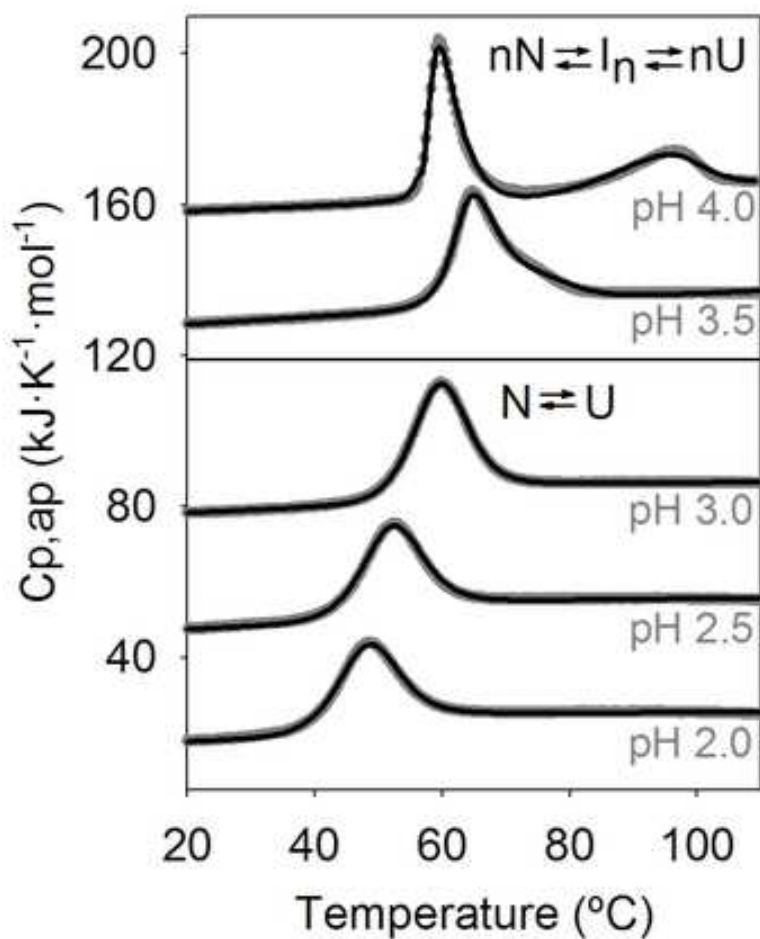


Figure 2
[Click here to download high resolution image](#)

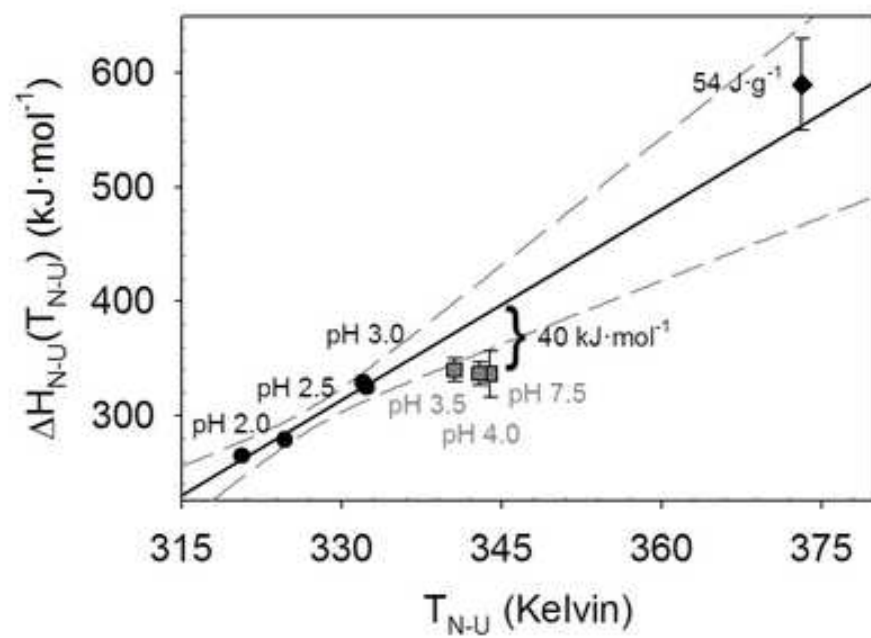


Figure 3

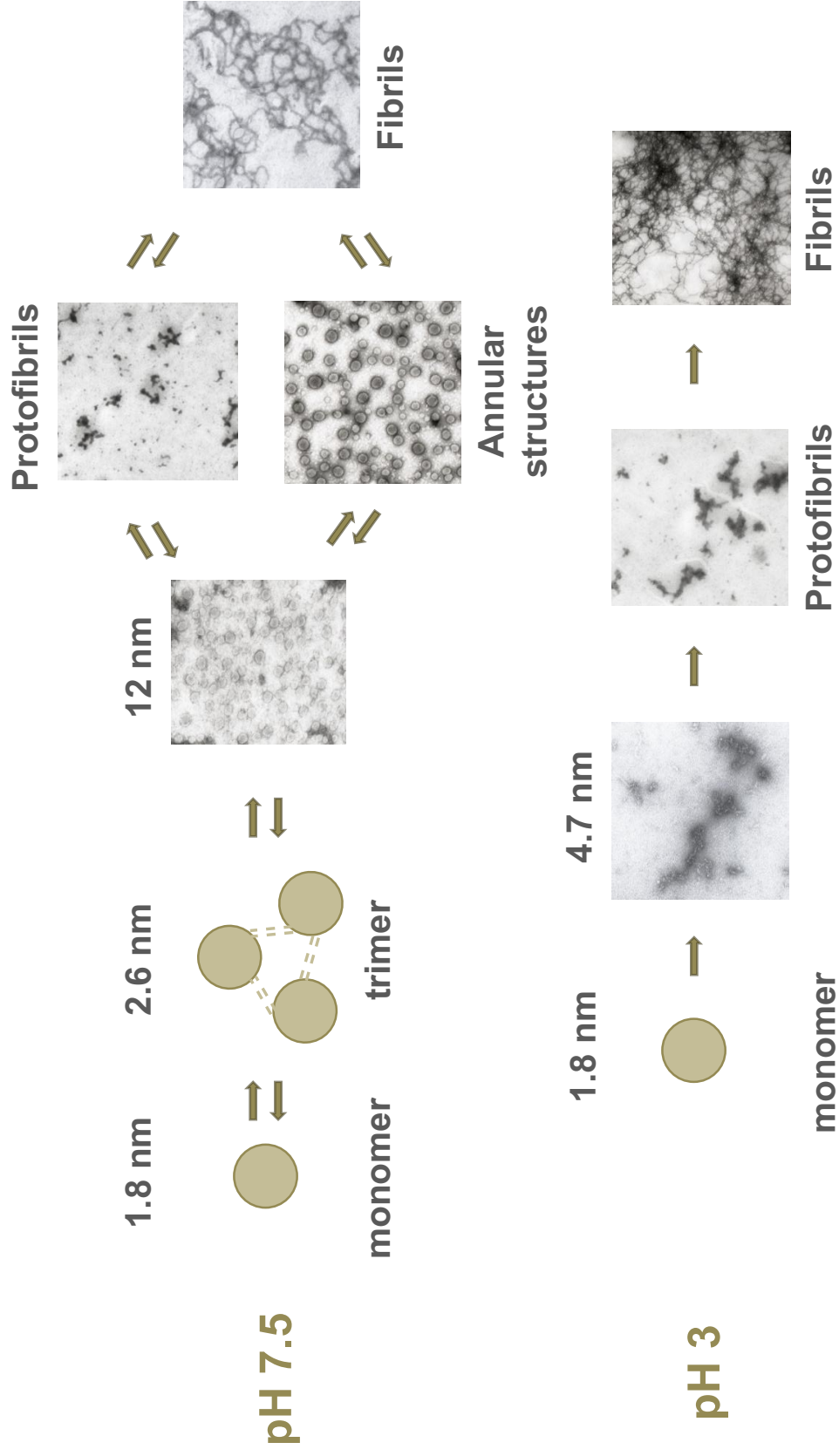


Figure 4
[Click here to download high resolution image](#)

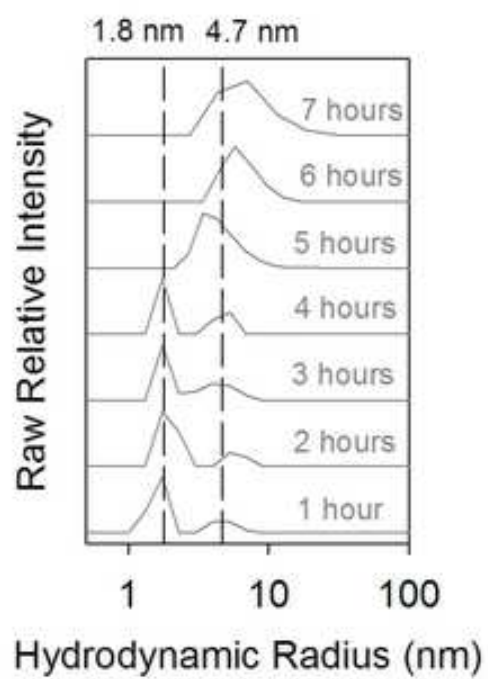
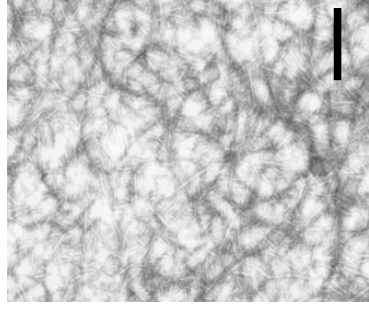
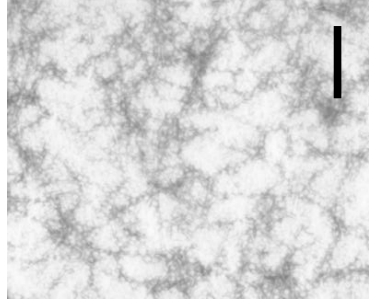


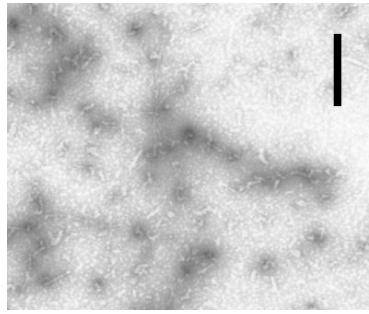
Figure 5



7 days



1 day



4 hours

Figure 6 upper panel
[Click here to download high resolution image](#)

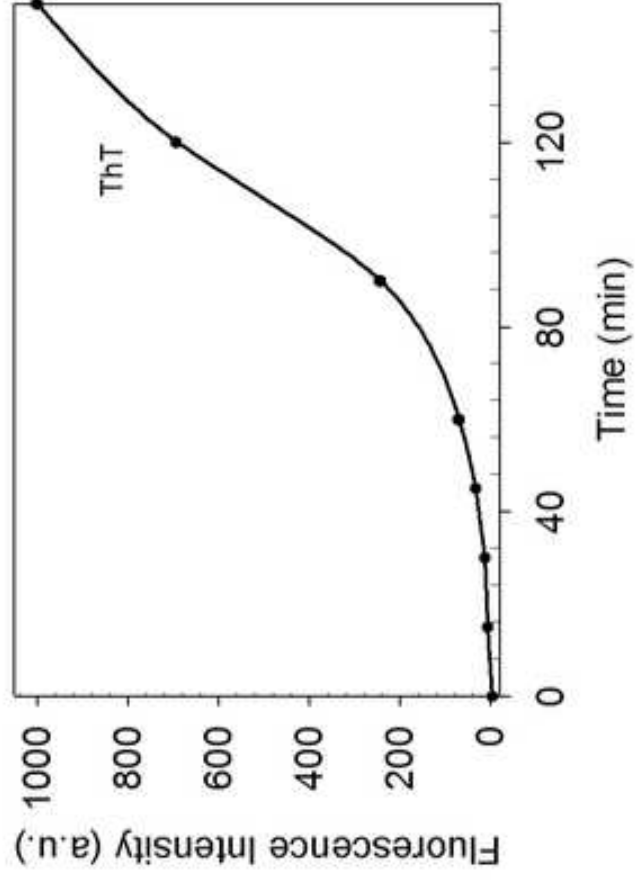


Figure 6 lower panel
[Click here to download high resolution image](#)

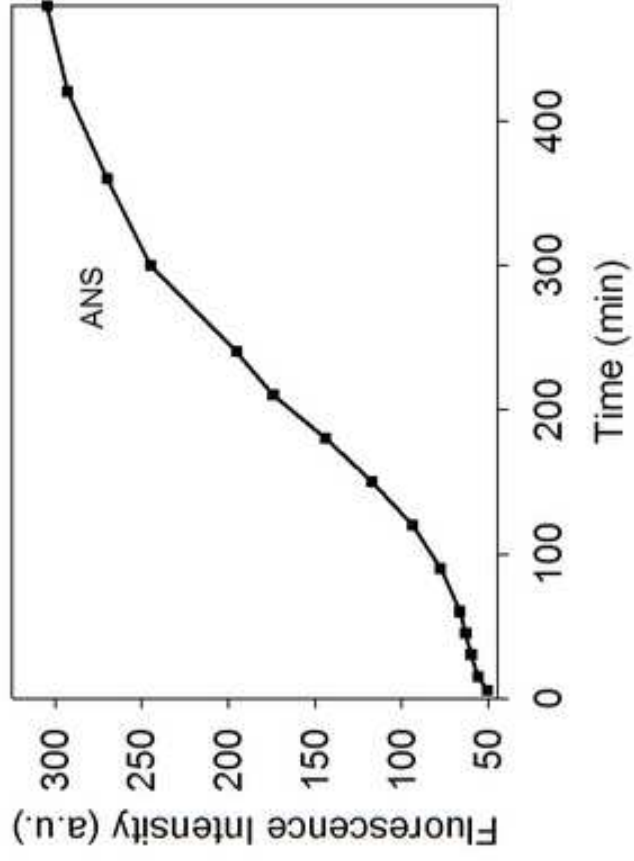


Figure 7

