



*ugr*

Universidad  
de **Granada**

**TESIS DOCTORAL**

**DOMINIOS WW: ESTUDIO DEL EQUILIBRIO  
CONFORMACIONAL Y DEL RECONOCIMIENTO  
DE LIGANDOS**

**Manuel Iglesias Bexiga**

Granada, 2011

Editor: Editorial de la Universidad de Granada  
Autor: Manuel Iglesias Bexiga  
D.L.: GR 4534-2011  
ISBN: 978-84-694-6019-1



# **DOMINIOS WW: ESTUDIO DEL EQUILIBRIO CONFORMACIONAL Y DEL RECONOCIMIENTO DE LIGANDOS**

Memoria presentada para aspirar al grado  
de Doctor por la Universidad de Granada

Fdo.: **Manuel Iglesias Bexiga**

Granada, 20 de Mayo de 2011

Los directores de la tesis:

**Dra. Irene Luque Fernández**  
Profesora Titular  
Departamento de Química-Física  
Universidad de Granada

**Dra. Eva Sánchez Cobos**  
Contratada del Programa de  
Incorporación de Doctores del Plan  
Propio de la UGR  
Departamento de Química-Física  
Universidad de Granada



## AGRADECIMIENTOS

Me resulta complicado expresar en unos cuantos párrafos toda la gratitud acumulada hacia todas las personas que me han ayudado durante estos años de tesis, que no es poca... Así que espero aprovechar bien este particular apartado y no despistarme para no dejarme a nadie en el tintero.

En primer lugar, quisiera agradecer de manera especial la dedicación y la supervisión de de mis directores de Tesis. Gracias Irene por darme la oportunidad de hacer lo que más me gusta, por confiar en mí, por tu infinita paciencia, por tu cariño y por enseñarme el camino. Mi más sincero agradecimiento a Eva, por sus inagotables consejos científicos y humanos, por su amistad y cercanía. Todas las reuniones que hemos tenido durante estos años, me han enseñado a razonar, a enfrentarme a un proyecto científico y me habéis transmitido vuestro entusiasmo. Pero mucho más importante, es que siempre he podido tener vuestro consejo y apoyo. ¡Gracias a las dos por todos estos años!

No puedo dar las gracias a mis directores de Tesis sin incluir también a José y Javier Ruiz que han sido para mí como los terceros directores de Tesis. Todo un lujo. Gracias por vuestros consejos (dentro y fuera del laboratorio), por vuestra simpatía, vuestra ayuda y vuestro apoyo incondicional. ¡Mil gracias!

Asimismo, también quiero agradecer a Pedro Luís Mateo el haberme permitido realizar esta Tesis Doctoral en su grupo de investigación.

A la Dra. María J. Macías (Instituto de Investigación Biomédica, Barcelona), por haberme aceptado en su grupo de investigación en repetidas ocasiones como si fuese uno más y por haberme supervisado y aconsejado en la resolución por RMN de las estructuras que se estudian en esta tesis.

Al Dr. Francisco Blanco (CIC Biogune, Bilbao), pues gracias a la colaboración mantenida con él, se pudieron producir las muestras de proteína y medir los espectros necesarios para la determinación estructural mediante RMN.

Gracias a todos los demás profesores del grupo de investigación y del resto del departamento por el apoyo y el calor humano que siempre me han brindado. Gracias Anabel, Salva, Quico, Nico, José Manuel Sánchez Ruiz, Beatriz, Asun, Antonio Parody, Mercedes, Isa, Enrique, Obdulio, etc. Gracias también a Ana Cámara de la Universidad de Almería por el interés que ha mostrado por mi en estos años y por sus sabios consejos.

Como no, también quiero hacer referencia a todos mis compañeros de laboratorio. A ellos, no sólo me une la ciencia, sino que también la convivencia a lo largo de estos años. Es mucho lo compartido durante tanto tiempo, por lo que además de haber influido en mi formación científica, también habéis participado en el enriquecimiento personal que ha supuesto este largo periplo. ¡Puedo decir que sé algo más del mundo en el que vivo gracias a vosotros! En especial quiero mostrar mi gratitud a Bertrand que, a pesar de tirar para adelante con gran cantidad de trabajo y estar muy ocupado, siempre ha tenido un hueco para enseñarme y siempre ha estado dispuesto a ayudar en lo que hiciese falta. No tengo palabras para agradecer todo lo que me enseñaste durante los años compartidos en el laboratorio. A María del Mar por sus charlas científicas y por sus inestimables explicaciones y consejos. A los compañeros que estaban cuando llegue y ya se fueron, pero que tanto me ayudaron cuando más perdido estaba: Adela, Andrés y José Manuel. A los compañeros que me han acompañado en todo el periplo, me han animado a diario y sin los que, posiblemente, esta tesis no hubiese llegado a buen puerto: Carles, Lorena, José Luís, Javi Murciano, Sara, Ana Mari, Fran y David. ¡Gracias por todo lo compartido!

A todos mis compañeros del grupo del lado que siempre habéis estado ahí, atentos y dispuestos a echar una mano: a los que se fueron Raquel, Fernando, Antonio y David, y a los que aún están, Rocí, Inma, Álvaro, Ángel, Mari Carmen, Sisi y Héctor.

A todos los compañeros que me ayudaron y me mostraron su apoyo en mis andaduras científicas allende Granada. En el IRB de Barcelona, a Román, Pau, Lidia, Tiago, Eric y Nina. En el CNIO de Madrid, Alicia, Simone, Richard, Dani y Eva.

A José Miguel y Araceli que siempre me han echado un cable con los fax, fotocopidora, envíos, recepción de correos, etc, y con las increíbles tandas de geles de acrilamida que requerían las purificaciones de proteína, respectivamente.

A los profesores y a los conferenciantes invitados en el Curso de Doctorado Inter-universitario “Estructura y Función de Proteínas”.

A la FEBS por sus ayudas económicas y becas para asistir a congresos.

Mil gracias a David Fairén, “mi pareja de hecho”, con el permiso de Lluna claro está (gracias a ti también). Gracias por hacer que esta ciudad fuese mi casa nada más aparecer por aquí, por hacer que me interesase por otras áreas de la ciencia, por la lucha precaria, por la dignidad en la investigación, por la fotografía, etc. Todo un referente moral e intelectual para mí. A los grandes amigos que he hecho en esta etapa “granaína” de mi vida, gracias por apoyo fundamental y por vuestro cariño a lo largo de estos años: Gonso, Carlita, Javi y Elena. Mención especial merece mi incondicional Tito de Jun, siempre dispuesto a hacer cualquier favor y a ayudar con lo que sea. No me olvido de los primeros años con la extraordinaria compañía de Quique, Malu, Quique-Aarón, su casita con palomar y los buenos ratos vividos en la taberna de Laif y Ruth. No me puedo olvidar tampoco de “Los Lila”, constantes, originales, divertidos y teatreros donde los haya: Ana, Elena, Javi, Borja, Susana, Nachete... Todos vosotros sois una parte importante de esta Tesis.

A mis amigos de toda la vida. A los científicos Toño, Brian, Mónica, Sotres y especialmente a Pablo Mateos (motivador de la ciencia y persona plena de curiosidad científica como ninguna...). A los que siempre cuento con ellos, primo Pablo y David. A los de casa y desde la infancia, Ricardo y Marquitos. A los argentinos, Kike, Carol, Ruben “flaquito”, Lala, Mel, Blanca e Isalena. A los del “Flori” (Jaume, Julio, Yago, Jorge, Laura, Eva, Darío, Pablo, Ana, Sonia, y un larguísimo etc). A Guille y Yoli de Málaga. Al grupo cuatro de México. ¡Gracias compañeros y amigos! Sin vuestros consejos, apoyo y comprensión esto habría sido mucho más difícil.

Y no por ser los últimos son los menos importantes, al contrario, quiero expresar mi reconocimiento y gratitud a las personas más importantes de mi vida y a quienes he dedicado este

trabajo: a mis padres, a mis hermanos y a Bego. Qué puedo decir de vosotros...A mis padres se lo debo todo; gracias por ser tan buenos, por vuestros consejos, por apoyarme, por vuestra generosidad, por haber creído siempre en mí y por hacerme sentir tan querido. A mis hermanos, Javier y Alejandro, por ser tan distintos como especiales para mí, por vuestro cariño, por los ánimos recibidos, porque sé que siempre contaré con vosotros (así como vosotros conmigo) y por toda la seguridad y ánimo que te da el saber esto “a ciencia cierta”.

Por último, a Bego, por cambiar mi vida y por estar día a día a mi lado (a pesar de la distancia que nos separa habitualmente...), apoyándome, aguantándome, animándome para superar los momentos bajos y alegrándose de los buenos. A su familia, que siempre ha sabido interesarse y apoyarme en todo. Sin duda, Bego, esta tesis también es tuya; nunca podré agradecerte suficientemente.





**A mis padres**

**A mis hermanos**

**A Bego**



“En Biología, nos enfrentamos a un único grupo de átomos, del que existe una sola copia, y produce acontecimientos ordenadamente, armonizados entre sí de modo maravilloso y con el ambiente siguiendo las leyes más sutiles. (...) En los grados de organización superiores, las copias se multiplican, es cierto. Pero ¿hasta qué punto? Tengo entendido que hasta algo así como  $10^{14}$  en un mamífero adulto. ¡Qué es esto! Sólo una millonésima parte del número de moléculas en una pulgada cúbica de aire. Aunque son voluminosas, esas copias no formarían, por coalescencia, más que una minúscula gota de líquido. Y obsérvese la forma en que están distribuidas en la realidad. Cada célula alberga sólo una de ellas (o dos, si pensamos en un organismo diploide). Puesto que conocemos el que esta minúscula oficina central posee en una célula aislado, ¿no se parecen a estaciones de gobierno local dispersas por todo el cuerpo, que se comunican entre sí con suma facilidad gracias a una clave común para todas ellas?”

(Erwin Schrödinger, “¿Qué es la vida?”, Dublín, 1944)



# ÍNDICE



<b>CAPÍTULO 1. Introducción</b>	<b>31</b>
<b>1.1. Reconocimiento molecular y diseño molecular en la era post-genómica</b>	<b>31</b>
<b>1.2. Interacciones de secuencias ricas en prolina con módulos de reconocimiento molecular.</b>	<b>32</b>
<b>1.3. Los dominios WW</b>	<b>36</b>
1.3.1. Estructura y clasificación de dominios WW	37
1.3.2. Plegamiento y estabilidad de dominios WW	40
<b>1.4. Los dominios WW como objeto del diseño racional de inhibidores</b>	<b>42</b>
1.4.1. Dominios L víricos como dianas para el diseño de antivirales de amplio espectro	42
1.4.2. Las funciones celulares de la proteína hYap, su regulación y el tándem de dominios WW	46
<b>1.5. Objetivos de la Tesis</b>	<b>52</b>
<b>CAPÍTULO 2. Estudio termodinámico de la estabilidad de dominios WW</b>	<b>55</b>
<b>2.1. Introducción</b>	<b>55</b>
<b>2.2. Caracterización del equilibrio conformacional del dominio hNedd4-WW4</b>	<b>57</b>
2.2.1. Espectros de dicroísmo circular	57
2.2.2. Caracterización del desplegamiento térmico del dominio hNedd4-WW4	60
2.2.2.1. Desplegamiento térmico mediante técnicas espectroscópicas	60
2.2.2.2. Desplegamiento térmico mediante calorimetría (DSC)	63
2.2.3. Experimentos de doble perturbación	72
<b>2.3. Caracterización de otros dominios WW</b>	<b>77</b>
2.3.1. Espectros de dicroísmo circular	77
2.3.2. Desnaturalización térmica en distintas condiciones de pH	79
2.3.3. Experimentos de doble perturbación	91

<b>2.4. Perspectivas y futuro en el estudio de plegamiento de dominios WW</b>	<b>94</b>
<b>CAPÍTULO 3. Estudio de unión de dominios L víricos al tercer dominio WW de la ligasa de ubiquitina humana Nedd4</b>	<b>99</b>
<b>3.1. Introducción</b>	<b>99</b>
<b>3.2. Estudio de unión de distintos dominios L víricos al dominio hNedd4-WW3</b>	<b>100</b>
3.2.1. Descripción de los ligandos	100
3.2.2. Caracterización termodinámica de la unión	<b>102</b>
<b>3.3. Especificidad de unión de los dominios L a diferentes dominios WW</b>	<b>110</b>
3.3.1. Análisis termodinámico de las interacciones de los dominios hNedd4-WW3, hYap-WW1, hYap-WW1(L30K) y hYap-WW2 con los dominios L víricos	111
3.3.2. Determinación de la capacidad calorífica de unión mediante ITC	<b>114</b>
<b>3.4. Caracterización estructural mediante RMN</b>	<b>116</b>
3.4.1. Identificación de los residuos del dominio hNedd4-WW3 implicados en el reconocimiento de los dominios L víricos HTLV1-b y Ébola-a	116
3.4.2. Asignación de los espectros de RMN del dominio hNedd4-WW3	119
3.4.3. Estructura secundaria de los complejos	120
3.4.4. Asignación de los ligandos en los espectros de RMN	122
3.4.5. Cálculo y evaluación de la calidad de las estructuras	<b>124</b>
<b>3.5. Descripción las estructuras de los complejos del dominio hNedd4-WW3</b>	<b>128</b>
3.5.1. Análisis del modo de unión de los dominios L víricos al dominio hNedd4-WW3	130
3.5.2. Comparación de los complejos formados por hNedd4-WW3 con otras estructuras	<b>133</b>

<b>CAPÍTULO 4. Reconocimiento molecular mediado por el tándem de dominios WW de la proteína humana hYap</b>	<b>139</b>
<b>4.1. Introducción</b>	<b>139</b>
<b>4.2. Caracterización termodinámica de la unión de ligandos naturales a los dominios WW de hYap</b>	<b>143</b>
4.2.1. Disección de las contribuciones termodinámicas a la afinidad de unión	147
4.2.2. Potencial electrostático y especificidad de unión	<b>149</b>
<b>4.3. Estudio del equilibrio conformacional de los dominios WW de la isoforma hYap2</b>	<b>153</b>
4.3.1. Caracterización del equilibrio conformacional de los dominios individuales hYap-WW1 y hYap-WW2	153
4.3.2. Caracterización del equilibrio conformacional del tándem de dominios WW de hYap	<b>156</b>
<b>4.4. Reconocimiento de ligandos naturales por el tándem de dominios WW de hYap</b>	<b>162</b>
<b>4.5. Implicaciones funcionales del segundo dominio WW en la isoforma hYap2</b>	<b>167</b>
<b>CAPÍTULO 5. Conclusiones</b>	<b>171</b>
<b>CAPÍTULO 6. Materiales y Métodos</b>	<b>175</b>
<b>6.1. Clonado, expresión y purificación de los dominios WW</b>	<b>175</b>
6.1.1. Dominios WW	175
6.1.2. Dominio hNedd4-WW3 enriquecido en <sup>13</sup> C y/o <sup>15</sup> N	<b>178</b>
<b>6.2. Preparación de muestras</b>	<b>180</b>
6.2.1. Preparación de disoluciones de proteína	180
6.2.2. Preparación de las disoluciones de ligando	181
6.2.3. Preparación de disoluciones de agente desnaturalizante (urea)	<b>182</b>
<b>6.3. Calorimetría diferencial de barrido (DSC)</b>	<b>183</b>
6.3.1. Introducción	183
6.3.2. Realización del experimento calorimétrico	186

6.3.3. Cálculo de la capacidad calorífica molar parcial de la proteína	187
6.3.4. Análisis de las trazas calorimétricas	189
<b>6.4. Dicroísmo circular</b>	<b>189</b>
6.4.1. Introducción	189
6.4.2. Obtención de espectros de dicroísmo circular	191
6.4.3. Estudio del desplegamiento térmico de dominios WW	192
6.4.4. Análisis de las curvas de desplegamiento térmico	193
<b>6.5. Espectroscopía de fluorescencia</b>	<b>193</b>
6.5.1. Introducción	193
6.5.2. Experimento de desplegamiento térmico seguido por fluorescencia	196
6.5.3. Experimento de titulación proteína-ligando mediante espectroscopía de fluorescencia	196
6.5.4. Análisis de un experimento de titulación mediante espectroscopía de fluorescencia	197
<b>6.6. Calorimetría isotérmica de titulación (ITC)</b>	<b>199</b>
6.6.1. Introducción	199
6.6.2. Realización de los experimentos de ITC	201
6.6.3. Análisis de los experimentos de ITC	202
<b>6.7. Resolución estructural de complejos proteína-ligando mediante Resonancia Magnética Nuclear (RMN)</b>	<b>204</b>
6.7.1. Introducción	204
6.7.2. Adquisición, procesado y asignación de los espectros de RMN	210
6.7.3. Cálculo de las estructuras	212
6.7.4. Evaluación de las estructuras	213
<b>6.8. Cálculo del potencial electrostático de una proteína mediante el programa DelPhi</b>	<b>214</b>
6.8.1. Generación de modelos estructurales	215
<b>BIBLIOGRAFÍA</b>	<b>219</b>

<b>APÉNDICES</b>	<b>247</b>
<b>Apendice I. Modelo de equilibrio de dos estados utilizado para el análisis de las trazas calorimétricas de DSC</b>	<b>247</b>
<b>Apendice II. Modelo de equilibrio de dos estados utilizado para la desnaturalización térmica seguida por CD</b>	<b>252</b>
<b>Apendice III. Modelo de unión de un ligando a una macromolécula con “n” sitios idénticos e independientes para ITC</b>	<b>255</b>
<b>Apendice IV. Modelo de unión de un ligando a una macromolécula con “m” clases de sitios diferentes e independientes para ITC</b>	<b>258</b>
<b>Apéndice V. Modelo de unión a una macromolécula (con un solo sitio de unión) de un ligando A con afinidad moderada por desplazamiento de otro B de afinidad baja</b>	<b>261</b>
<b>Apéndice VI. Artículo: "Thermodynamic Characterization of the Folding Equilibrium of the Human Nedd4-WW4 Domain: At the Frontiers of the Cooperative Folding"</b>	<b>264</b>



## ABREVIATURAS



## ABREVIATURAS

<b>ADN</b>	Ácido desoxirribonucleico
<b>ARN</b>	Ácido ribonucleico
<b>Abs</b>	Absorbancia
<b>C</b>	Concentración de proteína
<b>CD (ó DC)</b>	Dicroísmo circular
<b>C<sub>p</sub> ó C<sub>p,p</sub></b>	Capacidad calorífica molar parcial de una proteína
<b>C<sub>p,ap</sub></b>	Capacidad calorífica aparente de una proteína
<b>C<sub>p,app</sub></b>	Capacidad calorífica aparente de exceso en bruto de una proteína
<b>C<sub>p,D</sub></b>	Capacidad calorífica molar parcial del estado desnaturalizado
<b>C<sub>p,N</sub></b>	Capacidad calorífica molar parcial del estado nativo
<b>C<sub>p,prot</sub></b>	Capacidad calorífica de protonación de una proteína
<b>C<sub>p,s</sub></b>	Capacidad calorífica molar parcial del disolvente
<b>DSC (ó CDB)</b>	Calorimetría diferencial de barrido
<b>D</b>	Estado desnaturalizado de la proteína
<b>Da</b>	Dalton
<b>EDTA</b>	Ácido etilendiamino tetracético
<b>F<sub>M</sub></b>	Intensidad de fluorescencia de la proteína libre
<b>F<sub>ML</sub></b>	Intensidad de fluorescencia del complejo proteína-ligando
<b>F<sub>p</sub></b>	Fracción de protonación de las cadenas laterales
<b>GST</b>	Glutation-S Transfereasa
<b>HSQC</b>	“Heteronuclear Single Quantum Coherente”
<b>ITC (ó CIT)</b>	Calorimetría isotérmica de titulación
<b>IPTG</b>	Isopropil β-D-tiogalactopiranosido
<b>K<sub>a</sub></b>	Constante de equilibrio de asociación
<b>K<sub>D</sub></b>	Constante de equilibrio de desplegamiento
<b>K<sub>d</sub></b>	Constante de equilibrio de disociación
<b>K<sub>i</sub></b>	Constante de equilibrio de unión tras una inyección
<b>Kan</b>	Kanamicina
<b>LB</b>	Medio de cultivo bacteriano Laura Bertani
<b>[L]</b>	Concentración de ligando libre
<b>[L]<sub>T</sub></b>	Concentración de ligando total

<b>[L]<sub>b</sub></b>	Concentración de ligando unido
<b>[L]<sub>0</sub></b>	Concentración de ligando en la jeringa
<b>[M]<sub>T</sub></b>	Concentración de macromolécula total
<b>[M]<sub>i</sub></b>	Concentración de macromolécula total en tras cada inyección "i"
<b>M<sub>p</sub></b>	Masa molecular de la proteína
<b>MALDI-TOF</b>	"Matriz-Assisted Laser Desorption/Ionization Time-of-Flight"
<b>N</b>	Estado nativo de la proteína
<b>NOE</b>	Efecto nuclear Overhauser ("Nuclear Overhauser Effect")
<b>NOESY</b>	Espectroscopía de efecto Overhauser nuclear
<b>n</b>	Número de sitios de unión de un ligando a una proteína
<b>[P]</b>	Concentración de proteína
<b>[P]<sub>0</sub></b>	Concentración de proteína inicial
<b>PDB</b>	Banco de datos de proteínas
<b>PCR</b>	Reacción en cadena de la polimerasa
<b>P<sub>m</sub></b>	Peso molecular
<b>ppm</b>	Partes por millón
<b>q</b>	Función de partición
<b>q<sub>i</sub></b>	Calor liberado o absorbido en una inyección
<b>Q(N)</b>	Calor total acumulado en N inyecciones
<b>RMN</b>	Resonancia magnética nuclear
<b>RMSD</b>	Desviación cuadrática media "Root Mean Square Deviation"
<b>rpm</b>	Revoluciones por minuto
<b>SDS</b>	Lauril sulfato sódico
<b>SDS-PAGE</b>	Electroforesis en gel de poliacrilamida en presencia de lauril sulfato sódico
<b>T</b>	Temperatura
<b>T<sub>H</sub></b>	Temperatura de inversión de la entalpía
<b>T<sub>m</sub></b>	Temperatura de la transición a la que $\Delta G_D = 0$
<b>T<sub>r</sub></b>	Temperatura de referencia
<b>T<sub>s</sub></b>	Temperatura de máxima estabilidad
<b>TEV</b>	Virus del mosaico del tabaco ("Tobacco Etch Virus")
<b>TOCSY</b>	Espectroscopia de correlación total ("Total Correlation Spectroscopy")

<b>t</b>	tiempo
<b>u.a.</b>	Unidades arbitrarias para la intensidad de fluorescencia
<b>UV</b>	Ultravioleta
<b>V<sub>c</sub></b>	Volumen de célula del calorímetro
<b>V<sub>in</sub></b>	Volumen de inyección
<b>V<sub>p</sub></b>	Volumen específico parcial de la proteína
<b>V<sub>s</sub></b>	Volumen específico parcial del disolvente
<b>X<sub>D</sub></b>	Fracción del estado desplegado
<b>X<sub>N</sub></b>	Fracción del estado nativo
<b>ΔC<sub>p</sub></b>	Cambio de capacidad calorífica de desnaturalización a presión constante
<b>ΔC<sub>p,ap</sub></b>	Cambio de capacidad calorífica aparente de unión
<b>ΔC<sub>p,int</sub></b>	Cambio de capacidad calorífica intrínseca de unión
<b>ΔC<sub>p,p</sub></b>	Cambio de capacidad calorífica de protonación
<b>ΔG<sub>ap</sub></b>	Cambio de energía libre aparente de Gibbs de la formación del complejo
<b>ΔG<sub>D</sub></b>	Cambio de energía libre de Gibbs de desplegamiento
<b>ΔG<sub>máx</sub></b>	Cambio de energía libre de Gibbs en el máximo de la curva de estabilidad
<b>ΔH<sub>ap</sub></b>	Entalpía aparente de unión
<b>ΔH<sub>conf</sub></b>	entalpía de unión asociada a cambios conformacionales en proteína y/o ligando
<b>ΔH<sub>D</sub></b>	Entalpía de desplegamiento de una proteína
<b>ΔH<sub>p</sub></b>	Entalpía de protonación de las cadenas laterales
<b>ΔH<sub>m</sub></b>	Cambio de entalpía a la temperatura T <sub>m</sub>
<b>ΔH<sub>vH</sub> ó ΔH<sub>exp</sub></b>	Entalpía de van't Hoff ó experimental
<b>ΔH<sub>cal</sub></b>	Entalpía calorimétrica
<b>Δh(T)</b>	Función de entalpía específica promedio
<b>&lt;H&gt;</b>	Entalpía molar parcial del sistema
<b>Δm<sub>s</sub></b>	La cantidad de disolvente desplazado por la proteína en disolución dentro de la célula calorimétrica
<b>ΔS<sub>ap</sub></b>	Entropía aparente de unión
<b>ΔS<sub>D</sub></b>	Entropía de desplegamiento de una proteína

$\Delta S_m$	Cambio de entropía a la temperatura $T_m$
$\Delta n$	Índice de refracción
$\Delta \delta$	Cambio del desplazamiento químico
$\delta$	Desplazamiento químico
$\theta$	Elipticidad
$\theta_{MMR}$	Elipticidad molar media por aminoácido
$\theta_D$	Elipticidad del estado desplegado
$\theta_{D,0}$	Ordenada en el origen de la dependencia con la temperatura del estado desplegado durante la desnaturalización térmica
$\theta_N$	Elipticidad del estado nativo
$\theta_{N,0}$	Ordenada en el origen de la dependencia con la temperatura del estado nativo durante la desnaturalización térmica
$\Theta_i$	Fracción de saturación
$\chi$	Parámetro de cooperatividad
$\varpi$	Centro de masas espectral
[urea]	Concentración de urea
$\Phi$	Angulo diedro "phi"
$\Psi$	Angulo diedro "psi"

# CAPÍTULO 1



## **CAPÍTULO 1. Introducción**

### **1.1. Reconocimiento molecular y diseño molecular en la era post-genómica**

El correcto funcionamiento celular requiere el establecimiento de interacciones transitorias entre moléculas biológicas, así como una correcta regulación espacial y temporal. En conjunto, las interacciones de las proteínas en el interior celular constituyen enormes redes que son la base del metabolismo y del funcionamiento de los seres vivos. Algunos ejemplos de las interacciones transitorias fundamentales mediadas por proteínas son la unión de sustratos e inhibidores a enzimas, la identificación de antígenos por sus correspondientes anticuerpos, la interacción entre hormonas y sus receptores celulares, el establecimiento de redes de transducción de señales, etc.

Una alteración de cualquiera de estos procesos de reconocimiento molecular puede afectar al desarrollo del ciclo de vida celular y ser el origen de determinadas enfermedades humanas, como son el cáncer, el SIDA, la leucemia, la distrofia muscular, etc. Por este motivo, una profunda y detallada comprensión de estos procesos permitiría predecir y modular la respuesta fisiológica, o incluso patológica, de la célula a un estímulo específico. Además, facilitaría el desarrollo de estrategias racionales de diseño molecular que conduzcan a la obtención de nuevas moléculas de interés biomédico o biotecnológico.

En la última década hemos sido testigos de una gran revolución en los campos de la genómica y de la proteómica que ha hecho aún más patente la necesidad de profundizar en el estudio del reconocimiento molecular. La secuenciación del genoma humano, junto con el de varios organismos modelo y de agentes patógenos, ha generado un nuevo y excitante panorama científico que requiere la revisión y actualización de las metodologías tradicionales de diseño molecular. Muchas de las proteínas identificadas en estos genomas constituyen dianas potenciales para el desarrollo de nuevos agentes terapéuticos cuyas interacciones hay que caracterizar y validar. Para cada nueva diana es necesario identificar o diseñar nuevos compuestos que modulen su actividad de la forma deseada y, para ello, se requiere una racionalización del proceso de unión. Es

importante profundizar en la interrelación existente entre la estructura, la estabilidad de una proteína y sus características funcionales e incorporar al proceso de diseño no sólo aspectos estructurales, sino también consideraciones termodinámicas y dinámicas (estabilidad y cooperatividad), genéticas y funcionales, las cuales son frecuentemente ignoradas en las estrategias de diseño tradicionales (Velazquez Campoy & Freire 2005; Lafont et al. 2007; Freire 2008).

### **1.2. Interacciones de secuencias ricas en prolina con módulos de reconocimiento molecular.**

Normalmente, las interacciones transitorias entre proteínas están mediadas por pequeños dominios modulares altamente conservados. Estos dominios están presentes en distintas proteínas en número variable y en distintas combinaciones reconociendo secuencias específicas en sus dianas naturales (Pawson 1995; Pawson & Scott 1997; Pawson et al. 2002). Los dominios modulares proporcionan un mecanismo muy eficiente para el control de la actividad celular, contribuyendo a la localización celular de enzimas y sustratos dentro de cascadas de señalización específicas, así como reconociendo modificaciones post-transduccionales e incluso propiciando la intercomunicación entre distintas rutas de señalización. En general, los dominios modulares de reconocimiento de proteínas presentan entre 30 y 150 aminoácidos y se encuentran plegados en estructuras compactas y estables, caracterizadas por uno o más sitios de unión.

Cada familia de dominios estructuralmente diferente reconoce en sus proteínas diana ciertas secuencias específicas de 3 a 6 aminoácidos. Por ejemplo, los dominios SH3 (en inglés, "Src-homology region 3") reconocen secuencias ricas en prolina, los dominios SH2 ("Src-homology region 2") y los dominios PTB ("Phospho-tyrosine binding domain") se unen a secuencias peptídicas con restos de tirosina fosforilados y los dominios PDZ (acrónimo de le nombre en inglés de la proteína en que fueron descubiertos, "Post Synaptic Density protein 95") interaccionan con secuencias correspondientes a extremos carboxilo terminales (Kuriyan & Cowburn 1997; Yan et al. 2002; Jelen et al. 2003; Ball et al. 2005). En ocasiones, estas secuencias diana están flaqueadas por restos que interaccionan con elementos variables en el sitio de unión. Dichas interacciones confieren

especificidad de unión dentro de cada familia de dominios; es decir, son responsables de la selección de las proteínas naturales funcionalmente complementarias.

Las interacciones mediadas por dominios modulares presentan dos características especialmente interesantes. Por un lado, reconocen epítomos continuos, lo que simplifica la identificación a nivel de proteoma de sus dianas fisiológicamente relevantes (Brannetti & Helmer-Citterich 2003; Reiss & Schwikowski 2004; Neduva et al. 2005). Por otra parte, presentan sitios de unión cóncavos con superficies de interacción pequeñas (de unos 1000 Å<sup>2</sup>), más similares a las observadas en las interacciones de enzimas con pequeños sustratos que a las de interacciones proteína-proteína. Estas propiedades hacen de estos módulos de reconocimiento atractivas dianas para el desarrollo de nuevos agentes terapéuticos. Una profunda y detallada comprensión de las reglas que gobiernan la especificidad de unión de los distintos dominios de reconocimiento posibilita el diseño de pequeños compuestos que bloqueen con alta especificidad estas interacciones (Kardinal et al. 2000; Lee et al. 2002; Porter et al. 2007; Sudarman et al. 2008; Wang et al. 2008).

Dentro del conjunto de dominios modulares existen cinco familias de reconocimiento de secuencias ricas en prolina: dominios SH3, dominios WW (que toman su nombre de dos restos triptófano altamente conservados), dominios EVH1 (“Enabled Vasodilator-stimulated-protein Homology”), dominios GYF (denominados así debido a la presencia de la triada de restos glicina-tirosina-fenilalanina que los caracteriza) y dominios UEV (“Ubiquitin E2 Variant”) (Zarrinpar & Lim 2000; Macias et al. 2002; Ball et al. 2005; Li & Lazaridis 2005). Estas familias de dominios de reconocimiento de secuencias ricas en prolina interactúan con epítomos o secuencias consenso muy similares. Varias familias, como los dominios SH3 y WW, comparten un mecanismo de reconocimiento común basado en la inserción de las prolinas principales del ligando en bolsillos hidrofóbicos de los dominios. Incluso, se ha llegado a sugerir que puedan interactuar con los mismos epítomos o secuencias consenso (Bedford et al. 1997).

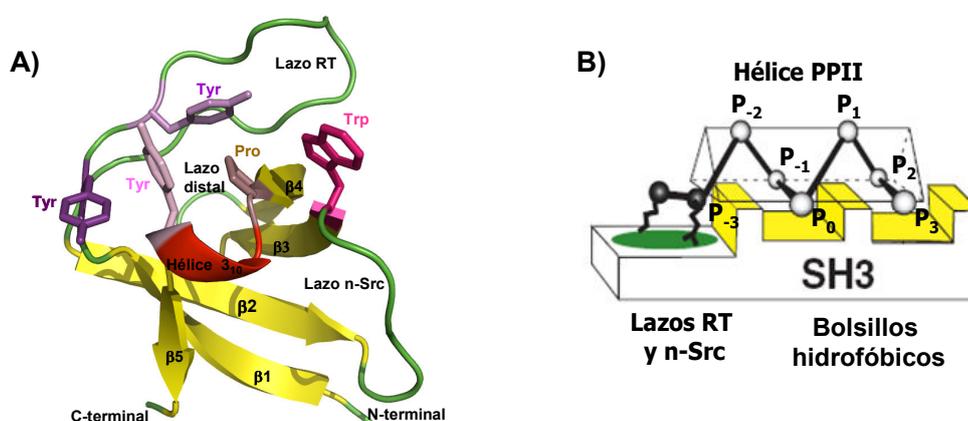
A día de hoy, se sabe que las interacciones de estos dominios con las secuencias ricas en prolina presentan afinidades moderadas o bajas, cuyas

constantes de disociación varían entre 1 y 500  $\mu\text{M}$ . Además, se postula que las divergencias entre los dominios de una misma familia son suficientes para establecer redes de interacciones de alta especificidad (Zarrinpar et al. 2003a). Sin embargo, con frecuencia se ha detectado un elevado grado de promiscuidad para alguno de estos dominios puesto que reconocen un número considerable de dianas con afinidades comparables. A la vista de esta aparente paradoja se ha propuesto la existencia de dos tipos de especificidad: una especificidad intrínseca encriptada en cada pareja dominio-ligando y una especificidad contextual en la que factores como la localización subcelular o el efecto cooperativo de interacciones múltiples desempeñan un papel importante (Ladbury & Arold 2000; Mayer 2001).

Dentro de los dominios de reconocimiento de secuencias ricas en prolina, los dominios SH3 han sido ampliamente estudiados mientras que se sabe relativamente poco de las demás familias. Los dominios SH3 son, con diferencia, los más abundantes en vertebrados, con un número estimado de 409 dominios en el proteoma humano (Zarrinpar & Lim 2000; Mayer 2001; Macias et al. 2002). Estos dominios forman parte de proteínas muy diversas (Stein 1998; McPherson 1999) y frecuentemente actúan como sitios de anclaje para el reconocimiento de sustratos conduciendo a la modificación enzimática y la regulación de las proteínas que los contienen (Barila & Superti-Furga 1998; Arold et al. 2001; Brasher et al. 2001; Brabek et al. 2002). En ocasiones, se han asociado a patologías tales como procesos cancerígenos (c-Src y c-Yes), el SIDA (Lyn, Hck, Lck), la leucemia (Abl), la osteoporosis (Src), o procesos inflamatorios, alérgicos y asmáticos (Dalgarno et al. 1997; Skorski et al. 1998).

Actualmente se dispone de una gran cantidad de información estructural sobre los dominios SH3, ya que existen más de 100 estructuras resueltas, tanto por RMN como por difracción de rayos X. Los dominios SH3 presentan un plegamiento característico que consiste en una estructura constituida por 5 hebras  $\beta$  antiparalelas conectadas por 3 lazos de longitud variable (RT, n-Src y distal) y una hélice  $3_{10}$ , como se observa en el Panel A de la Figura 1.3. Por otro lado, además, se ha caracterizado exhaustivamente su equilibrio conformacional desde un punto de vista termodinámico (Viguera et al. 1994; Martínez et al. 1999; Sadqi et al. 2002; Martín-Sierra et al. 2003) y cinético (Martínez et al. 1998; Martínez &

Serrano 1999; Korzhnev et al. 2004; Candell et al. 2009). Los dominios SH3 reconocen ligandos que contienen la secuencia xPxxP (donde la x suele ser un resto hidrofóbico, y la P, un resto prolina) y adoptan una conformación de hélice poliprolina tipo II (PPII). Las superficies de unión de los dominios SH3 están formadas por varios aminoácidos aromáticos muy conservados (Panel A, Figura 1.1) dispuestos tridimensionalmente en tres bolsillos: dos bolsillos hidrofóbicos y un bolsillo de especificidad (Panel B, Figura 1.1). Los dos bolsillos hidrofóbicos acomodan dipéptidos del ligando tipo xP. En cambio, el bolsillo de especificidad está acotado por los lazos RT y n-Src, los cuales son flexibles y variables en secuencia (Zarrinpar et al. 2003b; Ball et al. 2005).



**Figura 1.1. Reconocimiento de ligandos ricos en prolina por dominios SH3.** **A)** Representación de la estructura tridimensional del dominio SH3 de *c-Yes* determinada por rayos X (código PDB: 2HDA). Las cinco hebras beta se representan como  $\beta 1$ ,  $\beta 2$ ,  $\beta 3$ ,  $\beta 4$  y  $\beta 5$  en amarillo y la hélice  $3_{10}$  que conecta las hebras  $\beta 4$  y  $\beta 5$  aparece en rojo. El lazo n-Src que conecta las hebras  $\beta 2$  y  $\beta 3$ , el lazo RT que conecta las hebras  $\beta 1$  y  $\beta 2$  y el lazo distal que conecta las hebras  $\beta 3$  y  $\beta 4$  se indican en color verde. Los restos aromáticos más conservados que son importantes para la unión de los ligandos se indican en representación de varillas. **B)** Representación esquemática del sitio de unión del dominio SH3, modelado con un ligando tipo PXXP en orientación tipo I. El sitio de unión contiene dos bolsillos hidrofóbicos xP formados por restos aromáticos conservados (amarillo) y un bolsillo de especificidad adyacente que es más variable (verde). Las letras P en el ligando indican las posiciones en su secuencia consenso,  $P_0X_1X_2P_3$ .

Las interacciones de los dominios SH3 se basan fundamentalmente en la intercalación de los restos prolina del ligando entre aminoácidos hidrofóbicos del sitio de unión del dominio, con muy pocas interacciones polares, por lo que cabe esperar un patrón termodinámico dirigido por el efecto hidrofóbico (con una entropía de unión favorable). Sin embargo, los estudios termodinámicos sobre las

interacciones mediadas por los dominios SH3 con secuencias ricas en prolina presentan un patrón termodinámico dominado por una entalpía de unión favorable a la que se opone una contribución entrópica desfavorable (Renzoni et al. 1996; Arold et al. 1998; Wang et al. 2001b; Ferreon & Hilser 2004; Palencia et al. 2004). Este comportamiento termodinámico no se puede racionalizar fácilmente considerando exclusivamente las interacciones hidrofóbicas. Algunos trabajos proponen que este comportamiento termodinámico está asociado a efectos dinámicos que implican una redistribución conformacional en los lazos de los dominios SH3 (Wang et al. 2001a) o en el ligando (Ferreon & Hilser 2004); mientras que otros trabajos realizados en nuestro laboratorio sugieren la presencia de moléculas de agua enterradas, mediando una compleja red de puentes de hidrógeno, en la interfase de unión entre el ligando y el dominio SH3 (Palencia et al. 2004; Palencia et al. 2010).

### **1.3. Los dominios WW**

Los dominios WW son dominios modulares presentes en numerosas proteínas (más de 125 en el genoma humano) tanto en el citoplasma como en el núcleo celular (Hu et al. 2004). La unión de los dominios WW a sus dianas naturales está relacionada con funciones muy variadas. Entre otras funciones, se hallan involucrados en la transcripción de ADN, en el procesamiento de ARN, en el reconocimiento de ubiquitina, en el tráfico de proteínas y de receptores de señalización y en el control del citoesqueleto (Kay et al. 2000; Sudol & Hunter 2000; Sudol et al. 2001; Macias et al. 2002). Por tanto, como ocurre con los dominios SH3, algunas alteraciones de las funciones de los dominios WW se han relacionado con varias enfermedades humanas como el síndrome de hipertensión de Lyddle, la distrofia muscular, procesos cancerígenos y las enfermedades de Alzheimer y de Huntington (Pereboev et al. 2001; Sze et al. 2004; Rotin & Schild 2008; Bellomaria et al. 2010). Además, algunos dominios WW están implicados en el desarrollo de infecciones víricas, como el Ébola o la leucemia humana (Bouamr et al. 2003; Timmins et al. 2003). Por todo ello, estos dominios también han despertado gran interés como potenciales dianas para el desarrollo de nuevos agentes terapéuticos.

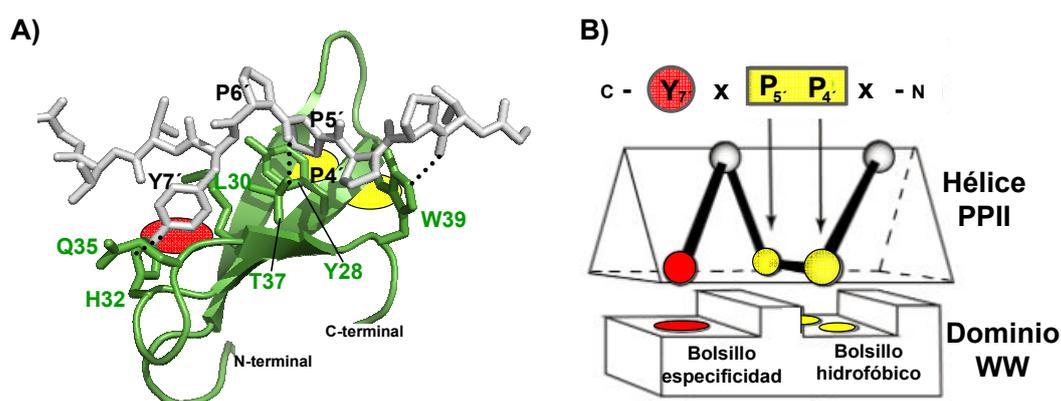
### 1.3.1. Estructura y clasificación de dominios WW

Al igual que los dominios SH3, los dominios WW están caracterizados por una estructura común y relativamente compacta. En este caso, la estructura es una sola lámina  $\beta$  formada por tres hebras antiparalelas retorcidas, unidas por dos lazos cuya extensión es variable. Estos dominios contienen aproximadamente 40 aminoácidos y son considerados los dominios modulares de proteínas naturales más pequeños que existen con un plegamiento simple en lámina  $\beta$ . Sin embargo, la información estructural disponible para los dominios WW es mucho más limitada que para los dominios SH3. Actualmente, existen unas 25 estructuras determinadas por RMN de complejos y de dominios WW libres y muy pocas estructuras cristalográficas.

Todos los dominios WW contienen dos cavidades en el sitio de unión: un bolsillo xP y un bolsillo de especificidad. En el Panel A de la Figura 1.2 se muestra, a modo de ejemplo, un complejo formado por un dominio WW y su correspondiente diana natural. En el Panel B de la misma figura se representa un esquema del sitio de unión de dicho complejo y se puede observar claramente que el bolsillo xP tiene algunas similitudes con el de los dominios SH3 (visto anteriormente en la Figura 1.1B). La principal similitud es la presencia de restos aromáticos conservados y muy parecidos en la segunda y tercera hebras del dominio WW (como son la tirosina Y28 y el triptófano W39, en la Figura 1.2A) conformando el bolsillo de interacción hidrofóbico donde se intercalan dipéptidos del tipo xP del ligando (en el ejemplo de la Figura 1.2A, las prolina P4' y P5'). El bolsillo de especificidad se localiza en el lazo que conecta la segunda y tercera hebras, y varía de un dominio WW a otro dependiendo del tipo de ligandos que reconozca (Sudol et al. 1995; Zarrinpar & Lim 2000; Zarrinpar et al. 2003a).

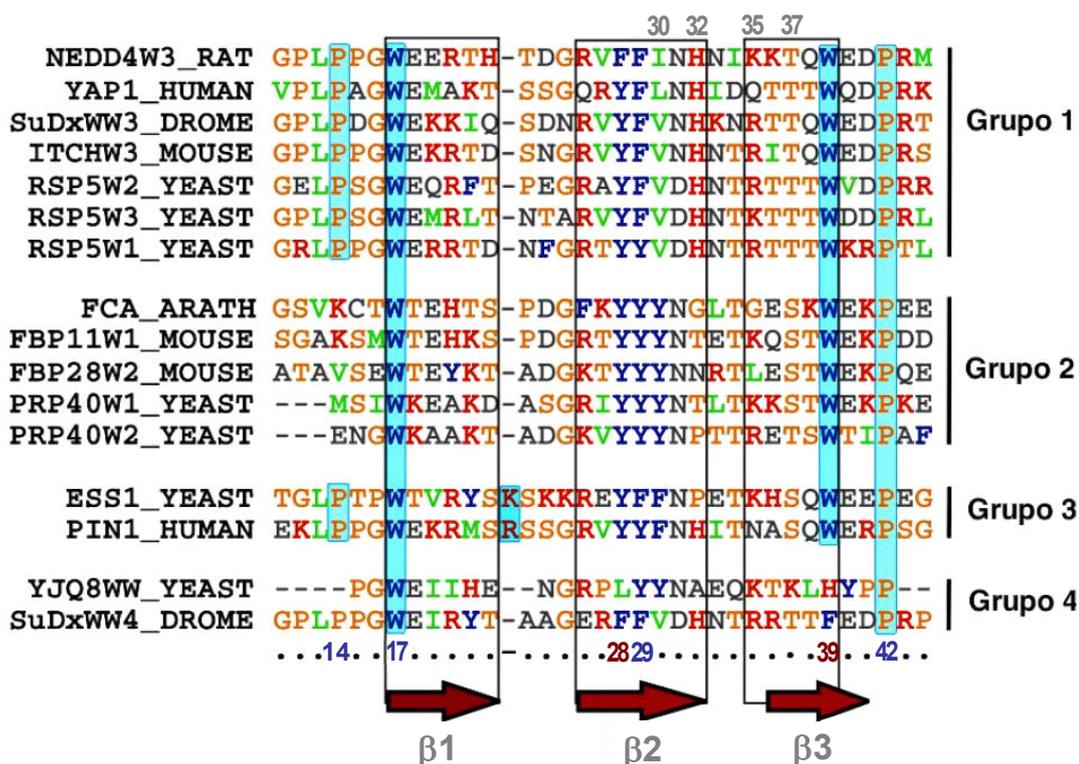
En la Figura 1.3 se muestra un alineamiento de las secuencias de diferentes dominios WW donde se pueden reconocer algunas de las características más importantes a nivel de estructura primaria. La numeración utilizada en el alineamiento de esta figura constituye la base de las futuras referencias para las secuencias de dominios WW en este trabajo. El nombre de estos dominios, "WW", deriva de dos restos triptófano muy conservados (W17 y W39) que aparecen espaciados entre 20 y 23 restos en sus secuencias. Estos dos

triptófanos junto a la prolina P42, configuran la principal señal de identidad de los dominios WW y se encuentran muy conservados en todos ellos. El W17, la P42 y el resto aromático en la posición 29 (en la figura en azul), están involucrados en la formación de un broche o hebilla de tipo hidrofóbico responsable de mantener la integridad del plegamiento de los dominios (Macias et al. 1996; Koepf et al. 1999a). Las siguientes posiciones más conservadas en los dominios WW, 28 y 39, están ocupadas por los restos aromáticos que conforman el bolsillo xP (en la figura, en granate). De hecho, el resto ubicado en la posición 39 suele ser el segundo triptófano que les da nombre a estos dominios, aunque en algunos casos puede ser otro residuo aromático como una fenilalanina o una tirosina (Sudol et al. 1995; Zarrinpar & Lim 2000; Zarrinpar et al. 2003a).



**Figura 1.2. Reconocimiento de ligandos ricos en prolina por dominios WW.** A) Representación esquemática de cintas de la estructura determinadas por RMN del dominio hYap-WW1 (L30K) formando complejo con el péptido GPPPYTVG (PDB: 1JMQ). El dominio WW está representado en color verde y las cadenas laterales de los restos que forman bolsillos en el sitio de unión Y28, L30, H32, Q35, T37 y W39 se han representado como varillas en verde. El péptido se ha representado como varillas de color gris. B) Representación esquemática del sitio de unión típico de un dominio WW con un ligando PPxY con la estructura de hélice PPII que caracteriza a estos ligandos. En ambos paneles, la región correspondiente al bolsillo xP se ha resaltado como círculos amarillos y la región correspondiente al bolsillo de especificidad como un círculo de color rojo.

Los dominios WW se agrupan en cuatro subfamilias o clases atendiendo a criterios relacionados con su secuencia (Figura 1.3) y con la especificidad de unión por los ligandos ricos en prolina que reconocen (Macias et al. 2000; Schleinkofer et al. 2004). Los dominios WW de clase I presentan un claro patrón de secuencia PWWP que hace referencia a los restos P14, W17, W39 y P42, y además reconocen la secuencia consenso PPxY. En la clase II, los dominios pierden la P14



**Figura 1.3. Clasificación de los dominios WW en los diferentes grupos de especificidad según el tipo principal de las secuencias ricas en prolina que reconocen.** Alineamiento de secuencias de una selección de dominios WW en la que se pueden observar, resaltados con un fondo azul celeste, los residuos conservados en cada tipo de dominio. Las posiciones implicadas en la formación del broche hidrofóbico de estabilidad se muestran en azul, las posiciones implicadas en la formación del bolsillo xP en granate y las posiciones implicadas en la formación del bolsillo de especificidad xY de los dominios WW de tipo 1 se muestran en gris. Los colores de los residuos son los que utiliza por defecto ClustalX e indican propiedades similares (aromático, positivo, apolar, etc.).

respecto del patrón de la clase I (patrón de secuencia WWP), interaccionan con secuencias consenso PPLP muy parecidas a las características de dominios SH3 y en este caso, el bolsillo de especificidad también es un bolsillo hidrofóbico de tipo xP (denominado bolsillo xP2). La clase III se caracteriza por el patrón PWWP más una inserción en su secuencia al final de la primera hebra  $\beta 1$  mientras que la secuencia consenso que reconocen contiene una prolina precedida por fosfoserina o fosfotreonina, del tipo p(S/T)P. Por último, la clase IV es la subfamilia más heterogénea donde, independientemente de los ligandos que puedan reconocer, se agrupan los dominios que tienen las secuencias más diferentes incluyendo aquellos que no conservan el W39 que forma el bolsillo xP. Aparte de los criterios de clasificación que acabamos de exponer, hay otros, relacionados fundamentalmente con la especificidad de unión, que hacen que actualmente la

clasificación de los dominios WW siga siendo motivo de controversia. Algunos autores proponen que el motivo de unión de la clase II debe ser definido como PPxPP (Macias et al. 2002), otros han sugerido un quinto posible grupo para algunos dominios que reconocen secuencias flanqueadas o interrumpidas por los aminoácidos arginina o lisina, del tipo R/LPPP(R) (Bedford et al. 1998; Bedford et al. 2000), o incluso para aquellos que puedan unir a tirosinas fosforiladas en la secuencia consenso (Otte et al. 2003).

Los dominios de la clase I son los dominios WW más abundantes, y a este grupo pertenecen todos los dominios empleados en los estudios de interacción presentados en este trabajo. El bolsillo de especificidad para los dominios WW de la clase I se ha denominado bolsillo xY por ser donde se acomoda la tirosina del motivo PPxY. Esta cavidad esta formada por restos situados al final de la segunda hebra y en la mitad de la tercera hebra. Estos son un aminoácido alifático en posición 30 (leucina, valina o isoleucina), un resto histidina en posición 32, un residuo hidrofílico en la posición 35 (glutámico, lisina o arginina) y por una treonina en posición 37. Análisis mutacionales de los residuos semiconservados en las posiciones 30 y 35 demuestran que la sustitución de éstos no afecta significativamente a la interacción (Toepert et al. 2001). En cambio, los aminoácidos estrictamente conservados H32 y T37, resultan fundamentales para la unión de dianas PPxY (Kasanov et al. 2001; Toepert et al. 2001). En la estructura del complejo mostrada en la Figura 1.2A se ilustra cómo la H32 interacciona con la tirosina del ligando, mientras que el hidroxilo de la cadena lateral de T37 establece un puente de hidrógeno con una de las prolinas del motivo de unión (Pires et al. 2001; Ball et al. 2005).

### 1.3.2. Plegamiento y estabilidad de dominios WW

Teniendo en cuenta el pequeño tamaño de los dominios WW (40 aminoácidos) y que constituyen una de las estructuras con secundaria en lámina  $\beta$  más extendidas y más sencillas de la naturaleza, los dominios WW son candidatos idóneos para abordar el análisis experimental y computacional de las distintas fuerzas que gobiernan el plegamiento y la estabilidad de este tipo de estructura secundaria (Ibragimova & Wade 1999; Koepf et al. 1999a; Koepf et al.

1999b; Crane et al. 2000; Jager et al. 2001; Nguyen et al. 2003; Deechongkit et al. 2004; Petrovich et al. 2006; Liu et al. 2009). Además, hay que sumar otras características que facilitan este tipo de estudios, como son su elevada solubilidad en disoluciones acuosas, la definición de un estado nativo monomérico, la ausencia de cofactores, grupos prostéticos, puentes disulfuro, prolinas en conformación cis o la elevada tolerancia a mutaciones sin que se desestabilice su estructura.

Actualmente, los estudios de plegamiento de dominios WW son escasos y, en general, se limitan a experimentos puntuales realizados mediante técnicas espectroscópicas, que proporcionan una información muy limitada. La primera caracterización termodinámica del equilibrio conformacional de un dominio WW fue realizada con el primer dominio WW de la proteína humana hYap. El estudio se llevó a cabo mediante el análisis de las desnaturalizaciones térmica y química y mostró el carácter monomérico y la elevada reversibilidad del estado nativo en la mayoría de las condiciones experimentales estudiadas (Koepp et al. 1999a; Koepp et al. 1999b). Los autores sugirieron posibles roles tanto estructurales como funcionales de los dos triptófanos altamente conservados que dan nombre a estos dominios. Tanto en este trabajo como en estudios posteriores sobre los dominios WW de las proteínas hPin1 (Karanicolas & Brooks 2004; Sharpe et al. 2007; Jager et al. 2008) y Fbp28 (Karanicolas & Brooks 2003; Petrovich et al. 2006; Sharpe et al. 2007), el equilibrio conformacional de los dominios WW se ha descrito de acuerdo al modelo de equilibrio de dos estados. No obstante, como ya hemos indicado, se trata de experimentos puntuales seguidos por técnicas espectroscópicas, en general poco informativas en cuanto a la cooperatividad del equilibrio de plegamiento.

A diferencia de los estudios termodinámicos, las caracterizaciones cinéticas del plegamiento llevadas a cabo con distintos dominios WW son más numerosas y completas, aunque describen diferentes comportamientos para los dominios WW (Crane et al. 2000; Ferguson et al. 2001a; Ferguson et al. 2001b; Jager et al. 2001; Petrovich et al. 2006; Sharpe et al. 2007). En general, estos estudios indican que el plegamiento es un proceso rápido que corresponde a una transición entre dos estados a través de un estado de transición compacto con la presencia de un

núcleo de plegamiento localizado en el lazo que conecta las hebras  $\beta 1$  y  $\beta 2$  (Jager et al. 2001; Petrovich et al. 2006). Otros trabajos sugieren que el núcleo de plegamiento es difuso y dependiente de la temperatura, y proponen que el estado de transición se estabiliza por interacciones hidrofóbicas asociadas a la formación de las zonas centrales de las hebras  $\beta$  (Ibragimova & Wade 1999; Karanicolas & Brooks 2004; Sharpe et al. 2007; Jager et al. 2008). Por ejemplo, la cinética de plegamiento del dominio WW de la proteína Fbp28, indica un complejo comportamiento que varía de un mecanismo de dos estados a un mecanismo de tres estados dependiendo de la temperatura (Karanicolas & Brooks 2003; Nguyen et al. 2003).

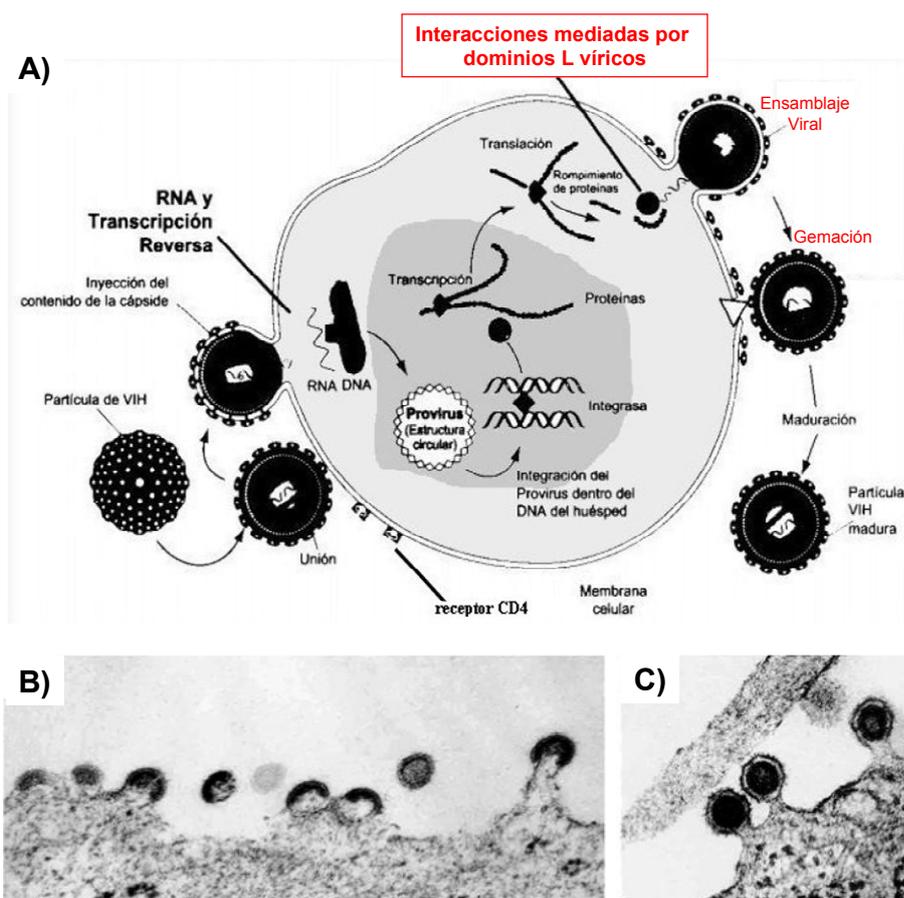
En vista del enorme interés que ha suscitado el plegamiento de pequeños dominios modulares y debido a la escasa información disponible sobre la termodinámica del plegamiento de dominios WW, uno de los objetivos que se plantean en esta Tesis es abordar por primera vez la caracterización termodinámica detallada y completa del equilibrio conformacional de algunos dominios WW utilizando diversas técnicas y metodologías de análisis.

### **1.4. Los dominios WW como objeto del diseño racional de inhibidores**

Con el objetivo de profundizar en la comprensión de los procesos de reconocimiento de secuencias ricas en prolinas por los dominios WW, a continuación se proponen dos situaciones en las que el estudio de estas interacciones es especialmente interesante, como son: el desarrollo de enfermedades de origen vírico y la regulación de la expresión de genes mediada por la proteína humana asociada a Yes (hYap).

#### **1.4.1. Dominios L víricos como dianas para el diseño de antivirales de amplio espectro**

La etapa final en el ciclo de vida de los virus encapsulados consiste en un proceso de fusión de la membrana de la célula huésped que tiene como resultado la liberación del virus (Figura 1.4). En la última década se han identificado pequeñas secuencias ricas en prolina que desempeñan un papel crítico en el



**Figura 1.4. Ejemplo de las etapas del ciclo de vida de los virus.** A) Representación simplificada del ciclo de vida del retrovirus VIH donde se representan las etapas más importantes, como la unión del virus a la membrana celular, la inyección de su ARN en la célula, la transcripción reversa del ARN en ADN, la integración de éste en el núcleo celular, su transcripción y translación, y, finalmente el ensamblaje en la pared celular y gemación necesarios para escapar de la célula infectada. Con letras rojas se han resaltado los procesos en los que participan los dominios L víricos. B) Proceso de formación de vesículas conteniendo las partículas virales y transporte de éstas hasta la pared de la membrana celular de la célula infectada. C) Proceso de formación de invaginaciones en la pared de la membrana celular y escape de las partículas virales de la célula. Imágenes tomadas de Freed 2002; Demirov & Freed 2004.

proceso de liberación del virus. Estas secuencias se denominan dominios L (“Late assembly domains”) y se hallan en diversas proteínas víricas como las poliproteínas Gag de los retrovirus o las proteínas de matriz de rhabdovirus y filovirus. Se ha demostrado que la eliminación de estos dominios o la alteración de sus secuencias bloquea la liberación del virus, impidiendo la propagación de las partículas infecciosas (Gottlinger et al. 1991; Yasuda & Hunter 1998; Yuan et al. 2000; Freed 2002; Yasuda et al. 2003). En la Figura 1.5 se muestran algunos ejemplos de los tres tipos de dominios L identificados como tres motivos muy

Familia	Virus	Secuencia
Retrovirus	HTLV1	----- P P P S S P T H D P P D S D P Q I P P P Y V E P T A P Q V L P
	HTLV2	V S P S A P A A P V P T P I C P T T T P P P P P P S P E A H V P P P Y V E P T T T Q C F P
	BLV	----- P G A S A P E E Q P P P Y D P P A I L P I I S
	GALV	- - - P V A G R P S A P P R P P I Y P A T D D L L L L S E P T P P P Y P A A L P P P L A P
	MLV	L Y P A L T P S L G A K P K P Q V L S D S G G P L I D L L T E D P P P Y R D P R P P P S D
	RSV	- P E E T A T P K T V G T S C Y H C G T A I G C N C A T A S A P P P P Y V G S G L Y P S L A
	MPMV	----- D G L E D E A A K Y H N P D W P P F L T R P P P Y N K A T P S A P T V
Rhabdovirus	HIV1	----- S R P E P T A P P E E S F R S G V E T T T P P Q K Q E P I D K E L Y P
	VSIV	----- K G K G K K S K K L G I A P P P Y E E D T S M E Y A P
Filovirus	RABV	----- R R D E D T Q K S S P A S A P L D D D D L W L P P P E Y V P L K E L T G K K
	ZEBOV	----- M R R V I L P T A P P E Y M E A I Y P V R S N
	MARV	----- M A S S S N Y N T Y M Q Y L N P P P Y A D H G A N Q L I P
Herpesvirus	HHV4	----- T P T P P N D E E R E S N E E P P P P Y E D P Y W G N G D R
Arenavirus	LASV	N R C P I C K M P L P T K L R P S A A P T A P P T G A A D S I R P P P Y S P

**Figura 1.5. Dominios L víricos.** Alineamiento múltiple de secuencias procedentes de distintas familias de virus que contienen una o varias secuencias denominadas dominios L. Las secuencias corresponden a los siguientes virus: leucemia humana tipo I (HTLV1), leucemia humana tipo II (HTLV2), leucemia bovina (BLV), leucemia de chimpancé (GALV), sarcoma murino de molones (MLV), sarcoma de Rous (RSV), Manson-Pfeifer de mono (MPMV), inmunodeficiencia humana tipo I (HIV1), estomatitis vesicular (VSIV), rabia (RABV), ébola del Zaire (ZEBOV), marburgovirus del Lago Victoria (MARV), herpes humano ó Epstein Barr (HHV4) o Lassavirus (LASV). El alineamiento se ha realizado centrandolo las distintas secuencias en el motivo PPxY que además se ha recuadrado. En cuadros rojos se han señalado algunos residuos parcialmente conservados. Los demás motivos típicos de dominios L, con secuencias PT/SAP o YxxL, aparecen subrayados.

conservados: PPxY, P(T/S)AP e YPDL. La amplia distribución de estas secuencias entre distintos tipos de virus, junto con el hecho de que estos motivos en ocasiones son intercambiables, ha llevado a plantear una estrategia común de los virus para su liberación de la célula huésped (Parent et al. 1995; Puffer et al. 1998; Chen et al. 2001; Martin-Serrano et al. 2001). Es importante mencionar que los tres motivos característicos de dominios L se han relacionado con las secuencias consenso propias de módulos de reconocimiento proteína-proteína: dominios SH3 (PxxP) (Mayer & Eck 1995; Dalgarno et al. 1997), dominios WW (PxxP ó PPxY) (Sudol 1996b; Sudol 1996a) y proteínas adaptadoras AP1 y AP2 (YxxL) (Ohno et al. 1995; Puffer et al. 1998).

Actualmente, hay muchas evidencias experimentales que indican que los dominios L no actúan directamente en la morfogénesis del virión, sino que su principal función es reclutar la maquinaria celular necesaria para el ensamblaje del virus y para su liberación. De hecho, está bien establecido que este proceso tiene lugar mediante la interacción directa de los dominios L víricos con proteínas celulares implicadas en los mecanismos de creación y gestión de vesículas (Figura 1.4). Los dominios de tipo PTAP, presentes en virus como VIH-1 o Ébola (Figura 1.5), interaccionan con el dominio UEV de la proteína Tsg101 (“Tumor

Susceptibility Gene 101”) (VerPlank et al. 2001; Bouamr et al. 2003); mientras que los dominios de tipo PPxY, presentes en los virus del Ébola, de Mason-Pfizer, del sarcoma Rous y otros onco-retrovirus como el virus de la leucemia humana en linfocitos T de tipo I (HTLV1), interaccionan con la ligasa de ubiquitina Nedd4 (“*Neural precursor cell Expressed Developmentally Down-regulated*”), y más concretamente con su tercer dominio WW (Harty et al. 2000; Harty et al. 2001; Freed 2002; Yasuda et al. 2003; Demirov & Freed 2004; Sakurai et al. 2004).

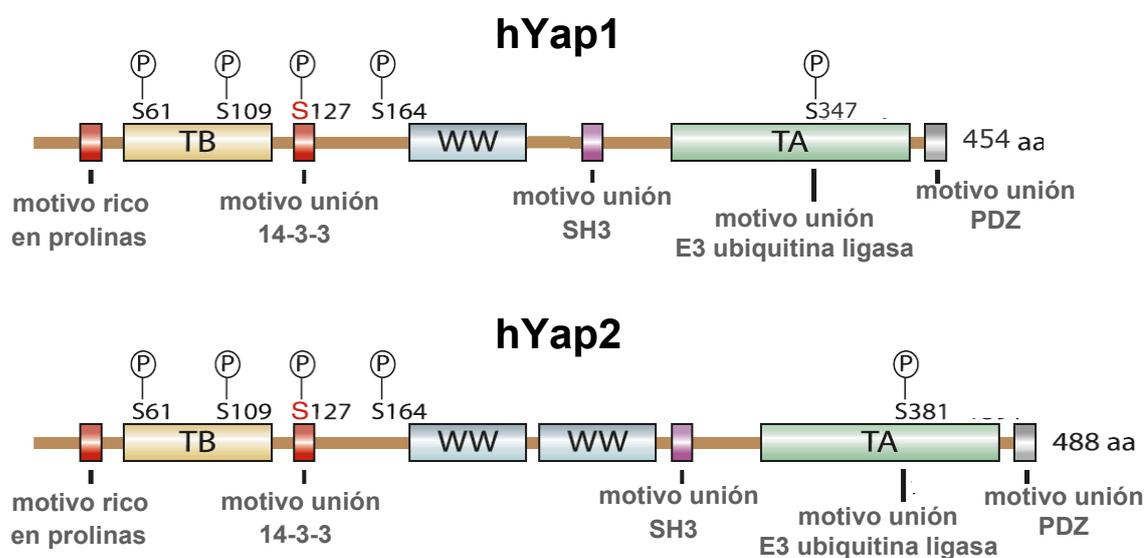
La ligasa de ubiquitina Nedd4 pertenece a la familia de ligasas tipo E3-HECT (“Homologous E6-AP C-terminus”), que se caracterizan por ser proteínas modulares compuestas por un dominio C2 en el extremo amino-terminal, entre uno y cuatro dominios WW en la región central y un dominio HECT en el extremo carboxilo-terminal (Harvey & Kumar 1999; Ingham et al. 2004). El dominio C2 es un dominio de unión a lípidos que media la localización de proteínas en membrana. Este dominio sugiere que la ligasa Nedd4 se encarga de marcar proteínas de membrana con ubiquitina, aunque se ha observado que también realiza el etiquetado de proteínas citoplasmáticas y nucleares. Nedd4 está implicada en procesos de degradación proteosómica, transporte de proteínas desde la región trans-Golgi a la membrana celular o en la internalización de proteínas de membrana para su posterior reciclaje o degradación lisosómica. Sin embargo, aunque se ha puesto de manifiesto que la presencia de algunas ligasas es imprescindible para la correcta salida de las partículas virales maduras de la célula huésped (Huang et al. 1995; Yasuda & Hunter 1998; Yuan et al. 2000; Martin-Serrano et al. 2004), no está del todo claro su papel en el final del ciclo viral (Garrus et al. 2001; Polo & Dubensky 2002). En este sentido, recientemente se ha demostrado que el virus HTLV1, que presenta dos dominios L consecutivos (PPPY+PTAP) en su poliproteína Gag, se une inicialmente, a través del motivo PPPY, a uno de los dominios WW de Nedd4 en la membrana plasmática. Posteriormente Nedd4 cataliza el etiquetado de Gag con ubiquitina y, finalmente, la proteína Gag ubiquitinada interacciona por medio del otro dominio L, PTAP con el dominio UEV de Tsg101, induciendo la gemación de la membrana celular (Blot et al. 2004).

Es particularmente interesante que organismos sencillos como los virus se hayan basado en la especificidad intrínseca de estos pequeños epítomos para reconocer y secuestrar la maquinaria celular necesaria para completar su ciclo vital. Por este motivo, el diseño de moléculas con capacidad para bloquear o modular las interacciones entre los dominios L y sus dianas se ha consolidado como una estrategia viable en el desarrollo de nuevos fármacos específicos para el tratamiento de numerosas enfermedades y antivirales de amplio espectro (Stauffer et al. 1997; Kardinal et al. 2000; Lee et al. 2002; Oneyama et al. 2002; Okumura et al. 2008). Así, otras moléculas de pequeño tamaño diseñadas para que bloqueen las interacciones víricas, podrían actuar como efectivos agentes terapéuticos contra patógenos tan letales como el VIH o el Ébola. Dado que el diseño eficaz de estos inhibidores requiere de la profunda comprensión de las fuerzas que dirigen las interacciones entre los dominios L y sus dianas, nos hemos propuesto, dentro de los objetivos de esta Tesis, caracterizar termodinámicamente y estructuralmente la interacción entre el tercer dominio WW de la proteína humana Nedd4 y un conjunto de dominios L. Este estudio constituye, además, uno de los primeros análisis termodinámicos exhaustivos de la interacción de dominios WW con sus ligandos.

### **1.4.2. Las funciones celulares de la proteína hYap, su regulación y el tándem de dominios WW**

El entramado de rutas de señalización celular se presenta como una red tridimensional de ramificaciones interconectada con otras vías de señalización en un proceso global más complejo como es el correcto funcionamiento del organismo. Algunos autores han descrito el panorama actual de la señalización celular como una gran tela de araña, en la que existen algunos nudos que son ocupados por proteínas responsables de la regulación de distintos procesos celulares (Bertini et al. 2009). La Proteína Humana Asociada a Yes (hYap) se encuentra en uno de los nudos de esta tela de araña y se ha descrito como pieza clave en diferentes vías de señalización que conducen a los más extremos y diferentes resultados para la célula, desde la proliferación o la transformación, hasta procesos opuestos, como son la muerte celular o la supresión de tumores.

Inicialmente la proteína hYap se descubrió como una proteína modular de 65 KDa (de ahí, su denominación original Yap65) que se asociaba por medio de una secuencia rica en prolinas al dominio SH3 de la quinasa de tirosina Yes. Fue en estos primeros trabajos con hYap donde se identificó y caracterizó por primera vez un dominio WW (Bork & Sudol 1994; Sudol 1994). Actualmente, se conocen varias isoformas de la proteína humana hYap y las dos más abundantes únicamente difieren entre si en la cantidad de dominios WW que contienen (Komuro et al. 2003). En la Figura 1.6 se muestra una representación esquemática de la arquitectura modular de las dos isoformas de la proteína hYap. La isoforma hYap1 corresponde a la forma originariamente descrita como Yap65 y contiene un sólo dominio WW, mientras que la otra, hYap2, contiene dos dominios WW. (Yagi et al. 1999). Además de las diferencias estructurales, también se ha comprobado que hYap2 es la forma mayoritaria en humanos y posee una mayor actividad transcriptor de genes (Komuro et al. 2003).



**Figura 1.6. Isoformas de la Proteína humana Asociada a Yes (hYap).** Representación esquemática de los dominios modulares en los que se organiza la proteína hYap. Arriba se muestra el esquema de la isoforma denominada hYap1 (454 aminoácidos), y abajo la de hYap2 (488 aminoácidos). Los módulos comunes a ambas isoformas son: una secuencia rica en prolinas, un dominio de unión a factores de transcripción (TB), un motivo de unión a la proteína 14-3-3, uno o dos dominios WW, un motivo de unión a un dominio SH3, un dominio transactivador de genes (TA), una secuencia de reconocimiento a una ubiquitina ligasa de la clase E3 y un motivo de unión a un dominio PDZ. Los óvalos con una P dentro indican las posiciones fosforilables correspondientes a los restos serina, con dianas del tipo HxRxx<sub>p</sub>(S/T), en la secuencia de la proteína hYap.

En la Figura 1.6 se observa que ambas isoformas de hYap presentan otros dominios o regiones: en el extremo N-terminal hay un sitio de unión a la proteína 14-3-3 (Rxx<sub>p</sub>SxP) y un dominio potenciador de factores de transcripción denominado dominio TB (“Transcription factor Binding domain”); en la zona central, ambas isoformas también presentan un motivo de unión a dominios SH3; en el extremo C-terminal, se encuentran un dominio transactivador de genes (TA) y un sitio de unión a dominios PDZ. Cada uno de estos dominios o sitios de unión presenta una función concreta. Así, el sitio de unión a 14-3-3 está relacionado con la regulación de la función de hYap (Kanai et al. 2000; Basu et al. 2003). El dominio TB reconoce factores de transcripción de la familia (TEF)/TEAD (Vassilev et al. 2001; Mahoney et al. 2005). En el caso del dominio TA, se ha sugerido que éste interacciona con algunas proteínas de membrana como las de la familia NHERF (“Na/H Exchanger Regulatory Factor”) (Kanai et al. 2000; Mohler et al. 2000) o como el receptor de la tirosina quinasa ErbB4 (Komuro et al. 2003; Omerovic et al. 2004). El sitio de unión a dominios PDZ interviene en la formación de un complejo con la proteína Zona Ocludens 2, por lo que es responsable de la localización de hYap en el núcleo celular (Oka et al. 2010).

Hoy en día no se tiene una comprensión total de todas las interacciones en las que interviene hYap. No obstante, se ha descrito su participación en diversas rutas de señalización y, como hemos mencionado, se puede intuir que muchas de ellas se hallan interconectadas por hYap. Además, esta proteína parece tener la habilidad de viajar del citosol al núcleo actuando sobre diferentes proteínas o expresando distintos genes dependiendo de su localización y de los factores que la regulan (Oka et al. 2008; Fernandez et al. 2009; Zhang et al. 2009). Este ir y venir de hYap entre el núcleo y el citoplasma se ha propuesto como un mecanismo importante para la regulación de su actividad. Normalmente, la forma inactiva de hYap se encuentra retenida en el citosol mediante la unión a la proteína 14-3-3. A su vez, esta interacción es debida a una modificación post-transduccional que consiste en la fosforilación de la serina en la posición 127 (S127) de hYap (Strano et al. 2001; Basu et al. 2003; Strano et al. 2005). Recientemente, se ha establecido que dicha S127 de hYap es una diana del tipo HxRxxS/T para las serina/treonina quinasas que pertenecen al subgrupo de la quinasas Ndr (“nuclear Dbf2-related”)

(Mah et al. 2005). En humanos se ha comprobado que dos de estas quinasas, Last1 (“Large tumor suppressor homolog 1”) y Lats2 (“Large tumor suppressor homolog 2”), son capaces de fosforilar S127 *in vivo* y se ha demostrado que son responsables de la unión de hYap a la proteína 14-3-3 (Zhao et al. 2007; Hao et al. 2008; Oka et al. 2008). Además de la diana de fosforilación de serinas, se han encontrado evidencias de que Lats1 interacciona por medio de una secuencia PPxY con los dominios WW de hYap2 (Tao et al. 1999).

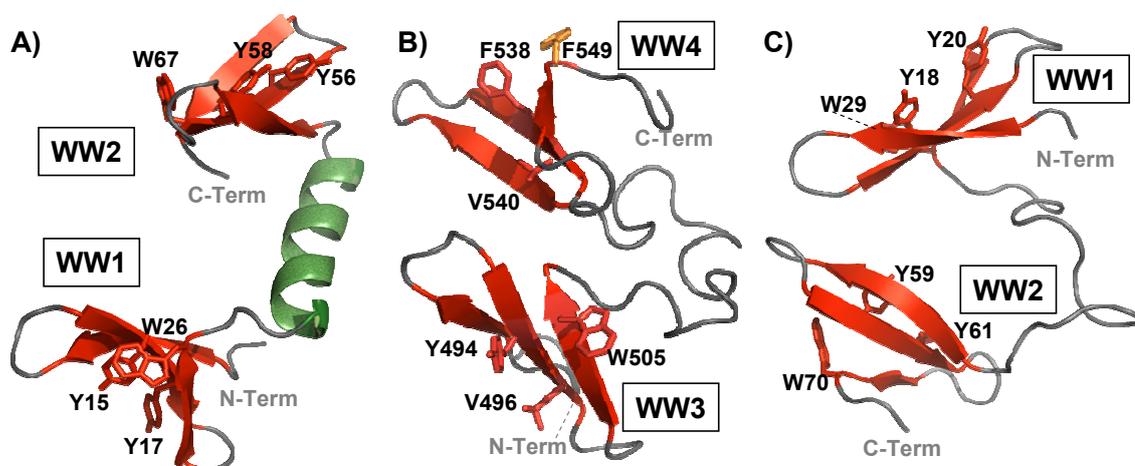
En otros intentos por comprender la regulación de hYap, sin embargo, se ha sugerido un efecto en sentido opuesto a partir de la fosforilación con Lats1 o Lats2. Es decir, se ha propuesto que la fosforilación en una posición diferente de la S127 provoca la liberación de hYap del anclaje a 14-3-3 (Matallanas et al. 2007). En la Figura 1.6, además de la S127 se ilustran otras 4 posiciones dentro de la secuencia de hYap que son posibles dianas de fosforilación para estas quinasas. En el caso de las serinas S61 y S109, ubicadas dentro del dominio de unión a factores de transcripción (TB), se ha planteado la posibilidad de que su fosforilación afecte negativamente a la unión de los factores de transcripción (TEF/TEAD), aunque esta hipótesis todavía no se ha validado (Wang et al. 2009). Otra posibilidad es que la fosforilación del dominio de co-activación transcripcional (TA) esté relacionada con el incremento de la afinidad por alguna ligasa de ubiquitina tipo E3, tal como se ha documentado entre la ligasa  $\beta$ -TRCP y el parólogo de hYap, TAZ (Tian et al. 2007).

Por otra parte, una cuestión importante y, aún no resuelta, es esclarecer las posibles implicaciones de la presencia de dos dominios WW en tándem en las características funcionales de la isoforma hYap2 con respecto a las hYap1 (con un solo dominio WW). A la vista de la mayor capacidad en la transcripción de genes en el núcleo de la forma hYap2 frente a hYap1, y sabiendo que la única diferencia entre ellas es la presencia del segundo dominio WW en hYap2, es razonable pensar que esta mayor actividad debe estar relacionada con la presencia del segundo dominio WW de la isoforma hYap2. En este sentido, es importante establecer cuál es el impacto de la presencia del segundo dominio WW en las propiedades de hYap. Concretamente, es nuestro objetivo investigar si la función del nuevo dominio es modular las interacciones con las dianas de la proteína, bien

modificando la afinidad de unión del primer dominio WW mediante interacciones cooperativas, o bien proporcionando nuevos sitios de anclaje para ligandos diferentes, afectando, por tanto, la especificidad de unión.

Es importante también establecer los posibles efectos que la configuración en tándem de los dos dominios WW pudiera tener sobre la estabilidad y cooperatividad estructural de la proteína. Como hemos señalado, el análisis del plegamiento y de las propiedades cooperativas es muy relevante para la comprensión de la especificidad de unión de los módulos de reconocimiento molecular y, por tanto, para el diseño racional de ligandos. En este sentido, en los últimos años se han descubierto múltiples proteínas cuyos dominios modulares de reconocimiento están ubicados consecutivamente. Esta configuración modular de las proteínas plantea la existencia de posibles efectos sinérgicos entre los diferentes dominios consecutivos que podrían dar lugar a efectos contextuales en la afinidad y especificidad de unión a sus ligandos. Por ejemplo, estudios recientes han evidenciado la importancia de la secuencia de conexión entre dominios SH3 y SH2 en la regulación de quinasas de la familia Src. Estas quinasas parecen haber evolucionado para restringir adecuadamente el espacio conformacional en disolución del tándem SH3-SH2 y de esta manera, regular la actividad de la enzima (Faraldo-Gomez & Roux 2007b; Faraldo-Gomez & Roux 2007a). Asimismo, estudios estructurales de distintas proteínas con dominios WW en tándem, revelan que la secuencia que los conecta tiene propiedades y funciones muy diferentes en cada caso. Esta secuencia puede estar estructurada y orientar a los dominios en caras opuestas, como sucede para la pareja de dominios WW1-2 de la proteína Prp40 (Figura 1.7A). En este caso, los dominios WW requieren de esta precisa orientación para unir proteínas específicas en los intrones de los precursores del ARNm en el espliceosoma (Wiesner et al. 2002). La secuencia que conecta los dominios también puede ser una secuencia más larga, flexible y desestructurada como la secuencia ubicada entre los dominios WW3-4 de la proteína Suppressor de Deltex Su(dx) (Figura 1.7B). Dicha secuencia estabiliza al dominio Su(dx)-WW4 que está parcialmente desestructurado antes de que se produzca la unión de un ligando al otro dominio Su(dx)-WW3 (Fedoroff et al. 2004; Jennings et al. 2007). En ocasiones, la secuencia de conexión puede ser

relativamente corta y flexible permitiendo que los dominios tengan cierta libertad conformacional pero obligando a que estén suficientemente próximos el uno del otro para aumentar la superficie de unión y reconocer ligandos con dos motivos de unión (Figura 1.7C), como se ha descrito para los dominios de WW1-2 de la proteína Fbp21 (Huang et al. 2009). Recientemente, se ha descrito otra forma de unión cooperativa para el primer y segundo dominios WW del tándem de Smurf1 (Smurf1-WW1-2) similar a la del tándem Fbp21-WW1-2, aunque en este caso el segundo dominio individual Smurf1-WW2 por sí sólo no tiene capacidad de unir ligandos ricos en prolina. Sólo cuando se encuentra formando un tándem con el otro dominio interaccionan entre ellos expandiendo la superficie de unión con el péptido. En este complejo, el dominio Smurf1-WW2 proporciona contactos auxiliares con los restos contiguos al único motivo PPxY del ligando (que une al otro dominio) con el consiguiente aumento de afinidad de unión (Chong et al. 2010).



**Figura 1.7. Representación de los sitios de unión:** **A)** dentro del tándem WW1-2 de la proteína factor de “splicing” de levadura Prp40, **B)** dentro del tándem WW3-4 de la proteína “Suppressor Deltex” Su(dx) y, **C)** dentro del tándem WW1-2 de Fbp21 (homólogo humano de Prp40). Los dominios WW se muestran como cintas de color rojo, el conector con estructura de hélice  $\alpha$  de Prp40 en color verde, las cadenas laterales de los residuos más importantes para el reconocimiento de secuencias ricas en prolina se muestran como varillas rojas y el resto, como cintas de color gris. Las estructuras de Prp40, de Su(dx) y Fbp21 corresponden a los ficheros depositados en el PDB con los códigos 1O6W, 1TK7 y 2JXW, respectivamente.

En este contexto, dentro de los objetivos propuesto en este trabajo se encuentran, por un lado, investigar la afinidad y la especificidad de unión de los dos dominios WW individuales de hYap con varias dianas naturales ricas en

prolina (entre ellas, las identificadas en las quinasas de serina Lats1 y Lats2) y, por otro lado, describir posibles efectos sinérgicos del tándem de dominios WW hYap2, tanto en su equilibrio conformacional, como en el reconocimiento de estos ligandos.

### 1.5. Objetivos de la Tesis

El objetivo general de esta Tesis es comprender los determinantes del reconocimiento de ligandos ricos en prolina por los dominios WW para una posterior incorporación al desarrollo de estrategias racionales de diseño de inhibidores eficaces y específicos de dichas interacciones. Más concretamente, los objetivos específicos son:

- I. Analizar la interrelación entre los procesos de unión y de plegamiento de los dominios mediante la caracterización de la estabilidad y de las propiedades conformacionales y dinámicas de los dominios WW.
- II. Realizar un análisis detallado de la energética de unión de varios dominios WW con ligandos ricos en prolina derivados de algunas de sus dianas naturales.
- III. Analizar los factores contextuales asociados a posibles interacciones cooperativas entre el tándem de dominios WW de la isoforma hYap2.

## CAPÍTULO 2



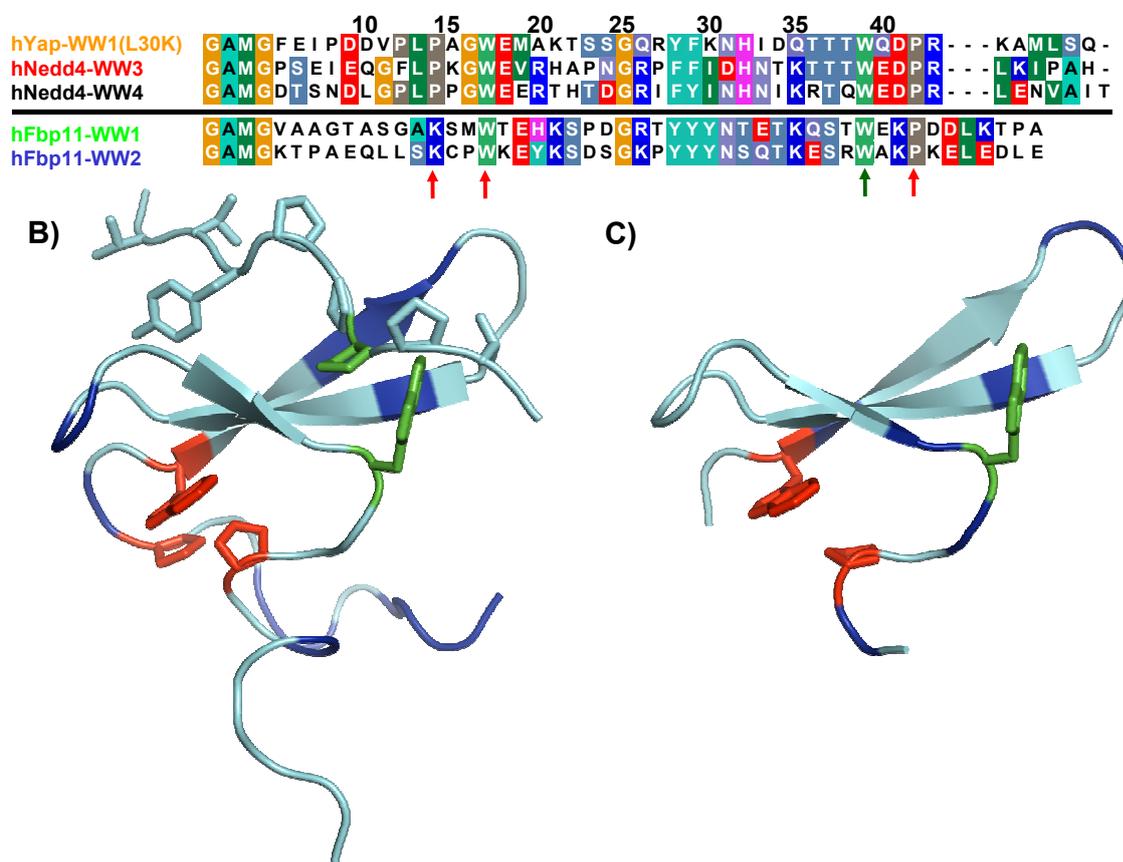
## **CAPÍTULO 2. Estudio termodinámico de la estabilidad de dominios WW**

### **2.1. Introducción**

A pesar del interés que despierta el estudio de la estabilidad de los dominios WW como sistemas modelo, actualmente hay relativamente pocos trabajos dedicados al estudio sistemático del equilibrio conformacional y estabilidad de estos dominios en la bibliografía. Únicamente se han publicado estudios puntuales de la estabilidad de algunos dominios, como pueden ser el primer dominio de la proteína humana Yap (Koepf et al. 1999a; Koepf et al. 1999b; Crane et al. 2000; Ferguson et al. 2001a; Karanicolas & Brooks 2004), el dominio WW de la proteína humana Pin1 (Karanicolas & Brooks 2004; Bruscolini et al. 2007; Liu et al. 2008) y el dominio WW de la proteína humana Fbp28 (Ferguson et al. 2001a; Karanicolas & Brooks 2004; Petrovich et al. 2006; Sharpe et al. 2007). En la mayoría de casos, el plegamiento de los dominios WW se ha descrito en términos de equilibrio de dos estados, a pesar del hecho de que estos estudios emplean únicamente técnicas espectroscópicas que no permiten establecer de forma inequívoca la naturaleza del equilibrio conformacional (Koepf et al. 1999b; Ferguson et al. 2001a; Jager et al. 2001). En este sentido, está bien establecido en la bibliografía que cualquier estrategia destinada a distinguir un equilibrio de dos estados de otros equilibrios más complejos debe investigar el desplegamiento de proteínas mediante diferentes técnicas que proporcionen información sobre diferentes características estructurales, a medida que se aumenta la temperatura o a medida que se incrementa la presencia de agentes desnaturalizantes (Freire 1995; Sanchez-Ruiz 1995; Martinez et al. 2011).

En este capítulo, hemos abordado la caracterización termodinámica del equilibrio conformacional de un conjunto de dominios WW clasificados en distintos grupos de especificidad de unión y procedentes de distintas proteínas, como se ilustra en el Panel A de la Figura 2.1. Para ello, hemos combinado técnicas calorimétricas y espectroscópicas en un amplio intervalo de condiciones, incluyendo cambios de temperatura, pH y concentración de agente

desnaturalizante, con el objetivo de caracterizar con detalle el equilibrio conformacional de estos dominios.



**Figura 2.1. Dominios WW de clases I y II. A)** Alineamiento de secuencias de todos los dominios WW utilizados. Se han dividido en dos grupos en función de su similitud de secuencia y su especificidad de unión: hYap-WW1, hNedd4-WW3 y hNedd4-WW4 del grupo I, y hFbp11-WW1 y hFbp11-WW2 del grupo II. Con diferentes colores se resaltan aquellos restos conservados entre las secuencias, en negro aquellos que son diferentes y con flechas se indican las posiciones de los restos que conforman un broche hidrofóbico responsable de la integridad de los dominios WW. Representación esquemática de cintas de las estructuras determinadas por RMN del dominio hYap-WW1 (L30K) formando complejo con el péptido GPPPPYTVG (PDB: 1JMQ) **B)** y del dominio hFbp1-WW1 (PDB: 1ZR7) **C)**. El esqueleto polipeptídico esta representado en color turquesa, las cadenas laterales de P14, W17 y P32 formando un empaquetamiento hidrofóbico con varillas, se han representado en rojo y las cadenas laterales del resto W39, que interacciona con la prolina del péptido se han representado como varillas verdes. Sobre la representación esquemática de cada dominio se han mostrado en color azul las diferencias entre los dominios WW3 y WW4 de Nedd4 y hYAP-WW1(L30K) en B), y entre el dominio hFbp11-WW1 y el dominio hFbp11-WW2 en C).

En primer lugar se ha llevado a cabo la caracterización termodinámica completa y exhaustiva del cuarto dominio WW de la ligasa de ubiquitina humana Nedd4, hNedd4-WW4. Este análisis nos ha permitido establecer las pautas

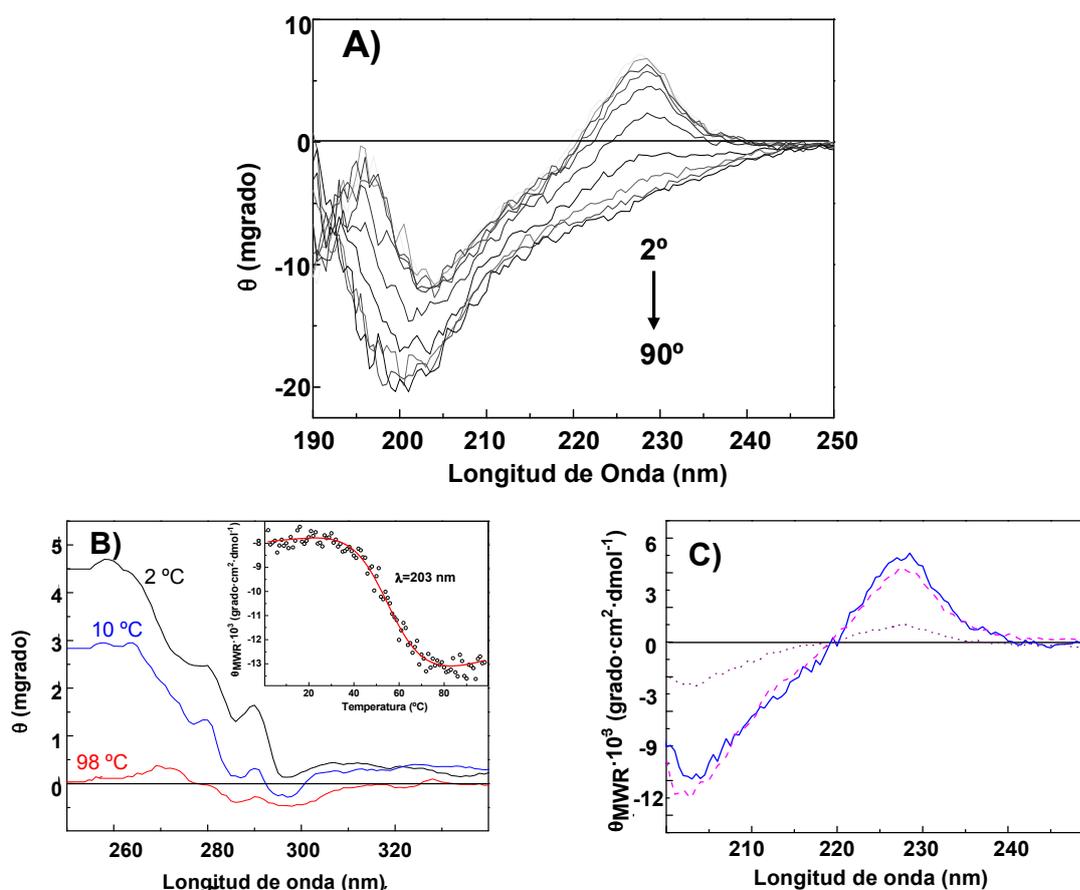
generales para la caracterización rigurosa de este tipo de dominios. Así, utilizando como referencia la caracterización termodinámica de hNedd4-WW4, se ha abordado el estudio de otros cuatro dominios WW: el tercer dominio WW de la misma proteína Nedd4, hNedd4-WW3; el mutante L30K del dominio WW1 de la proteína humana Yap, hYap-WW1(L30K), cuya estructura en complejo con el péptido GTPPPYTVG se ha representado en el Panel B de la Figura 2.1; y los dos dominios WW presentes en la proteína humana Fbp11, hFbp11-WW1, cuya estructura se muestra en el Panel C de la Figura 2.1, y hFbp11-WW2.

## **2.2. Caracterización del equilibrio conformacional del dominio hNedd4-WW4**

### **2.2.1. Espectros de dicroísmo circular**

Con el fin de obtener información sobre las propiedades conformacionales del dominio hNedd4-WW4 e identificar las mejores condiciones para realizar una caracterización termodinámica completa de este dominio, se registraron diferentes espectros de dicroísmo circular (CD).

El espectro de CD del estado nativo de hNedd4-WW4 en la región del UV-lejano se caracteriza por un pequeño pico positivo centrado en 230 nm y un intenso pico negativo a 203 nm, como se aprecia en el Panel A de la Figura 2.2. Este comportamiento es similar al descrito en la bibliografía para otros dominios WW (Koepf et al. 1999a; Koepf et al. 1999b; Ferguson et al. 2001a; Jager et al. 2001; Jager et al. 2007), y para otros dominios o proteínas cuya estructura secundaria es fundamentalmente lamina  $\beta$  (Viguera et al. 1994; Knapp et al. 1998; Reid et al. 1998). El máximo a 230 nm parece estar relacionado con interacciones locales que implican cadenas laterales aromáticas, principalmente las de restos triptófano en un microambiente hidrofóbico (Woody 1994), lo que en nuestro caso, está probablemente asociado al empaquetamiento aromático entre el primer triptófano W17, la prolina en la posición 42, P42, y otros residuos como la prolina en la posición 14, P14 (Figura 2.1), formando lo que se ha denominado un broche ó “core” hidrofóbico (Koepf et al. 1999a).



**Figura 2.2. Espectros de CD del dominio hNedd4-WW4 en fosfato sódico 20 mM pH 7.0.** A) Espectros de dicroísmo circular en UV-lejano cada diez grados durante el calentamiento de 2° a 90 °C. B) Espectros de CD en el UV-cercano a diferentes temperaturas: a 2 °C, en negro; a 10 °C, en azul; y a 98 °C, en rojo. En el recuadro se representa la dependencia de la señal de CD seguida a 203 nm (UV-lejano) con la temperatura para el dominio hNedd4-WW4. Para una mejor ilustración de la forma sigmoidal, en línea roja, se ha representado el mejor ajuste de acuerdo con el modelo de dos estados (Apéndice II). C) Espectros de dicroísmo circular en UV-lejano a 2 °C en diferentes condiciones de fuerza iónica. El espectro representado como una línea continua azul fue tomado antes de calentar a baja fuerza iónica (cloruro sódico 0 mM), aquel en línea discontinua magenta se registró después de calentar a baja fuerza iónica, y el que está representado como una línea de puntos de color violeta se tomó después de calentar a elevada fuerza iónica (cloruro sódico 100 mM pH 7.0).

La intensa banda negativa en torno a 200nm no es propia de un espectro de CD en el UV-lejano típico de una lamina  $\beta$  caracterizado por una banda negativa a 217 nm y una positiva a 195 nm como se muestra en el ejemplo de la Figura 6.3B de Materiales y Métodos sino que más bien corresponde a una conformación de ovillo-desordenado (“random-coil”), con una banda negativa a unos 200 nm y otra positiva hacia los 185-190 nm (Manavalan & Johnson 1987). Los espectros de este tipo se han descrito para proteínas que contienen láminas  $\beta$  distorsionadas o hebras cortas e irregulares, conocidas como proteínas  $\beta$ -II. Estas

proteínas tienen una menor proporción de lámina  $\beta$  con respecto a la estructura de ovillo-desordenado que las proteínas  $\beta$  normales ( $\beta$ -clase I), que presentan los espectros típicos de CD en UV-lejano descritos anteriormente (Sreerama & Woody 2003). La estructura de los dominios WW puede presentar largos extremos que parecen estar desordenados, como se observa en la Figura 2.1B para el mutante hYap-WW1(L30K), cuyos espectros suelen ser del tipo “random coil”. A la vista de estos resultados, el espectro del dominio hNedd4-WW4 nos indica que este dominio podría clasificarse como una proteína de tipo  $\beta$ -II. En general, para distinguir la lámina tipo  $\beta$ -II de conformaciones desordenadas se han propuesto principalmente dos criterios: I) la presencia de bandas de absorción en el UV-cercano, y II) una dependencia sigmoideal de la elipticidad con la temperatura a la longitud de onda del mínimo a unos 200 nm (Wu et al. 1992). En el Panel B de la Figura 2.2 se muestran varios espectros de CD en el UV-cercano a distintas temperaturas para el dominio hNedd4-WW4, así como la dependencia de la temperatura para este dominio seguida a 203 nm (en el recuadro de la Figura 2.2B). En ellos se observan varias bandas de absorción correspondientes a la región del UV-cercano a 2 °C y 10 °C que desaparecen con la desnaturalización del dominio a alta temperatura. Además, se obtuvieron transiciones sigmoideas para la dependencia de la señal de CD con la temperatura a las distintas longitudes de onda. Estos resultados indican que el dominio hNedd4-WW4 claramente cumple estos criterios y que, en general, los dominios WW se pueden clasificar como proteínas  $\beta$ -II.

En el Panel C de la Figura 2.2 se muestran los espectros de CD en el UV-lejano para el dominio hNedd4-WW4 bajo diferentes condiciones de fuerza iónica. En principio no se observan diferencias en la forma de los espectros a baja fuerza iónica (tampón fosfato sódico 40 mM, pH 7.0) o a elevada fuerza iónica (tampón fosfato sódico 40 mM cloruro sódico 100mM, pH 7.0). Sin embargo, hay importantes diferencias en la reversibilidad del desplegamiento al variar la concentración de sal. Cuando la fuerza iónica es baja, el espectro a 2 °C después de calentar hasta 98 °C es muy similar al de la proteína nativa a 2 °C antes de calentar. Por el contrario, si la fuerza iónica es mayor, el espectro nativo no se recupera. Para mantener las condiciones de equilibrio, todos los experimentos de

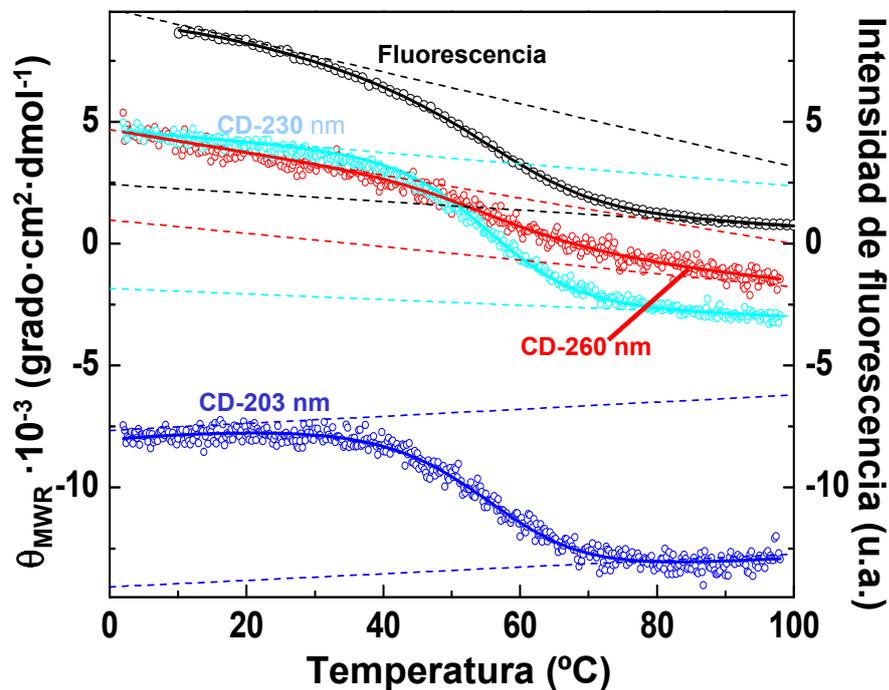
desnaturalización térmica de hNedd4-WW4 se han realizado en condiciones de baja fuerza iónica.

### **2.2.2. Caracterización del desplegamiento térmico del dominio hNedd4-WW4**

#### **2.2.2.1. Desplegamiento térmico mediante técnicas espectroscópicas**

Una estrategia particularmente útil a la hora de identificar si el proceso de desplegamiento de una proteína se describe adecuadamente mediante el modelo de dos estados es la obtención de idénticos parámetros termodinámicos a partir de diferentes observables, como serían señales espectroscópicas que describan distintos aspectos estructurales de la proteína (Ginsburg & Carroll 1965; Sanchez-Ruiz 1995). A la vista de los espectros de CD del dominio hNedd4-WW4 que se muestran en la Figura 2.2, hemos seleccionado varias longitudes de onda de distintas regiones del espectro que parecen adecuadas para seguir el desplegamiento térmico del dominio WW, aquellas en las que la variación de intensidad con la temperatura es más acusada. Para la región del UV-lejano (Figura 2.2A) hemos escogido dos longitudes de onda, el mínimo a 203 nm y el máximo a 230 nm. En la región del UV-cercano, las diferencias de intensidad son muy pequeñas en todo el intervalo de longitud de onda (Figura 2.2B), por lo que se ha elegido la longitud del máximo relativo a 260 nm. Por otro lado, para completar la información acerca de la desnaturalización térmica de este dominio WW se ha seguido la variación de la fluorescencia de los restos triptófano del dominio, es decir, se ha seguido un observable espectroscópico local que proporciona información relativa a la estructura terciaria del dominio WW. En la Figura 2.3 se muestra que los perfiles de desnaturalización térmica seguidos mediante los cuatro observables espectroscópicos (fluorescencia a 344 nm y dicroísmo circular a 203, 230 y 260 nm) en tampón fosfato sódico 20 mM pH 7.0 presentan una forma sigmoidal con un punto medio de transición en torno a los 55 °C.

Sobre los datos experimentales en la Figura 2.3 se ha representado el mejor ajuste global de las cuatro curvas de acuerdo con el modelo de dos estados (Sección 6.4.4 de Materiales y Métodos y Apéndice II). En este análisis, los parámetros termodinámicos  $T_m$  y  $\Delta H_m$  son comunes a las cuatro curvas,



**Figura 2.3. Desnaturalización térmica del dominio hNedd4-WW4 seguida mediante diferentes técnicas espectroscópicas.** Los círculos representan los datos experimentales recogidos en tampón fosfato sódico 20 mM pH 7.0. El experimento de CD en el UV-cercano seguido a 260 nm en rojo se ha realizado con 220  $\mu$ M de proteína. Los experimentos de CD en el UV-lejano seguidos a 230 y 203 nm en turquesa y azul, respectivamente, se han registrado con muestras de concentración 50  $\mu$ M. El experimento de fluorescencia seguido a 344 nm en negro se ha realizado a una concentración de 20  $\mu$ M de proteína. Las líneas continuas que acompañan los datos experimentales corresponden al mejor ajuste global según el modelo de dos estados (Apéndice II) compartiendo los parámetros  $\Delta H_m$  y  $T_m$ . Las líneas rectas discontinuas representan las líneas base obtenidas a partir de la estrategia de ajuste global para los estados nativo y desplegado en cada curva de desnaturalización.

mientras que las correspondientes líneas base se obtuvieron considerándolas como líneas rectas independientes para cada curva. Los parámetros termodinámicos de este ajuste se recogen en la Tabla 2.1. En esta figura se puede observar que los datos experimentales pueden ser descritos adecuadamente por el modelo de dos estados, que proporciona un conjunto común de parámetros termodinámicos para los cuatro perfiles de desnaturalización térmica, así como líneas base muy similares, casi paralelas, con una pequeña dependencia de la temperatura, como cabe esperar de un equilibrio de dos estados (Viguera et al. 1994; Nicholson & Scholtz 1996; Reid et al. 1998; Padmanabhan et al. 1999; Crane et al. 2000; Schuler et al. 2002; Felitsky & Record 2003; Streicher & Makhatadze 2007). Sin embargo, aunque se están utilizando distintas técnicas y

diferentes longitudes de onda, es importante mencionar que realmente los observables estructurales seguidos son muy similares en todas las curvas. Tanto la fluorescencia de triptófano como la señal de CD en UV-cercano o la banda a 230 nm en UV-lejano, proporcionan una información que es resultado de la variación de las interacciones entre aromáticos durante el desplegamiento. Por tanto, este tipo de análisis meramente espectroscópico no nos permite tener una visión multidimensional del equilibrio conformacional o de la cooperatividad durante la desnaturalización térmica que confirme el desplegamiento según el modelo de dos estados del dominio hNedd4-WW4.

Un equilibrio de dos estados se asocia a la existencia de un punto isobéptico (longitud de onda a la que dos especies químicas, por lo general relacionadas linealmente, tienen la misma capacidad de absorción molar,  $\epsilon$ ). Aunque en los espectros de CD del dominio hNedd4-WW4 recogidos a distintas temperaturas no se observó ningún punto isodicroico (Figura 2.2A), parece que estos espectros tienden a cruzarse alrededor de los 195 nm. Es posible que este punto no quede bien definido ya que se encuentra en una región del espectro con poca resolución, en el que la relación señal/ruido es elevada. En cualquier caso, la ausencia de este punto isodicroico también ha sido descrita previamente para otras proteínas dos-estados con un espectro similar de hebra  $\beta$ -II, como son los dominios SH3 (Knapp et al. 1998) o las proteínas “cold-shock” (Reid et al. 1998).

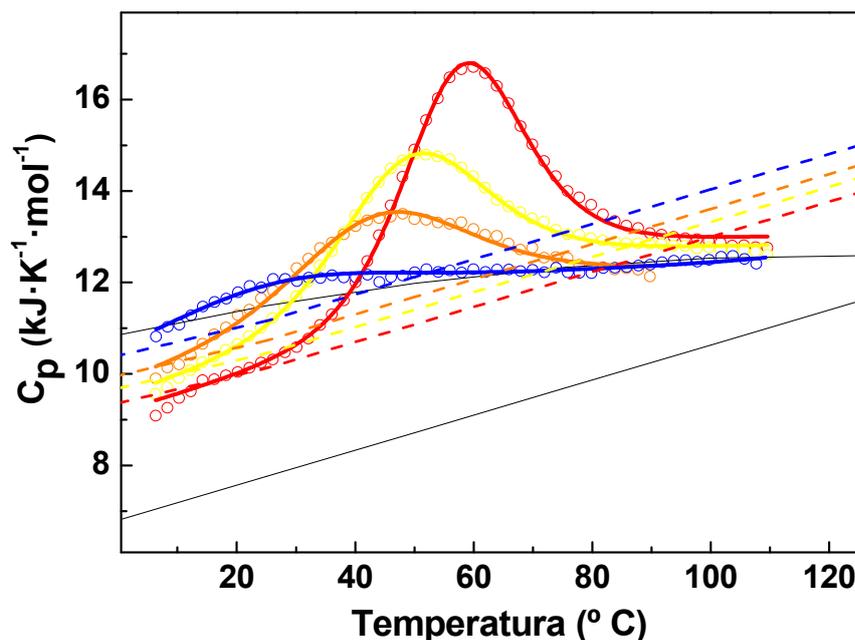
**Tabla 2.1. Parámetros termodinámicos de plegamiento del dominio hNedd4-WW4 obtenidos mediante el ajuste global de experimentos de desnaturalización térmica seguida por DSC y CD<sup>a</sup>.**

Método de ajuste global	pH	T <sub>m</sub> (°C)	$\Delta H_m$ (kJ·mol <sup>-1</sup> )	$\Delta H_{m,exp}$ (kJ·mol <sup>-1</sup> )	$\Delta C_p$ (kJ·K <sup>-1</sup> ·mol <sup>-1</sup> )	$\Delta G(25^\circ\text{C})$ (kJ·mol <sup>-1</sup> )	T <sub>s</sub> (°C)	$\Delta G_{m\acute{a}x}(T_s)$ (kJ·mol <sup>-1</sup> )
CD+DSC (varios pH) <sup>c</sup>	7	58.0	133	120		11.6	-25	19.5
	6	46.5	100	85		5.9	-16	11.0
	5	37.4	73	65		2.6	-9	6.0
	4 <sup>b</sup>	22.6	41			-0.3	-3	1.8
CD + fluorescencia	7 <sup>b</sup>	54.2	120		2.5	10.4		
Doble perturbación CD	7 <sup>b</sup>	55.1	150		2.6	9.2		

<sup>a</sup>Los errores se han estimado en 5% para la T<sub>m</sub>, 10% para  $\Delta H_m$ , 30% para  $\Delta H_{m,exp}$  y 20% para el resto de parámetros. <sup>b</sup>En estas condiciones de pH, los experimentos de desnaturalización térmica únicamente se obtuvieron mediante CD. <sup>c</sup>Valores para los coeficientes que definen las líneas base compartidas entre los diferentes conjuntos de experimentos de acuerdo con la estrategia de ajuste según el modelo de dos estados, donde T<sub>r</sub>, temperatura de referencia, es 25 °C y los valores subrayados fueron estimados a partir de la contribución individual de cada aminoácido (Privalov & Makhatadze 1990):  $\Theta_{230,N} = 15.4 - 0.04 \cdot (T - T_r)$ ;  $\Theta_{230,D} = 1.21 - 0.0112 \cdot (T - T_r)$ ;  $C_{p,N} = 10.11 - 0.038 \cdot (T - T_r)$ ; y  $C_{p,N} = 10.52 - 0.0238 \cdot (T - T_r) - 0.1163 \cdot 10^{-3} \cdot (T - T_r)^2$ .

### 2.2.2.2. Desplegamiento térmico mediante calorimetría (DSC)

En la Figura 2.4 se muestran las trazas de DSC correspondientes al proceso de desnaturalización térmica del dominio hNedd4-WW4 en el intervalo de pH entre 7.0 y 3.0. La desnaturalización térmica se ha registrado a baja fuerza iónica, de modo que la reversibilidad del proceso de desnaturalización ha sido superior al 80% en todos los casos. Tampoco se han apreciado efectos debidos a la velocidad de barrido o a la concentración de la muestra ni en las trazas de DSC ni en las curvas de CD. Por tanto, el dominio puede ser considerado como una proteína monomérica que se despliega en equilibrio bajo estas condiciones experimentales. Esto ha sido confirmado mediante experimentos de dispersión dinámica de luz (realizados como se describe en Cobos et al. 2004), en los que se comprobó que el 99.8% de la muestra se encontraba en forma monomérica, con un radio estimado de 1.8 nm.

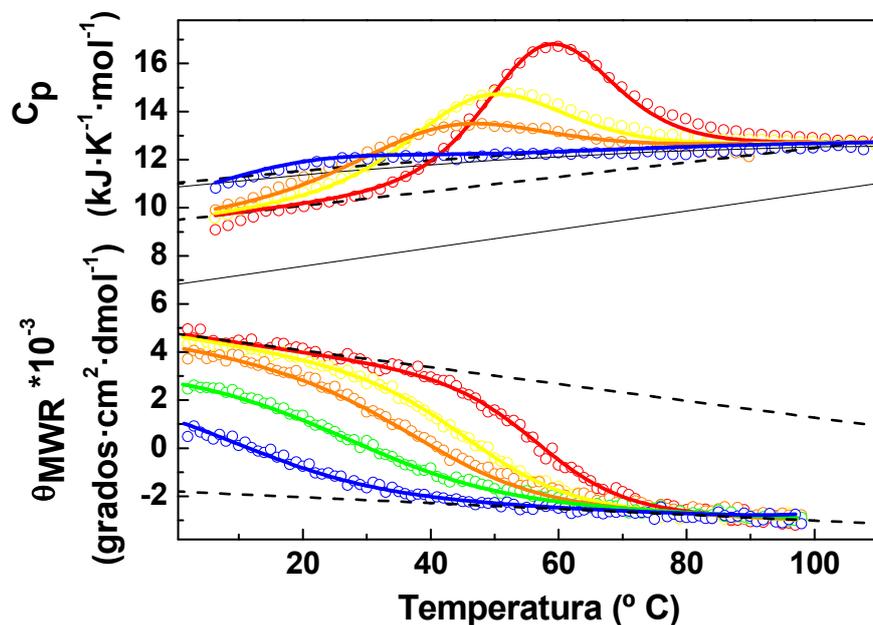


**Figura 2.4. Dependencia de la capacidad calorífica del dominio hNedd4-WW4 con la temperatura medida en diferentes condiciones experimentales.** Los datos se han representado como círculos vacíos con diferente color según la condición de pH que representan: rojo para 20 mM fosfato sódico pH 7.0, amarillo para 20 mM MES pH 6.0, naranja para 20 mM acetato sódico pH 5.0 y azul para 20 mM glicina pH 3.0. El mejor ajuste individual, considerando funciones  $C_{p,N}(T)$  y  $C_{p,D}(T)$  para cada curva según el modelo de dos estados, se ha mostrado como una línea continua acompañando a los datos experimentales en el mismo código de colores. Las líneas discontinuas en los mismos colores corresponden a las funciones  $C_{p,N}(T)$  obtenidas a partir del ajuste para cada curva. Las líneas continuas en negro corresponden a la función  $C_{p,D}(T)$  calculada mediante la adición de las contribuciones de cada aminoácido (Privalov & Makhatadze 1990) (arriba) y a la función teórica  $C_{p,N}(T)$  estimada a partir del peso molecular del dominio (Freire 1995) (abajo).

Las curvas de DSC en la Figura 2.4 presentan una transición muy amplia que abarca gran parte del intervalo experimental de temperatura. A bajas temperaturas, además se observa una importante desviación de los valores de  $C_p$  en cada curva, mientras que a elevadas temperaturas presentan valores muy similares. Los cuatro perfiles de DSC se han analizado (como se detalla en el Apartado 6.3 de Materiales y Métodos) mediante un ajuste individual al modelo de dos estados (Apéndice I). Para cada curva se han considerado como parámetros ajustables los valores de  $T_m$  y  $\Delta H_m$ . Además se han definido para cada una de ellas dos funciones de capacidad calorífica, una para el estado nativo,  $C_{p,N}(T)$ , y otra para el estado desnaturalizado,  $C_{p,D}(T)$ . Como se ilustra en la Figura 2.4, este análisis proporciona una descripción razonable de los datos experimentales, con valores de  $R$  y  $R^2$  superiores a 0,999. Los coeficientes dependientes de la temperatura de la función  $C_{p,D}(T)$  se han estimado a partir de las contribuciones tabuladas de cada uno de los aminoácidos (Privalov & Makhatadze 1990), con el fin de disminuir parámetros en el ajuste. Esta función teórica (también representada en la Figura 2.4) coincide con la capacidad calorífica de los datos experimentales después de la transición principal, como se ha descrito para otras proteínas pequeñas (Martinez et al. 1994; Viguera et al. 1994; Cobos et al. 2004).

El análisis del comportamiento de las líneas base en el contexto del modelo de plegamiento de dos estados constituye una cuestión importante, ya que unas líneas base mal definidas podrían enmascarar evidencias de la existencia de estados intermedios poblados o de otras contribuciones como las producidas cuando el desplegamiento implica pasar por múltiples estados conformacionales (Zhou et al. 1999; Privalov & Dragan 2007). Es importante señalar que la intersección de las funciones  $C_{p,N}(T)$  y  $C_{p,D}(T)$  obtenidas mediante el ajuste individual se produce cerca de los valores de  $T_m$  (Figura 2.4). Estas temperaturas distan bastante de la temperatura considerada tradicionalmente como la temperatura de convergencia universal de las proteínas globulares, 140 °C. Estos valores sugieren que ambas funciones  $C_{p,N}(T)$  y  $C_{p,D}(T)$  no están correctamente definidas a partir del ajuste individual de las curvas de DSC, ya que a elevada temperatura se obtienen valores positivos para el cambio de capacidad calorífica durante el desplegamiento,  $\Delta C_{p,N-D}$ , lo cual no tiene ningún sentido físico.

En la Figura 2.5 se han representado las mismas curvas de desplegamiento térmico seguidas por DSC mostradas en la Figura 2.4, junto con otras curvas de desnaturalización térmica seguidas mediante CD a 230 nm en diferentes condiciones de pH. Al igual que sucede con los perfiles de DSC, las desnaturalizaciones térmicas seguidas por CD también presentan importantes desviaciones de la señal a bajas temperaturas mientras que los valores de las distintas curvas coinciden a elevadas temperaturas.



**Figura 2.5. Desnaturalización térmica del dominio hNedd4-WW4 seguida por CD (a 230 nm) y por DSC bajo distintas condiciones de pH.** Los datos experimentales de ambas técnicas se han representado como círculos vacíos con diferente color según la condición de pH que representan: rojo para fosfato sódico 20mM pH 7.0, amarillo para MES 20mM pH 6.0, naranja para acetato sódico 20mM pH 5.0, verde para acetato sódico 20mM pH 4.0 y azul para glicina 20mM pH 3.0. La desnaturalización térmica a pH 4.0 sólo se obtuvo mediante CD. Las líneas continuas corresponden al ajuste global de acuerdo con el modelo de dos estados, considerando comunes las variables  $\Delta H_m$  y  $T_m$  para cada condición de pH. Las líneas discontinuas en negro corresponden a las líneas base comunes:  $\theta_N(T)$ ,  $\theta_D(T)$ ,  $C_{p,N}(T)$  y  $C_{p,D}(T)$ , para cada conjunto de curvas obtenidas a partir del ajuste. Las líneas rectas discontinuas en negro bajo las trazas de DSC representan la función teórica  $C_{p,N}(T)$  correspondiente a una estimación basada en el peso molecular de cada dominio (Freire 1995), mientras que la línea continua con cierta curvatura representa la función teórica  $C_{p,D}(T)$  calculada mediante la adición de las contribuciones individuales de cada aminoácido en la secuencia de hNedd4-WW4 (Privalov & Makhatadze 1990).

En la Figura 2.5, se muestra el análisis de los dos conjuntos de datos experimentales mediante un ajuste global al modelo de dos estados. La idea es la misma que la planteada para el análisis global siguiendo la desnaturalización

térmica mediante diferentes observables espectroscópicos-estructurales (Figura 2.3): obtener una visión de la desnaturalización del dominio desde diferentes perspectivas. En este caso, el cambio de las interacciones entre restos aromáticos del dominio (CD 230 nm) y el cambio de la capacidad calorífica del sistema debido al desplegamiento del dominio WW,  $C_p(T)$ . Para este análisis se ha asumido que las curvas de DSC obtenidas para los diferentes valores de pH comparten una función común de  $C_{p,N}(T)$  y otra de  $C_{p,D}(T)$  (al igual que en la Figura 2.4, los coeficientes dependientes de la temperatura se han estimado a partir de la función teórica  $C_{p,D}(T)$  para disminuir el número de variables en el ajuste) y se ha asumido que  $T_m$  y  $\Delta H_m$  son variables independientes para cada valor de pH. También se ha asumido que las curvas de CD comparten las mismas líneas base para el estado nativo,  $\theta_N(T)$ , y para el estado desnaturalizado,  $\theta_D(T)$ . Este tipo de análisis se ha realizado para otras proteínas caracterizadas como dos-estados, como son los dominios SH3 (Viguera et al. 1994). Los parámetros termodinámicos de este ajuste se recogen en la Tabla 2.1.

El valor obtenido para la temperatura de intersección de las funciones  $C_{p,N}(T)$  y  $C_{p,D}(T)$  compartidas en el análisis global de los datos de CD y de DSC, está en torno a 100 °C, y se aproxima más a la temperatura de convergencia que el valor obtenido en el análisis individual de estas curvas (Figura 2.4). De hecho, está relativamente cerca del valor más realista, 120 °C, propuesto recientemente en una revisión realizada por Privalov y Dragan (Privalov & Dragan 2007). Además, el análisis global también proporciona unas líneas base de CD muy razonables que no se cruzan en el rango de temperatura de los experimentos y son muy similares a las líneas base obtenidas para otras proteínas pequeñas de hoja  $\beta$ , tales como los dominios SH3, las proteínas “cold-shock” de choque en frío (Viguera et al. 1994; Knapp et al. 1998; Reid et al. 1998) y otras pequeñas proteínas globulares (Nicholson & Scholtz 1996; Cobos et al. 2002; Felitsky & Record 2003).

A pesar de que las transiciones calorimétricas del dominio hNedd4-WW4 a los distintos valores de pH son muy anchas debido a una aparente baja cooperatividad conformacional del dominio, el análisis global con el modelo de dos estados considerando una  $C_{p,N}(T)$  común para las curvas registradas en diferentes

condiciones de pH describe adecuadamente los datos experimentales. La robustez del análisis global con el modelo de dos estados también se ve reflejada en el ajuste combinado de las curvas de CD puesto que también describe la dispersión a baja temperatura considerando una única línea base. Los valores de entalpía obtenidos de este análisis global según el modelo de dos estados para cada curva son comparables con los estimados directamente en las áreas bajo las curvas de capacidad calorífica,  $\Delta H_{m,exp}$  (Tabla 2.1), aunque hay que tener en cuenta que estos últimos están afectados por un error considerable. Ambos datos proporcionan una relación entalpía calorimétrica/entalpía van't Hoff para las curvas de pH 7.0, 6.0 y 5.0 que está en torno a 0.9, aproximadamente lo esperado para una proteína dos estados en lo que se ha convenido denominar “criterio calorimétrico” (Bhakuni et al. 1991; Bruylants et al. 2005). En resumen, el equilibrio de desplegamiento del dominio hNedd4-WW4 puede ser razonablemente descrito de acuerdo con el modelo de equilibrio de dos estados. La estrecha concordancia entre la calorimetría y las desnaturalizaciones térmicas seguidas mediante distintos observables espectroscópicos se refleja en los análisis combinados mostrados en las Figuras 2.3 y 2.5. No obstante, hay que destacar que algunos de los parámetros resultantes directamente de los diferentes ajustes muestran una dispersión importante, como por ejemplo los valores de  $T_m$  a pH 7.0 cuyas diferencias son de 4 °C (Tabla 2.1).

Los parámetros termodinámicos normalizados proporcionados por el análisis global de acuerdo con el modelo de dos estados del desplegamiento del dominio hNedd4-WW4, están justo en el límite superior de lo esperado según el estudio estadístico a 60 °C realizado con varias proteínas dos estados por Robertson y Murphy (Robertson & Murphy 1997). En el caso del dominio hNedd4-WW4, los valores obtenidos son:  $\Delta H_{N-D}(60\text{ °C}) = 3.1 \pm 0.4\text{ kJ}\cdot(\text{mol de residuos})^{-1}$ ,  $\Delta S_{N-D}(60\text{ °C}) = 9.5 \pm 1.5\text{ J}\cdot\text{K}^{-1}\cdot(\text{mol de residuos})^{-1}$  y  $\Delta C_{p,N-D} = 50 \pm 6\text{ J}\cdot\text{K}^{-1}\cdot(\text{mol de residuos})^{-1}$ . Estos resultados se pueden considerar dentro de la estadística debido a que ésta presenta un margen de error asociado muy elevado (7-15%).

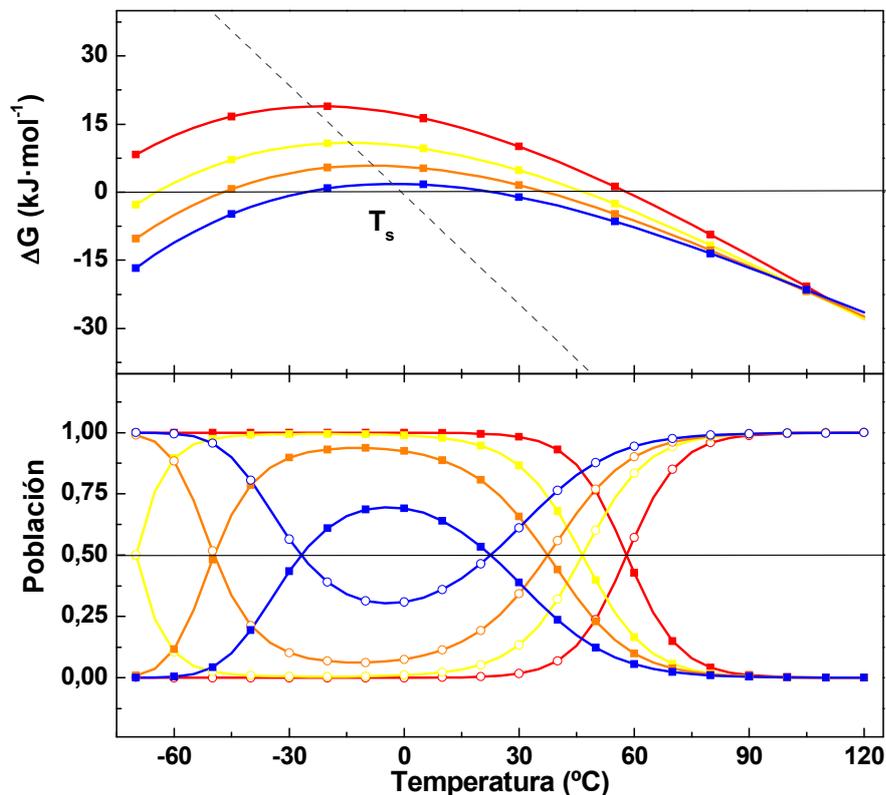
La función de entalpía específica promedio [ $\Delta h(T) = \Delta H(T)/P_m$ ] de las proteínas globulares normalmente converge a 110 °C en un valor medio de 54 J·g<sup>-1</sup>. Este comportamiento de convergencia de las proteínas tradicionalmente se ha

atribuido al hecho de que las contribuciones hidrofóbicas (a la entropía y entalpía) se aproximan a cero a la temperatura de la convergencia, por lo que, a esta temperatura, la entalpía de desplegamiento refleja sobre todo las interacciones polares y de van der Waals (Fu & Freire 1992). En este contexto, el hecho de que la mayoría de las proteínas globulares tengan similares entalpías específicas promedio a 110 °C indica un alto grado de homogeneidad en cuanto a la densidad y la calidad de las contribuciones entálpicas por residuo (Hilser et al. 1996). El valor de entalpía específica promedio puede calcularse mediante la extrapolación lineal de la entalpía de desplegamiento a 110 °C teniendo en cuenta el cambio de capacidad calorífica derivado de la regresión lineal de los valores de  $T_m$  frente  $\Delta H_m$  bajo diferentes condiciones de pH (Fu & Freire 1992; Martinez et al. 1994). En el caso del dominio hNedd4-WW4 se ha usado el valor de cambio de capacidad calorífica  $2.5 \pm 0.5 \text{ kJ}\cdot\text{K}^{-1}\cdot\text{mol}^{-1}$  ( $r = 0.995$ ), y se ha obtenido un valor de  $45 \text{ J}\cdot\text{g}^{-1}$ . A pesar de que el error asociado a la determinación de entalpía específica promedio es muy elevado, del orden del 20%, en nuestro caso dicho parámetro también se encuentra muy próximo al límite esperado para una proteína que se despliega en dos estados y además, por definición refleja la baja cooperatividad de este dominio.

### ▫ Curva de estabilidad

En la Figura 2.6, se han representado la dependencia con la temperatura de la función energía de Gibbs a diferentes valores de pH (Panel A) y las dependencias de las poblaciones correspondientes a los estados nativo y desnaturalizado (Panel B). Estas gráficas indican claramente que hNedd4-WW4 es sólo marginalmente estable, ya que, en el mejor de los casos, el valor máximo encontrado para la energía de Gibbs de este dominio es inferior a  $20 \text{ kJ}\cdot\text{mol}^{-1}$  (Tabla 2.1).

Las curvas de estabilidad obtenidas en diferentes condiciones de pH convergen a unos 120 °C (Figura 2.6A), lo cual está de acuerdo con las estadísticas de parámetros termodinámicos de desplegamiento de otras proteínas dos estados (Robertson & Murphy 1997). Sin embargo, es importante señalar que,



**Figura 2.6.** Curvas de estabilidad (panel superior) y análisis de la población de los estados nativo y desplegado con la temperatura (panel inferior) del dominio hNedd4-WW4. Los parámetros termodinámicos empleados para realizar estas representaciones son los obtenidos a partir del análisis global de las curvas de desnaturalización térmica seguida por CD a una longitud de onda de 230 nm y de DSC bajo condiciones desde pH neutro hasta pH ácido (Tabla 2.1). En el panel superior, las líneas continuas con cuadrados representan los cambios de energía libre de gibbs con la temperatura,  $\Delta G_{N-D}(T)$ , según la condición de pH: rojo para fosfato sódico 20mM pH 7.0, amarillo para MES 20mM pH 6.0, naranja para acetato sódico 20mM pH 5.0 y azul para glicina 20mM pH 3.0. Las líneas discontinuas en gris oscuro corresponden la dependencia con el pH del valor de  $T_s$ . En el panel inferior, usando el mismo colores se han simulado las fracciones molares correspondientes a los estados nativo (cuadrados llenos) y desplegado (círculos vacíos) considerando el mismo código de colores para cada condición de pH descrito para el panel superior.

en el caso de hNedd4-WW4, la temperatura del máximo de estabilidad ( $T_s$ ) es anormalmente baja, son valores entre  $-25\text{ }^{\circ}\text{C}$  a pH 7.0 y  $-3\text{ }^{\circ}\text{C}$  a pH 4.0 (Tabla 2.1). A pesar de que la información acerca de los valores de  $T_s$  para diferentes proteínas es escasa, una recopilación de valores de  $T_s$  (Felitsky & Record 2003) muestra claramente que, incluso para proteínas pequeñas (50-60 residuos) los valores de  $T_s$  son generalmente muy superiores a  $0\text{ }^{\circ}\text{C}$ , cercanos, incluso, en algunos casos, a  $30\text{ }^{\circ}\text{C}$ . De hecho, el valor más bajo que hemos encontrado hasta la fecha en la bibliografía es el de la proteína CI2, con un valor de  $-8\text{ }^{\circ}\text{C}$  (Jackson & Fersht 1991), muy superior a los  $-25\text{ }^{\circ}\text{C}$  obtenidos para el dominio hNedd4-WW4.

En este punto, surge la pregunta sobre el significado de una  $T_S$  tan baja. Es interesante recordar que el incremento del valor de  $T_S$ , se ha descrito como una de las estrategias utilizadas por proteínas termófilas para lograr una mayor estabilidad a altas temperaturas mediante una rigidificación significativa de su estado nativo (Nojima et al. 1977; Jaenicke 2000; Razvi & Scholtz 2006). En este caso, argumentando en sentido contrario, se podría especular que, dado que los dominios WW están generalmente involucrados en la mediación de interacciones proteína-proteína, el valor de  $T_S$  anormalmente bajo que caracteriza al dominio hNedd4-WW4 podría reflejar la necesidad de una gran plasticidad y flexibilidad del estado nativo para el reconocimiento efectivo de sus ligandos.

### ▫ Capacidad calorífica del estado nativo

La capacidad calorífica absoluta de este dominio a bajas temperaturas [ $C_{p,N}(20\text{ °C}) = 1.73 \pm 0.2\text{ J}\cdot\text{K}^{-1}\cdot\text{g}^{-1}$ ] es muy superior al valor de  $1.3 \pm 0.3\text{ J}\cdot\text{K}^{-1}\cdot\text{g}^{-1}$  que corresponde al promedio de diferentes bases de datos de proteínas dos-estados, y que tradicionalmente se ha aceptado como un valor universal para las proteínas globulares (Privalov & Khechinashvili 1974; Freire 1995). Este valor se encuentra, asimismo, en el límite superior del intervalo de capacidades caloríficas para el estado nativo obtenido en estadísticas más recientes utilizando un conjunto más amplio de proteínas globulares, que se encuentran entre los 1.25 y los 1.8  $\text{J}\cdot\text{K}^{-1}\cdot\text{g}^{-1}$  (Makhatadze 1998). Este elevado valor de la capacidad calorífica, al igual que los valores de  $T_S$  o el  $\Delta h(110\text{ °C})$ , está en concordancia con un estado nativo que presente una importante flexibilidad conformacional.

Hay importantes diferencias entre las representaciones de la función de capacidad calorífica del estado nativo teórica (Freire 1995) y las determinadas en los ajustes de los datos experimentales (Figuras 2.4 y 2.5). La aparición de líneas base por encima de los valores típicamente observados para otras proteínas globulares, se ha descrito previamente para otros sistemas, tales como ciertas cremallera de leucina (Dragan & Privalov 2002; Privalov & Dragan 2007), las proteínas BBL (Naganathan et al. 2005) y gpW (Fung et al. 2008). Sin embargo, en estos casos, además de un valor elevado de  $C_{p,N}(20\text{ °C})$ , se observó también una fuerte dependencia de la capacidad calorífica del estado nativo con la temperatura.

Teniendo en cuenta que las líneas base pre- y post- transición están estrechamente correlacionadas con las fluctuaciones de entalpía de los estados macroscópicos (Privalov & Khechinashvili 1974; Freire et al. 1993; Makhatadze & Privalov 1995; Robertson & Murphy 1997; Privalov & Dragan 2007), estos valores anormalmente altos han sido interpretados en términos de un equilibrio más complejo que el de dos estados. En el caso concreto de las cremalleras de leucina, se ha establecido que este comportamiento responde a un equilibrio multiestados, mientras que en el caso de las proteínas gpW y BBL se ha atribuido a una baja cooperatividad en el equilibrio conformacional, de modo que BBL se ha propuesto como un claro ejemplo de un plegamiento sin barreras, “downhill” (Naganathan et al. 2005), y gpW parece estar en la línea divisoria entre un plegamiento de dos estados y un sistema “downhill” (Fung et al. 2008). En estos casos, la fuerte dependencia de la capacidad calorífica del estado nativo con la temperatura refleja el cambio progresivo en las propiedades del conjunto de conformaciones en equilibrio (Naganathan et al. 2005). En nuestro caso, aunque hNedd4-WW4 presenta valores de  $C_{p,N}(T)$  anormalmente altos, la pendiente de la función capacidad calorífica es la esperada para una proteína de este tamaño. Así, como se puede observar en la Figura 2.5, la función  $C_{p,N}(T)$  teórica (Freire 1995) es paralela a la obtenida en el ajuste al modelo de dos estados. La pendiente nos indica que las propiedades conformacionales del dominio hNedd4-WW4 evolucionan con el incremento de temperatura de igual manera que teóricamente lo hace un único macroestado nativo de una proteína del mismo tamaño que despliega según el modelo de dos estados.

Una explicación alternativa para el valor anómalo de la  $C_{p,N}(20\text{ °C})$  experimental podría ser la presencia de otro estado plegado nativo con una conformación diferente, lo cual también sería coherente con la dispersión hallada para los valores de  $T_m$  y la ausencia de un punto de isobélico (Arbely et al. 2009; Settanni & Fersht 2009). En este caso, el equilibrio entre los diferentes estados nativos podría dar lugar a una contribución en el cambio de entalpía que justificarían los elevados valores para  $C_{p,N}(T)$  (Settanni & Fersht 2009). Una situación parecida se ha descrito para otro dominio WW, el dominio hFbp28-WW, donde el cambio de conformación en el estado nativo está provocado por un único

residuo, como respuesta a la presencia de diferentes tipos de ligandos (Karanicolas & Brooks 2004).

En resumen, a pesar de que los datos experimentales se pueden describir en términos de un equilibrio de dos estados, el análisis del desplegamiento térmico del dominio hNedd4-WW4 revela varias características de los espectros y valores anómalos para algunos de los parámetros termodinámicos que son indicios de un proceso más complejo, como puede ser el desplegamiento sin barreras o la presencia de otras conformaciones alternativas del estado nativo.

### **2.2.3. Experimentos de doble perturbación**

Para investigar más a fondo la naturaleza del equilibrio conformacional de hNedd4-WW4, se ha analizado el acoplamiento de la desnaturalización térmica y química. Estos experimentos de doble perturbación se han presentado como una herramienta adecuada para discernir entre los diferentes tipos de equilibrios (Oliva & Munoz 2004).

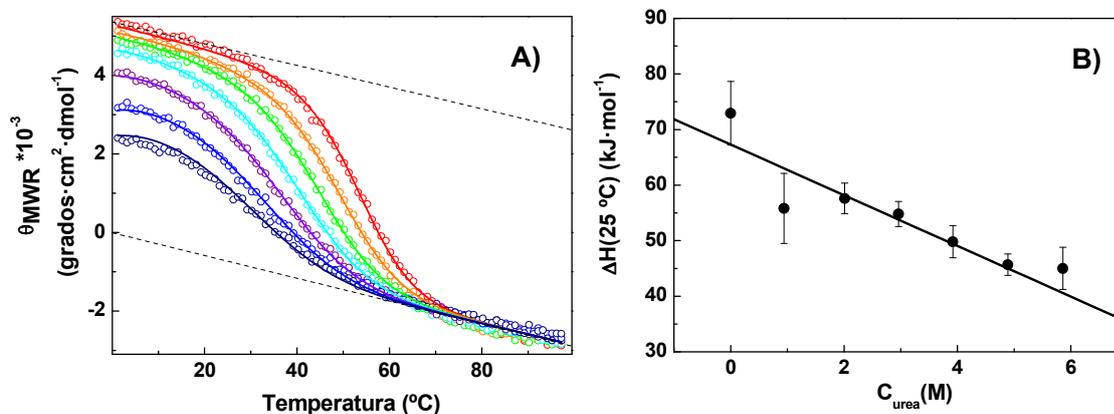
En una situación de plegamiento limitado por una barrera energética, el desplegamiento térmico seguido mediante un observable local de la estructura de la proteína tendrá una forma sigmoidal, la señal cambia gradualmente a medida que lo hace el grado de estructura. En este caso, las contribuciones del estado nativo y del estado desplegado a los parámetros termodinámicos obtenidos del análisis de desplegamiento (entalpía y entropía) estarán relacionadas de forma lineal. La presencia de un agente desnaturalizante, como es la urea, hace que disminuya la temperatura del punto medio de la transición de desplegamiento,  $T_m$ , sin modificar propiedades conformacionales del estado nativo (y de este modo, manteniendo la línea base pre-transición) hasta que la desnaturalización por frío comienza a ser detectable (Cobos et al. 2002; Felitsky & Record 2003).

En un equilibrio sin barrera ("downhill"), las curvas de desplegamiento también presentan una forma sigmoidal, aunque menos pronunciada, y con importantes variaciones en las líneas base pre- y post-transición, como una disminución de la señal del estado nativo (línea base pre-transición) incluso en condiciones de elevada estabilidad (Oliva & Munoz 2004). También se ha propuesto que dicho cambio continuo de la distribución conformacional produce

una relación no lineal entre los parámetros termodinámicos (Naganathan et al. 2005), así como la intersección de las líneas base fenomenológicas derivadas del análisis dentro del rango de temperatura del experimento. La intersección de las líneas base es incompatible con cualquier proceso que implique una barrera de energía, ya que éstas reflejan las propiedades estructurales del conjunto, y su cruce implicaría que los estados nativo y desplegado tendrían el mismo contenido estructural a la temperatura de intersección (Oliva & Munoz 2004).

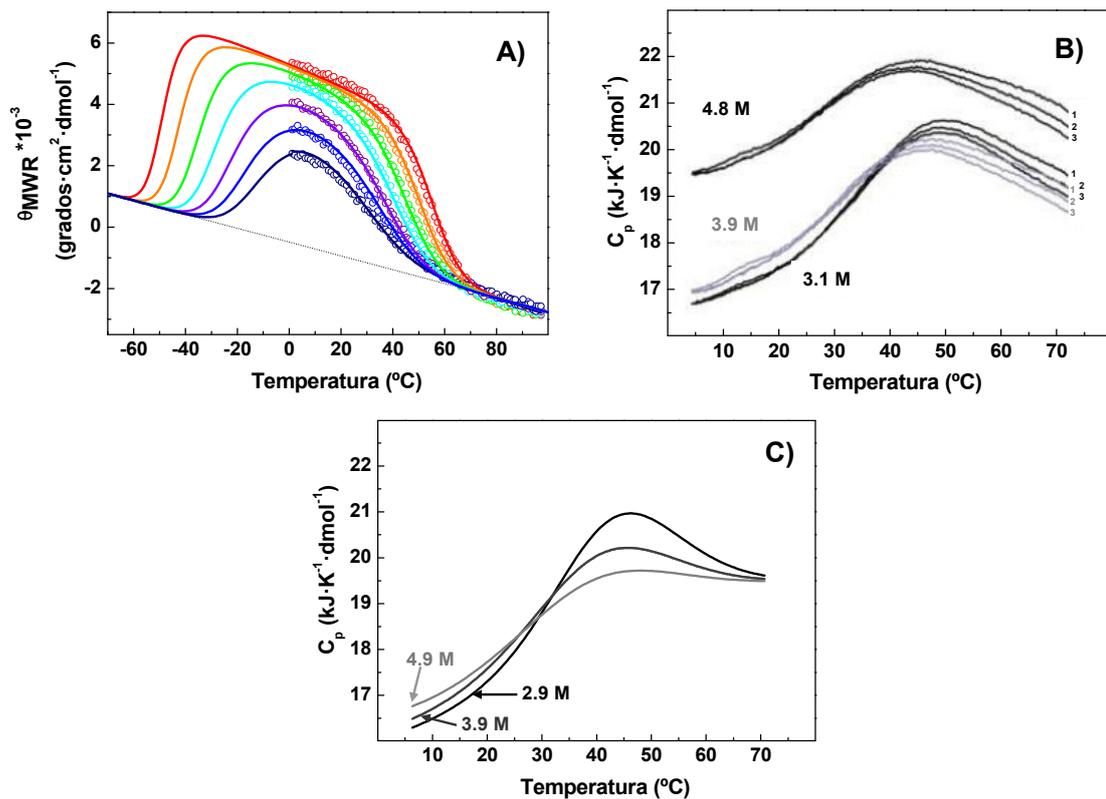
Los resultados de los experimentos de doble perturbación llevados a cabo con el dominio hNedd4-WW4 se resumen en la Figura 2.7. En el Panel A se muestran los perfiles de desnaturalización térmica seguidos por CD a 230 nm correspondientes a concentraciones de urea entre 0 M y 6 M de urea en tampón fosfato sódico 20 mM a pH 7.0. La característica más sobresaliente de este conjunto de curvas de desnaturalización térmica es una disminución progresiva de la elipticidad molar por resto para el estado nativo a bajas temperaturas concomitante con el aumento en la concentración de urea. Esto, en principio, podría ser indicativo de un plegamiento tipo “downhill” (Oliva & Munoz 2004), en cuyo caso reflejaría el cambio de propiedades estructurales de la distribución conformacional del estado nativo con la concentración de urea. No obstante el ajuste de los datos experimentales utilizando el modelo de dos estados reproduce correctamente estos efectos, como se puede ver en la Figura 2.7A. A raíz de este análisis, parece que nuestros datos son compatibles con el inicio de la desnaturalización por frío en el contexto de un equilibrio de dos estados. Además, las líneas base obtenidas en este análisis son paralelas, lo que indica que la dependencia con la temperatura del estado nativo y desnaturalizado son similares y, por tanto, no se cruzan dentro del rango de temperaturas experimentales (Nicholson & Scholtz 1996; Knapp et al. 1998; Cobos et al. 2002; Felitsky & Record 2003; Oliva & Munoz 2004; Naganathan et al. 2005). Los valores de  $T_m$  y  $\Delta H_m$  presentan una muy buena correlación lineal [ $r = 0.995$  (datos no mostrados)] con la concentración de urea, y el cambio de capacidad calorífica es comparable al obtenido del estudio de la desnaturalización térmica en distintas condiciones de pH (Tabla 2.1). Tampoco se observan grandes desviaciones del comportamiento lineal de la entalpía de desplegamiento de hNedd4-WW4 a 25 °C con la concentración

de urea (Figura 2.7B), como se ha propuesto para aquellas proteínas cuyo desplegamiento se produce sin barreras de energía (Oliva & Munoz 2004; Naganathan et al. 2005; Fung et al. 2008; Naganathan et al. 2010).



**Figura 2.7. Dependencia de la desnaturalización térmica seguida por CD (a 230 nm) en presencia de distintas concentraciones de urea para el dominio hNedd4-WW4. A)** En círculos vacíos se han representado los datos experimentales en fosfato sódico 20mM pH 7.0, en presencia de urea, atendiendo al siguiente código de colores: 0 M urea en rojo, 0.94 M urea en naranja, 2.01 M urea en verde, 2.96 M urea en magenta, 3.92 M urea en purpura, 4.89 M urea en azul y 5.98 M en azul marino. Las líneas continuas, que acompañan los datos experimentales, corresponden al mejor ajuste global según el modelo de dos estados (ecuación II.IX, Apéndice II) compartiendo las líneas base del estado nativo y del estado desnaturalizado para cada conjunto de datos. Las líneas rectas discontinuas representan dichas líneas base obtenidas a partir de la estrategia de ajuste global para los estados nativo y desplegado en cada conjunto de experimentos de desnaturalización térmica. **B)** Representación de los valores de entalpía de desplegamiento a extrapolados a 25  $^{\circ}\text{C}$  frente a la concentración de urea. La línea negra continua corresponde al ajuste lineal de los datos. Las barras representa la propagación de errores del ajuste.

Para distinguir claramente entre los dos posibles comportamientos durante el plegamiento en el caso del dominio hNedd4-WW4, sería deseable obtener una evidencia experimental de la desnaturalización fría pronosticada por el análisis de dos estados de los datos de CD, como se ilustra en la simulación de la Figura 2.8A. Para ello, hemos llevado a cabo varios experimentos DSC en diferentes concentraciones de urea, que se muestran en el Panel B la Figura 2.8. Sin embargo, estos experimentos no aportan evidencias de que se produzca una transición a bajas temperaturas. Lo que se observa es un cambio en la  $C_{p,N}(T)$  hacia valores más altos coincidiendo con el aumento de la concentración de urea (9-10  $\text{kJ} \cdot \text{K}^{-1} \cdot \text{mol}^{-1}$  en ausencia de urea y 20  $\text{kJ} \cdot \text{K}^{-1} \cdot \text{mol}^{-1}$  en 4.8 M de urea), y que es paralelo al cambio en la señal de estado nativo que presentan las curvas de CD (Figura 2.7A). Como se puede apreciar en el panel C de la Figura 2.8, la



**Figura 2.8. Simulaciones y dependencia con la urea de la desnaturalización térmica del dominio hNedd4-WW4 en tampón fosfato 20 mM, pH 7.0.** **A)** Simulación del perfil de desnaturalización térmica entre  $-70\text{ }^{\circ}\text{C}$  y  $100\text{ }^{\circ}\text{C}$  en presencia de diferentes concentraciones de urea de 0, 0.94, 2.01, 2.96, 3.92, 4.89 y 5.98 M (líneas continuas) junto con los datos experimentales de CD registrados a una longitud de onda de 230 nm (círculos abiertos). Ambos se han representado de acuerdo con el código de colores empleado en la Figura 2.7. **B)** Curvas normalizadas de DSC obtenidas a concentraciones intermedias de urea. Se muestran tres calentamientos consecutivos de cada muestra en presencia de 3.1, 3.9 y 4.8 M de urea realizados hasta  $70\text{ }^{\circ}\text{C}$  para evitar efectos térmicos asociados a fenómenos de degradación de la urea. **C)** Simulación de los perfiles de capacidad calorífica esperados para distintos experimentos de DSC en presencia de 2.9, 3.9 y 4.9 M urea. Se representan tres curvas teniendo en cuenta los parámetros termodinámicos obtenidos a partir del ajuste global al modelo de dos estados del experimento de doble perturbación seguido por CD (Figura 2.7).

simulación de los perfiles de capacidad calorífica esperados según los parámetros obtenidos en el experimento de doble perturbación por CD proporcionan trazas calorimétricas muy similares a las obtenidas experimentalmente (Figura 2.8B). Esta simulación nos indica que los experimentos no son concluyentes porque cuando se estudian proteínas pequeñas con transiciones anchas y poco cooperativas, aún tratándose de un equilibrio de dos estados, sería muy difícil observar experimentalmente un aumento claro de la capacidad calorífica a baja temperatura asociado a la desnaturalización fría, tal como se ha descrito para otras

proteínas dos estados mucho más cooperativas (Privalov 1990; Azuaga et al. 1992).

En resumen, las dependencias lineales de la entalpía de desplegamiento con la temperatura y la concentración de urea indican, en principio, un equilibrio de dos estados. No obstante, no hemos encontrado pruebas concluyentes de la existencia de una desnaturalización a bajas temperaturas y los experimentos de doble perturbación, muy similares a los descritos para proteínas con desnaturalización tipo “downhill” (Oliva & Munoz 2004; Naganathan et al. 2005; Fung et al. 2008), indican que el nivel de cooperatividad en el equilibrio es muy bajo. En este sentido, es importante señalar que, aunque el modelo de dos estados describe el comportamiento experimental de manera razonable, los parámetros termodinámicos obtenidos para el dominio hNedd4-WW4 no son propios de proteínas dos estados, sino más similares a los de proteínas “downhill” como BBL (Naganathan et al. 2005) y gpW (Fung et al. 2008).

En conjunto, los datos de hNedd4-WW4 indican que el equilibrio conformacional de este dominio parece estar en el límite de lo que podría considerarse como un equilibrio cooperativo de dos estados y parecen reflejar una disminución de la barrera de energía entre macro-estados inducida por cambios en las condiciones experimentales, tales como la presencia de desnaturalizantes químicos. En este sentido, un reciente estudio computacional con el dominio WW de la proteína Pin1 ha llegado a conclusiones similares, sugiriendo que este dominio podría exhibir un comportamiento “downhill” a temperaturas inferiores a la fisiológica (Bruscolini et al. 2007). Estos efectos a baja temperatura son similares a los observados con el dominio hNedd4-WW4 a consecuencia de la adición de urea.

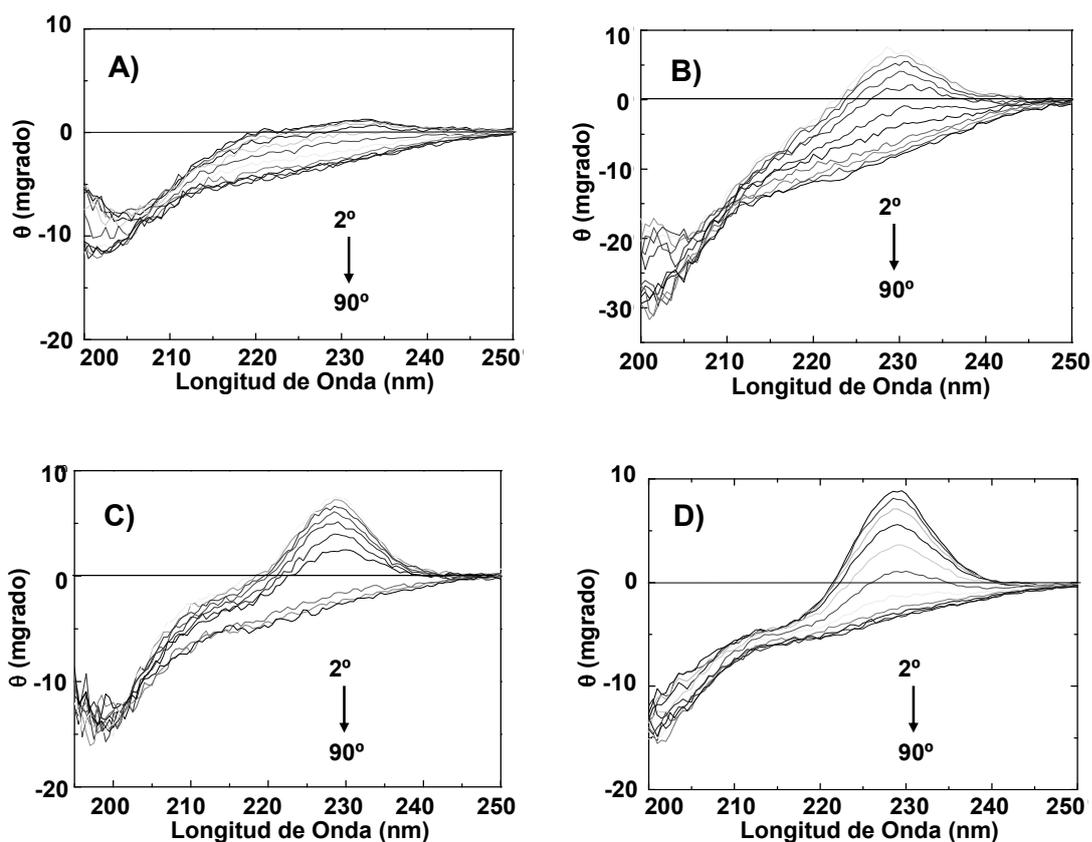
Para profundizar en el estudio del desplegamiento de este dominio WW y determinar claramente la naturaleza de su equilibrio conformacional estamos realizando estudios adicionales, en colaboración con el grupo de investigación del Dr. Víctor Muñoz, para estimar la magnitud de la barrera de plegamiento, tanto a partir del análisis de los datos calorimétricos con el modelo de barrera variable (Munoz & Sanchez-Ruiz 2004) como mediante estudios cinéticos detallados (Naganathan et al. 2005; Fung et al. 2008; Li et al. 2009; Naganathan et al. 2010).

### **2.3. Caracterización de otros dominios WW**

La caracterización y comparación entre los experimentos cinéticos y de equilibrio de dominios WW homólogos, así como de otras proteínas marginalmente estables, brinda la oportunidad de diseccionar experimentalmente los factores determinantes en la secuencia, en la estructura, y en la evolución de las barreras de energía de plegamiento y de la cooperatividad para la estructura de tipo  $\beta$ . El uso de dominios homólogos ofrece la ventaja de la comparación de las proteínas naturales con diferencias en la secuencia debido a la evolución divergente. De hecho, en un trabajo reciente realizado con el dominio BBL se ha propuesto que la combinación del análisis cinético y del equilibrio conformacional de otros homólogos estructurales con cinéticas de plegamiento ultrarrápidas proporciona una prueba empírica de la validez y de la sensibilidad de las diferentes características del equilibrio que se han propuesto para el plegamiento tipo “downhill” (Naganathan et al. 2010). Por tanto, en base a estas ideas y dado que el dominio hNedd4-WW4 parece estar en el límite de lo que se considera como un equilibrio de dos estados, hemos ampliado la caracterización del equilibrio conformacional a un conjunto de dominios WW de distintas procedencias, entre los que se encuentra un dominio homólogo de hNedd4-WW4. Por un lado, hemos seleccionado otros dos dominios WW del grupo I de especificidad, además de hNedd4-WW4, su homólogo hNedd4-WW3 y el dominio hYap1-WW1(L30K), y por otro lado, otros dos dominios homólogos del grupo II, los dominios WW1 y WW2 de la proteína hFbp11. En la Figura 2.1A, se han recogido los alineamientos de las secuencias correspondientes a ambos conjuntos de dominios, para los cuales la similitud de secuencia obtenida es bastante elevada (65%).

#### **2.3.1. Espectros de dicroísmo circular**

En la Figura 2.9 se presentan los espectros de CD de cada dominio WW en el rango del UV-lejano a temperaturas entre 2 y 98 °C. Como podemos observar, no todos los espectros de CD de estos dominios WW presentan exactamente el mismo aspecto. El máximo característico a 230 nm (Figura 2.2A) presenta diferente intensidad en los distintos dominios: en los espectros de hNedd4-WW3 ni siquiera se aprecia dicho pico, en los espectros del dominio hYap-WW1(L30K)



**Figura 2.9. Espectros de dicroísmo circular en UV-lejano de distintos dominios WW.** Se han registrado diez espectros en fosfato sódico 20 mM pH 7.0 en el intervalo de temperaturas de 2° a 90 °C. **A)** Dominio hNedd4-WW3, **B)** dominio hYap-WW1(L30K), **C)** dominio hFbp11-WW1 y **D)** dominio hFbp11-WW2.

tiene una intensidad muy baja, y en los dominios WW1 y WW2 de hFbp11 es una banda muy intensa. De acuerdo con la bibliografía (Koepf et al. 1999a; Koepf et al. 1999b; Ferguson et al. 2001a; Jager et al. 2001; Jager et al. 2007), estos efectos en el máximo a 230 nm se relacionan con diferencias en el empaquetamiento del W17 en el corazón hidrofóbico de los dominios WW. Por otra parte, todos estos dominios WW muestran la banda intensa y de valor negativo en torno a 200nm, que es muy similar a la observada para hNedd4-WW4 y a la descrita para otras proteínas que contienen láminas  $\beta$  distorsionadas (Viguera et al. 1994; Knapp et al. 1998; Reid et al. 1998). No obstante, como se aprecia en la Figura 2.9, este pico presenta una menor dependencia con la temperatura en los dos dominios WW de la proteína hFbp11. Por último, del mismo modo que sucede con el dominio hNedd4-WW4, en los espectros de CD a distintas temperaturas de estos cuatro

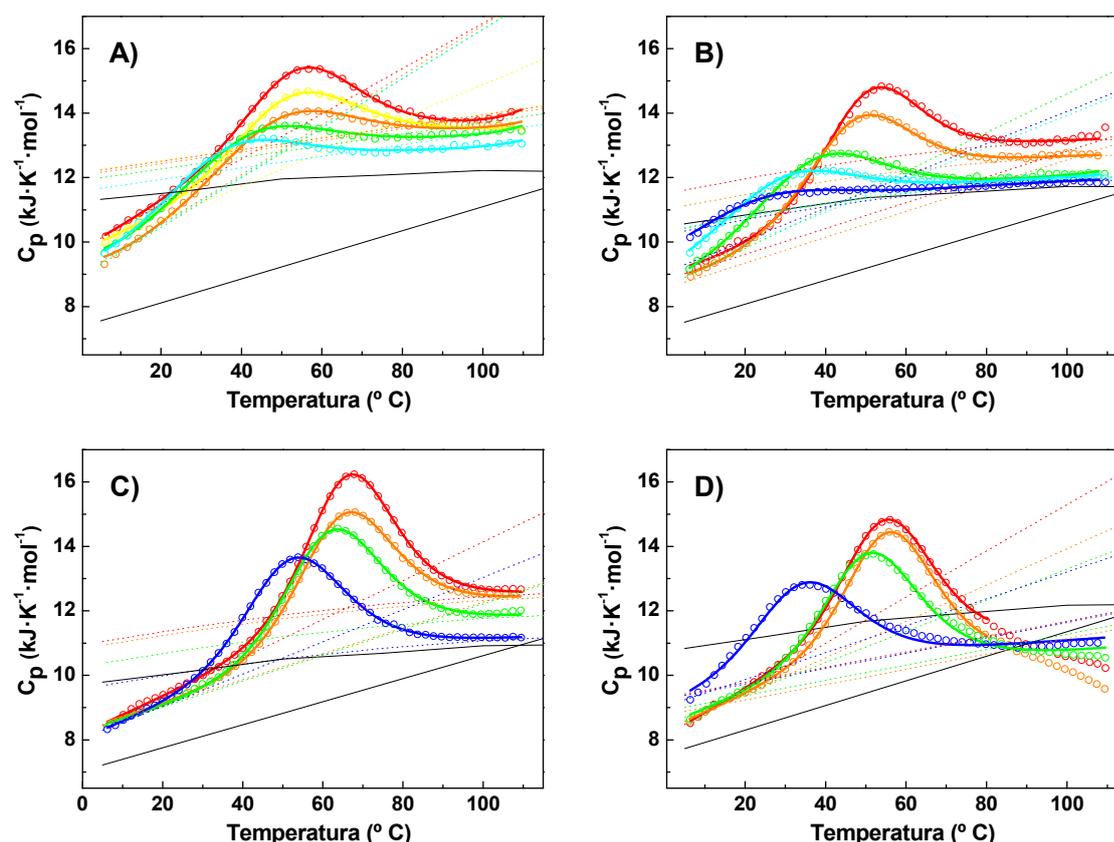
dominios tampoco se observa el punto isobéptico (o isodicroíco) característico de las proteínas dos estados.

### 2.3.2. Desnaturalización térmica en distintas condiciones de pH

En la Figura 2.10 se presentan los perfiles de DSC para el desplegamiento térmico de los dominios hNedd4-WW3 (panel A), hYap-WW1 (L30K) (panel B), hFbp1-WW1 (panel C) y hFbp1-WW2 (panel D), en distintas condiciones de pH. Tras comprobar que cada proceso de desplegamiento es reversible, se realizó el ajuste individual al modelo de dos estados (ecuación I.X en el Apéndice I) de cada una de las curvas. Las únicas excepciones para las que no se ha podido realizar el análisis de las curvas completas, han sido los experimentos a pH 7.0 y 5.0 del dominio hFbp11-WW2 (Figura 2.10D), porque presentan una disminución brusca en la  $C_p$  por encima de los 80 °C. El proceso de desplegamiento es irreversible por encima de esta temperatura incluso en presencia de agentes reductores. En este caso se ha podido comprobar que las trazas presentan una reversibilidad más elevada en ambas condiciones de pH, si se calienta justo hasta completar la transición principal (70-80 °C), lo que confirma que se trata de una agregación post-transición. Estos fenómenos de agregación del dominio hFbp11-WW2 se pueden explicar adecuadamente como un proceso irreversible por medio del modelo Lumry-Eyring ( $N \rightleftharpoons D \rightarrow F$ ) y se sabe que afectan mínimamente a los parámetros termodinámicos obtenidos a partir de la endoterma principal (Sanchez-Ruiz 1992; Martinez et al. 1994). Por tanto, en éste y en los próximos análisis que se realicen, se han considerado los datos experimentales de estas dos trazas cortándolas en 80 °C.

El ajuste individual de cada endoterma de desplegamiento de acuerdo con el modelo de dos estados reproduce perfectamente los datos experimentales para los cuatro dominios WW (con valores de  $r$  y  $r^2$  entre 0.999 y 0.99), como se observa en la Figura 2.10. Sin embargo, al igual que sucede con el dominio hNedd4-WW4 (Figura 2.4), las líneas base que se obtienen de estos ajustes para el estado nativo,  $C_{p,N}(T)$ , y desplegado,  $C_{p,D}(T)$ , no son del todo realistas, pues se produce su intersección a temperaturas cercanas a la  $T_m$ , análogamente a lo que sucede en el análisis individual de las curvas de DSC del dominio hNedd4-WW4

(Figura 2.4). En el caso de las funciones para la  $C_{p,N}(T)$  obtenidas en los ajustes individuales a distintos pH para cada dominio tampoco se parecen a la función lineal teórica estimada a partir del peso molecular y típicamente aceptada para proteínas globulares (Freire 1995).



**Figura 2.10. Dependencia de la capacidad calorífica con la temperatura de cuatro dominios WW distintos en diferentes condiciones de pH. A) Dominio hNedd4-WW3. B) Dominio hYap-WW1(L30K). C) Dominio hFbp11-WW1. D) Dominio hFbp11-WW2.** Los datos se han representado como círculos vacíos con diferente color según la condición de pH: rojo para fosfato sódico 20 mM pH 7.0, amarillo para MES 20 mM pH 6.0, naranja para acetato sódico 20 mM pH 5.0, verde para acetato sódico 20 mM pH 4.0, azul cian para glicina 20 mM pH 3.5 y azul para glicina 20 mM pH 3.0. El mejor ajuste individual, considerando funciones  $C_{p,N}(T)$  y  $C_{p,D}(T)$  para cada curva según el modelo de dos estados, se ha mostrado como una línea continua acompañando a los datos experimentales en el mismo código de colores. Las líneas punteadas en los mismos colores corresponden a las funciones  $C_{p,N}(T)$  y  $C_{p,D}(T)$  obtenidas a partir del ajuste para cada curva. Las líneas continuas en negro corresponden a la función  $C_{p,D}(T)$  calculada mediante la adición de las contribuciones de cada aminoácido (Privalov & Makhatadze 1990) (arriba) y a la función teórica  $C_{p,N}(T)$  estimada a partir del peso molecular del dominio (Freire 1995) (abajo).

En cada conjunto de curvas de DSC a distintos pH para estos dominios se aprecia un aumento progresivo de la capacidad calorífica a alta temperatura con el pH. Las funciones de  $C_{p,D}(T)$  a pH 3.0 de cada dominio coinciden prácticamente

con la función teórica calculada a partir de los valores tabulados para cada aminoácido (Privalov & Makhatadze 1990), salvo en el caso de los experimentos de hFbp11-WW2, que muestran valores de  $C_{p,D}(T)$  ligeramente por debajo de la predicción teórica debido a la distorsión de la traza por los procesos irreversibles post-transición. Al aumentar el pH, se observa un incremento en el valor de la función  $C_{p,D}(100\text{ °C})$  de 1 a  $2\text{ kJ}\cdot\text{K}^{-1}\cdot\text{mol}^{-1}$  entre pH 3.0 y pH 7.0, dependiendo del dominio (Figura 2.10).

La variación de la función  $C_{p,D}$  con el pH está posiblemente relacionada con las diferentes contribuciones a la capacidad calorífica de los distintos estados de ionización de las cadenas laterales cargadas (Gomez et al. 1995). En el intervalo de pH de los experimentos, entre pH 7.0 y 3.0, son cuatro los grupos susceptibles de variar su estado de ionización. Además del grupo indol de la histidina, están los grupos carboxílicos del extremo amino-terminal, de la cadena lateral del ácido aspártico y del ácido glutámico. En la Tabla 2.2, se muestran varias estimaciones

**Tabla 2.2. Cálculo teórico del cambio de la capacidad calorífica correspondiente a la protonación del carboxilo terminal y de las cadenas laterales de los restos histidina, ácido glutámico y ácido aspártico.** Para el cálculo se han tenido en cuenta el número de restos presentes en los dominios hNedd4-WW3, hYap1-WW1 y hFbp11-WW1. Se han utilizado los siguientes valores de cambio de capacidad calorífica y entalpía de protonación:  $125.4\text{ J}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$  y  $-2\text{ J}\cdot\text{mol}^{-1}$  respectivamente, para los grupos carboxílicos; y  $16.7\text{ J}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$  y  $29.6\text{ J}\cdot\text{mol}^{-1}$  respectivamente, para los grupos indol (Gomez et al. 1995)

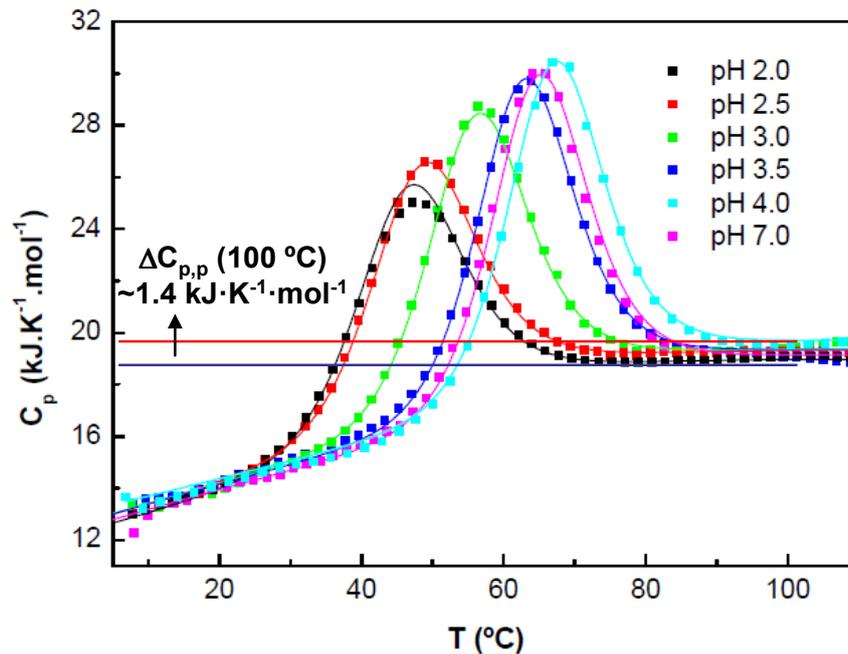
Proteína	Grupos ionizables	$F_p \cdot \Delta C_{p,\text{prot}}, F_{p,\text{todos}}=1$ ( $\text{J}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$ )	$*F_p \cdot \Delta C_{p,\text{prot,máximo}}, F_{p,\text{His}}=0.5$ ( $\text{J}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$ )
hNedd4-WW3	His	3	50.1
	Glu	4	501.6
	Asp	2	250.8
	C-terminal	1	125.4
	<b>Total</b>		927.9
hYap1-WW1 (L30K)	His	1	16.7
	Glu	2	250.8
	Asp	4	501.6
	C-terminal	1	125.4
	<b>Total</b>		894.5
hFbp11-WW1	His	1	16.7
	Glu	3	376.2
	Asp	4	376.2
	C-terminal	1	125.4
	<b>Total</b>		894.5

\*Para calcular el  $\Delta C_{p,\text{prot}}$  máximo se ha considerado  $F_p = 1$  para todos los grupos ionizables menos para las histidinas porque a pH 7, su  $pK_a \sim \text{pH}$ . Por tanto, el valor para el grado de protonación es  $F_p = 0.5$ , y el valor de  $C_{p,\text{prot}}$  asociado a la fluctuación térmica de la protonación se ha calculado de acuerdo con la relación:  $F_p(1-F_p)\Delta H_p^2/RT^2$  (Gomez et al. 1995).

de los valores de capacidad calorífica de protonación  $C_{p,prot}$  que caben esperar para nuestros sistemas cuando el factor (o grado) de protonación total,  $F_p$ , para estos grupos en cada dominio es 1 ó 0.5. Cuando la fracción de histidinas ionizables es 0.5, la contribución de  $C_{p,prot}$  es máxima debido a las grandes fluctuaciones entálpicas asociadas a la ionización de su grupo indol (Gomez et al. 1995). Si bien es cierto que la contribución debida al efecto de la ionización sobre el solvente es equivalente pero de signo opuesto al estimado para los grupos carboxílicos, aún así los valores de  $C_{p,prot}$  asociados a las fluctuaciones entálpicas de la ionización de las histidinas (Tabla 2.2) son pequeños, pero del mismo orden de magnitud que las desviaciones de  $C_p$  con el pH observadas para cada dominio en la Figura 2.10.

Estas contribuciones asociadas a la protonación de restos cargados son universales para todas las proteínas. Normalmente, los grupos cargados de suelen estar en la superficie de la proteína donde la constante dieléctrica es relativamente baja (Sturtevant 1977), por tanto, el grado de exposición de los grupos cargados al disolvente no cambia mucho en los procesos desplegamiento y no se suelen considerar explícitamente en el análisis de los perfiles de DSC. Sin embargo, en proteínas de menor tamaño y poco cooperativas como son estos dominios WW, la magnitud de este efecto no es despreciable con respecto a los cambios de capacidad calorífica asociados a las transiciones de desplegamiento. Con el fin de ilustrar esto, en la Figura 2.11 se recoge la dependencia con la temperatura de la  $C_p$  a distintas condiciones de pH, entre 2.0 y 7.0, medidas en nuestro grupo de investigación para una proteína de mayor tamaño, donde se pueden observar variaciones ( $\Delta C_{p,prot}$  a 100 °C) de magnitud similar a las observadas en las transiciones de los dominios WW que no se han considerado en su análisis.

Para analizar adecuadamente las líneas base de las curvas de DSC obtenidas a diferentes valores de pH hemos realizado un ajuste global de acuerdo con el modelo de dos estados (Apéndice I). En este análisis, las funciones de capacidad calorífica tanto del estado nativo como del estado desnaturalizado se han considerado comunes a todas las curvas. Además, para tener en cuenta la protonación de las cadenas laterales, se ha añadido un término  $F_p \cdot \Delta C_{p,prot}(100 \text{ °C})$  al cambio de la capacidad calorífica:  $\Delta C_p = (a-d) + (b-e) \cdot T + c \cdot T^2 + F_p \cdot \Delta C_{p,prot}(100 \text{ °C})$ . El

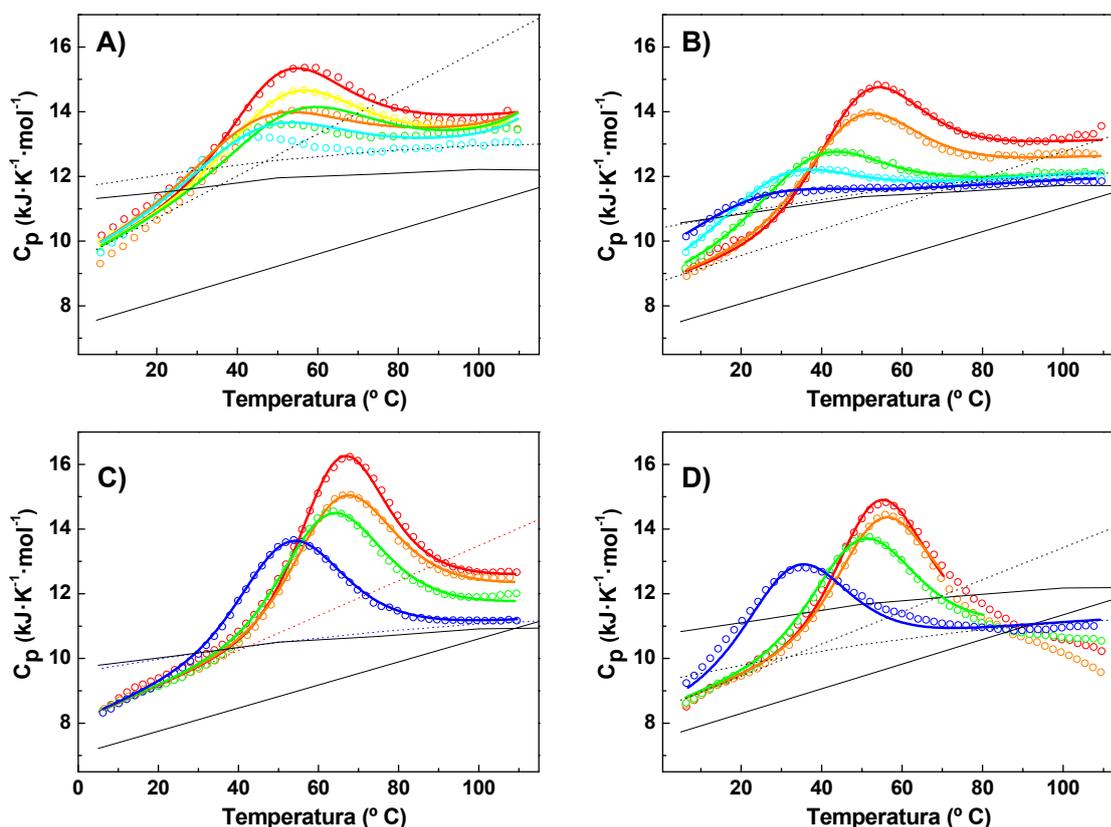


**Figura 2.11. Efecto de la protonación de las cadenas laterales en la  $C_p$  para el desplegamiento de la proteína SPCp41.** Dependencia de la capacidad calorífica para la desnaturalización térmica de la proteína química SPCp41 basada en el permutante circular de el dominio SH3 de  $\alpha$ -espectrina y el péptido p41 (76 aminoácidos) (figura tomada de Martin-Sierra et al. 2003). Las líneas continuas se corresponden al mejor ajuste por mínimos cuadrados no lineales al modelo de dos estados, y las líneas rectas horizontales indican en el eje de ordenadas los valores de  $C_p(100\text{ °C})$  a pH neutro (en color rojo) y ácido (en color azul).

grado de protonación,  $F_p$ , es una de las variables del ajuste global para cada condición de pH y la contribución máxima debida a la protonación de cadenas laterales en todo el rango de pH estudiado,  $\Delta C_{p,prot}(100\text{ °C})$ , una constante que se ha estimado directamente de los perfiles de DSC como la diferencia del valor de  $C_p$  a  $100\text{ °C}$  entre la curva a pH 7.0 y la curva a pH 3.0.

En la Figura 2.12 se puede comprobar cómo este análisis global de acuerdo con el modelo de dos estados describe adecuadamente las curvas de DSC de todos los dominios, menos en el caso del dominio hNedd4-WW3, que muestra importantes desviaciones para las curvas menos cooperativas a pH ácido. Las funciones de  $C_{p,N}(T)$  obtenidas en este ajuste global presentan valores más elevados que las estimadas a partir del peso molecular (Freire 1995), no solo para la ordenada en el origen como se ha descrito para el análisis individual de las curvas de DSC de hNedd4-WW4 (Figura 2.5), sino que ahora también sus pendientes son bastante más pronunciadas en los perfiles de DSC de hNedd4-

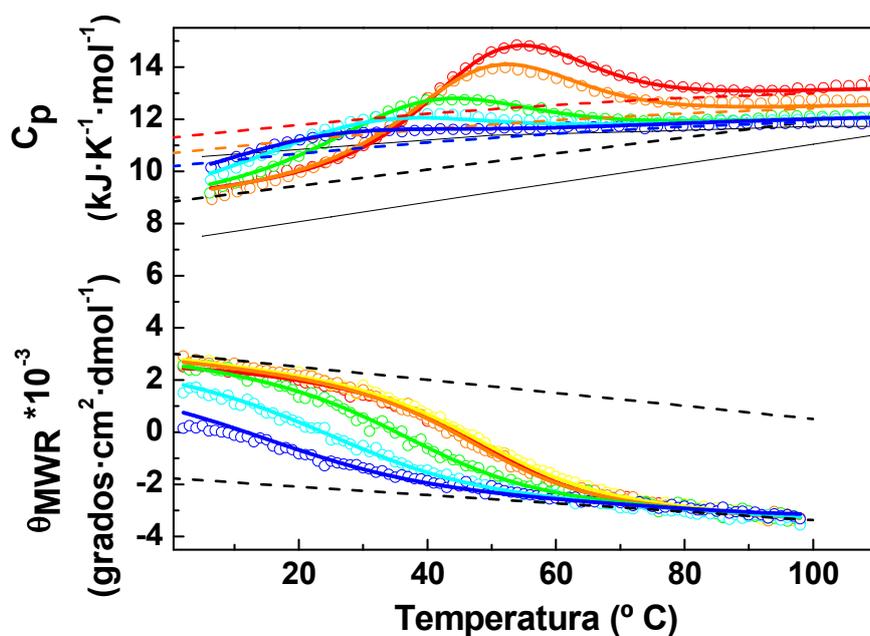
WW3, hFbp11-WW1 y hFbp11-WW2. Esta elevada pendiente de la función  $C_{p,N}(T)$  sigue produciendo una intersección con la función  $C_{p,D}(T)$  a temperaturas poco realistas, muy por debajo de los 120 °C convenidos como valor de convergencia para las proteínas globulares (Privalov & Dragan 2007).



**Figura 2.12. Dependencia de la capacidad calorífica con la temperatura en diferentes condiciones de pH. A) Dominio hNedd4-WW3. B) Dominio hYap-WW1(L30K). C) Dominio hFbp11-WW1. D) Dominio hFbp11-WW2.** Los datos se han representado como círculos vacíos según la condición de pH: rojo para fosfato sódico 20 mM pH 7.0, amarillo para MES 20 mM pH 6.0, naranja para acetato sódico 20 mM pH 5.0, verde para acetato sódico 20 mM pH 4.0, azul cian para glicina 20 mM pH 3.5 y azul para glicina 20 mM pH 3.0. El mejor ajuste global según el modelo de dos estados se muestra como una línea continua en el mismo código de colores. En el ajuste global se consideran las funciones  $C_{p,N}(T)$  y  $C_{p,D}(T)$  comunes a todas las curvas de cada dominio y un término adicional  $F_p \cdot \Delta C_{p,prot}(100^\circ\text{C})$  para la protonación de las cadenas laterales. Las líneas punteadas en color negro corresponden a dichas funciones  $C_{p,N}(T)$  y  $C_{p,D}(T)$ . Las líneas continuas en color negro corresponden a la función  $C_{p,D}(T)$  calculada mediante la adición de las contribuciones de cada aminoácido (Privalov and Makhatadze 1990) (arriba) y a la función teórica  $C_{p,N}(T)$  estimada a partir del peso molecular del dominio (Freire 1995) (abajo)

Para evaluar si el proceso de desplegamiento de estos dominios WW tiene el mismo comportamiento descrito para el dominio hNdd4-WW4 hemos obtenido sus parámetros termodinámicos a partir de diferentes observables que describen distintos aspectos conformacionales de la proteína, como son la señal

espectroscópica de CD a 230 nm y la variación de la capacidad calorífica mediante DSC (Ginsburg & Carroll 1965; Sanchez-Ruiz 1995). Como hemos descrito para el dominio hNedd4-WW4, las curvas de desplegamiento térmico de DSC y CD de cada dominio se analizaron mediante un ajuste múltiple asumiendo líneas base comunes y valores de  $\Delta H_m$  y  $T_m$  comunes para cada condición de pH. Como se puede apreciar en la Figura 2.13, este análisis ha proporcionado una descripción razonable de los datos experimentales para el dominio hYap-WW1(L30K), aunque los ajustes del resto de dominios WW presentaron desviaciones importantes, lo que indica que los distintos observables reflejan diferentes propiedades conformacionales para hNedd4-WW3, hFbp11-WW1 y hFbp11-WW2 durante el desplegamiento. Este resultado, junto a la elevada pendiente de la función  $C_{p,N}(T)$



**Figura 2.13. Desnaturalización térmica del dominio hYap-WW1 (L30K) seguida por DSC y por CD bajo distintas condiciones de pH.** Los datos experimentales de ambas técnicas se han representado como círculos vacíos con diferente color según la condición de pH que representan: rojo para fosfato sódico 20mM pH 7.0, amarillo para MES 20mM pH 6.0, naranja para acetato sódico 20mM pH 5.0, verde para acetato sódico 20mM pH 4.0, cian para glicina 20mM pH 3.5 y azul para glicina 20mM pH 3.0. La desnaturalización térmica a pH 6.0 sólo se obtuvo mediante CD. Las líneas continuas corresponden a al ajuste global simultáneo de acuerdo con el modelo de dos estados, considerando común la función  $\Delta C_p(T)$  para ambos conjuntos de datos de CD (230 nm) y DSC, siendo directamente estimado del análisis de las trazas experimentales de DSC. Las líneas discontinuas en negro corresponden a las líneas base comunes para cada conjunto de curvas obtenidas a partir del ajuste. La línea recta continua en negro bajo las trazas de DSC representan la función teórica  $C_{p,N}(T)$  correspondiente a una estimación basada en el peso molecular de cada dominio (Freire 1995), mientras que la línea continua con cierta curvatura (casi coincide con la traza a pH 3.0) representa la función teórica  $C_{p,D}(T)$  calculadas mediante la adición de las contribuciones individuales de cada aminoácido (Privalov & Makhatadze 1990).

**Tabla 2.3. Parámetros termodinámicos de plegamiento de los dominos hYap-WW1(L30K), hNedd4-WW3, hFbp11-WW1 y hFbp11-WW2.** Los distintos parámetros se han obtenido mediante el ajuste de los experimentos de DSC o mediante el ajuste global de los experimentos de DSC junto a las desnaturalizaciones térmicas seguidas por CD<sup>a</sup>.

Método	pH	T <sub>m</sub> (°C)	ΔH <sub>m</sub> (kJ·mol <sup>-1</sup> )	ΔG(25°C) (kJ·mol <sup>-1</sup> )	T <sub>s</sub> (°C)	ΔG <sub>máx</sub> (T <sub>s</sub> ) (kJ·mol <sup>-1</sup> )	ΔC <sub>p</sub> (kJ·K <sup>-1</sup> ·mol <sup>-1</sup> )	Δh <sub>m</sub> (110 °C) <sup>d</sup> (kJ·g <sup>-1</sup> )	C <sub>p,N</sub> (20°C) (kJ·K <sup>-1</sup> ·mol <sup>-1</sup> )
<b>hNedd4-WW3</b>									
DSC distintos pH <sup>c</sup> (ajuste individual)	7	58.8	83	8.7	-9	13.3		42.5	
	6	50.4	79	5.1	5	6.7		--	
	5	48.3	63	3.7	11	4.3		39.9	11.3
	4	37.7	54	1.9	12	2.9	2.1	39.6	(ó 2.05 por g)
	3.5	30.3	52	0.9	9	2.5		36.0	
	3 <sup>b</sup>	12.4	29	--	--	--		--	
Doble perturbación CD	2 <sup>b</sup>	--	--	--	--	--		--	
	7 <sup>b</sup>	51.5	134	7.7	--	--	2.9	--	
<b>hYap-WW1 (L30K)</b>									
CD+DSC distintos pH <sup>c</sup> (ajuste global)	7	47.0	95	5.8	-24	11.4		43.9	
	6 <sup>b</sup>	47.9	91	--	--	--		45.0	
	5	46.0	93	5.4	-23	10.9	2.0	46.5	10.0
	4	35.6	75	2.5	-20	7.2		44.2	(ó 1.81 por g)
	3.5	21.6	54	-0.6	-17	3.7		42.5	
	3	5.0	31	-3.1	-16	1.3		--	
Doble perturbación CD	7 <sup>b</sup>	48.9	132	7.8	--	--	2.2	--	

**Tabla 2.3. Parámetros termodinámicos de plegamiento de los dominos hYap-WW1(L30K), hNedd4-WW3, hFbp11-WW1 y hFbp11-WW2 (continuación).**

<b>hFbp11-WW1</b>									
DSC distintos pH <sup>c</sup> (ajuste individual)	7	67.9	129	17.4	-34	32.5		41.0	
	5	69.2	111	15.7	-31	27.9		40.0	9.2
	4	66.3	107	14.0	-30	25.2	1.02	37.0	(ó 1.73 por g)
	3	55.3	100	9.4	-25	18.8		36.8	
	7 <sup>b</sup>	64.2	112	10.0	--	--	1.24	--	
Doble perturbación CD									
<b>hFbp11-WW2</b>									
CD+DSC distintos pH <sup>c</sup> (ajuste global)	7	52.2	121	8.8	-26	16.0		41.0	
	6 <sup>b</sup>	54.0	118	8.9	--	--		40.8	
	5	46.1	98	5.6	-19	10.6		38.1	9.5
	4	36.3	91	3.1	-18	8.6	1.89	40.6	(ó 1.70 por g)
	3.5 <sup>b</sup>	26.2	76	0.3	--	--		40.8	
	3	19.1	56	-1.2	-15	3.5		39.7	
	2 <sup>b</sup>	11.5	45	--	--	--		40.6	
Doble perturbación CD	7 <sup>b</sup>	51.8	128	8.2	--	--	2.10	--	

<sup>a</sup>Los errores se han estimado en 5% para la  $T_m$ , 10% para  $\Delta H_m$  y 20% para el resto de parámetros. <sup>b</sup>En estas condiciones de pH, los experimentos de desnaturalización térmica únicamente se obtuvieron mediante CD. <sup>c</sup>Valores para los coeficientes que definen las líneas bases compartidas entre los diferentes conjuntos de experimentos de acuerdo con la estrategia de ajuste según el modelo de dos estados, donde  $T_r$ , temperatura de referencia, es 25 °C y los valores subrayados fueron estimados a partir de la contribución individual de cada aminoácido (Privalov & Makhatdze 1990). <sup>d</sup>Este parámetro se ha calculado mediante la ecuación II.VI (Apéndice II) a partir de los valores de  $T_m$ ,  $\Delta H_m$  y  $\Delta C_p$  (pendiente de la correlación lineal entre  $T_m$  y  $\Delta H_m$ ) obtenidos a partir del análisis global de las desnaturalizaciones térmicas seguidas por CD.

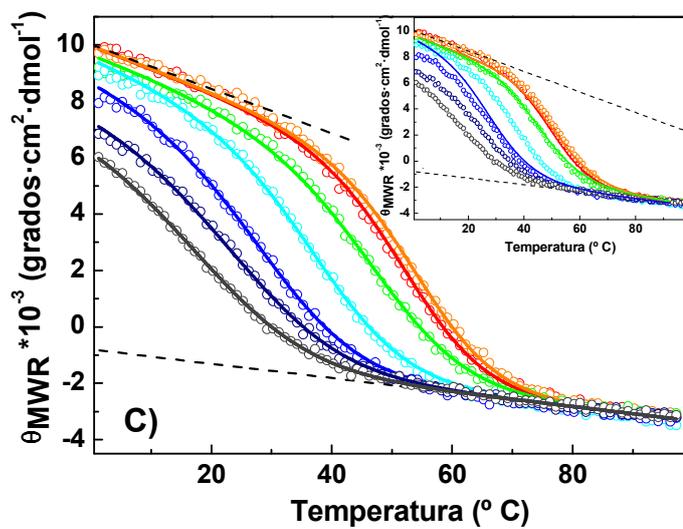
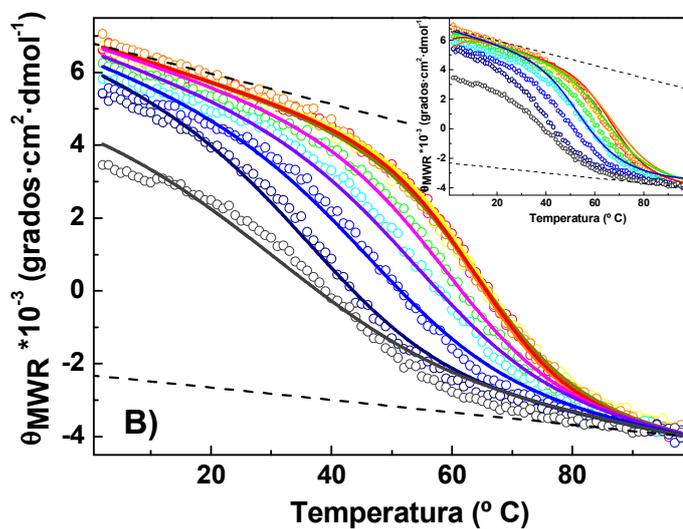
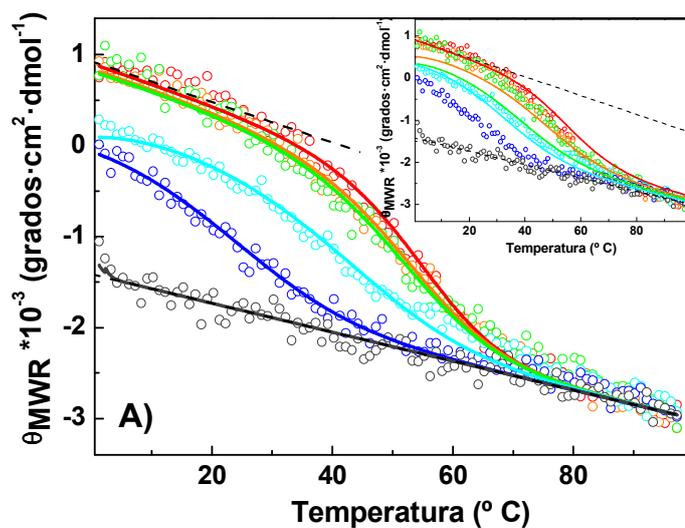
obtenida del ajuste global de DSC, son evidencias claras de que los equilibrios conformacionales de los dominios hNedd4-WW3, hFbp11-WW1 y hFbp11-WW2 presentan una mayor complejidad que el típico comportamiento de dos estados.

En la Tabla 2.3 se han recogido los parámetros termodinámicos obtenidos para la desnaturalización térmica de cada dominio WW en diferentes condiciones de pH mediante CD y DSC. Para el dominio hYap-WW1(L30K) se han considerado los parámetros obtenidos del análisis combinado de CD y DSC y para los otros dominios sólo los obtenidos mediante DSC.

En la Figura 2.14 se muestran los análisis globales del desplegamiento térmico seguido por CD (230 nm) para los dominios hNedd4-WW3, hFbp11-WW1 y hFbp11-WW2 en distintas condiciones de pH. En esta figura se puede observar cómo el modelo de dos estados es capaz de ajustar globalmente los conjuntos de curvas de CD de los dominios hNedd4-WW3 y hFbp11-WW2, pero no las curvas del dominio hFbp11-WW1.

Además, en los recuadros interiores de cada panel de la Figura 2.14, para ilustrar las diferencias de los parámetros termodinámicos resultantes de los distintos ajustes globales de las curvas de CD o DSC se han simulado las curvas de CD que resultarían de los parámetros obtenidos en los ajustes de DSC. Para los dos dominios hNedd4-WW3 y hFbp11-WW1 las diferencias de los parámetros obtenidos por ambas técnicas son notables. En ambos casos parecen sobrestimados los parámetros obtenidos mediante DSC respecto de los de CD. Estas diferencias son máximas para el dominio hNedd4-WW3 (Figura 2.14A), para el que ni siquiera ha sido posible realizar un ajuste global de las curvas de DSC. Este resultado constituye la principal evidencia de que el modelo de dos estados no es adecuado para el análisis del equilibrio conformacional de este dominio WW.

**Figura 2.14. Desnaturalización térmica de varios dominios WW seguida mediante CD a una longitud de onda de 230 nm y bajo distintas condiciones de pH: A) dominio hNedd4-WW3, B) dominio hFbp11-WW1 y C) dominio hFbp11-WW2. Los datos experimentales se han representado como círculos vacíos con diferente color según la condición de pH: rojo para fosfato sódico 20mM pH 7.0, amarillo para MES 20mM pH 6.0, naranja para acetato sódico 20mM pH 5.0, verde para acetato sódico 20mM pH 4.0, cian para glicina 20mM pH 3.5, azul para glicina 20mM pH 3.0, azul marino para glicina 20mM pH 2.5 y gris para glicina 20mM pH 2.0. Las líneas continuas corresponden al ajuste global de acuerdo con el modelo de dos estados, considerando común un valor de  $\Delta C_p(T)$  estimado como la pendiente de la correlación de  $\Delta H_m$  frente a  $T_m$ . Las líneas discontinuas en negro corresponden a las líneas bases comunes para cada conjunto de curvas e invariantes durante el ajuste. En cada recuadro, se ha representado una simulación de las curvas de CD (líneas continuas) sobre los datos experimentales (círculos). Las simulaciones se han realizado utilizando las mismas líneas base y los parámetros de los mejores ajustes al modelo de dos estados de las respectivas curvas de DSC de cada dominio WW**



Por otro lado, en las curvas simuladas con los datos de DSC de hFbp11-WW1 (Figura 2.14B), se observa una incipiente desnaturalización fría por debajo de 15 °C que es diferente a la disminución de señal a baja temperatura con el pH que se aprecia en los datos experimentales de CD, son diferentes efectos sobre el estado nativo. No obstante, para el dominio hFbp11-WW2 (Figura 2.14C) las curvas simuladas son mucho más parecidas a las experimentales y sólo hay diferencias importantes entre los parámetros a pH 3.0.

Los valores de convergencia para la entalpía específica media a 110 °C de cada dominio estimados como se ha indicado para el dominio hNedd4-WW4 son también inferiores a los 54 J·g<sup>-1</sup> propuestos para proteínas globulares; para los cuatro dominios, estos valores convergen en una entalpía específica media de unos 40-45 J·g<sup>-1</sup> independientemente de las condiciones experimentales (Tabla 2.3). Este resultado, como se ha comentado para hNedd4-WW4, refleja la baja cooperatividad de estos dominios en comparación con el resto de las proteínas globulares (Hilser et al. 1996). Además, los valores de entalpía, entropía o cambio de capacidad calorífica a 60 °C de estos dominios WW son bastante diferentes a los valores determinados para otras proteínas globulares (Robertson & Murphy 1997). Al igual que para el dominio hNedd4-WW4, los valores de convergencia para hYap-WW1(L30K) están próximos a los límites de los valores tradicionalmente aceptados para proteínas globulares. Esto es debido a que la intersección de las funciones  $C_{p,N}(T)$  y  $C_{p,D}(T)$  de hNedd4-WW4 y hYap-WW1(L30K) se producen a temperaturas mayores que en los otros dominios WW (Figuras 2.5 y 2.12B, respectivamente), de modo que la dependencia de los parámetros termodinámicos con la temperatura no se ven tan afectados por el valor de cambio de capacidad calorífica,  $\Delta C_p$ , en estos casos.

La temperatura de máxima estabilidad,  $T_S$ , de estos dominios también presenta valores muy bajos en comparación con los obtenidos en la caracterización de la estabilidad de otras proteínas globulares (Jackson & Fersht 1991; Felitsky & Record 2003). En las condiciones experimentales de mayor estabilidad (aparentemente a pH 7), estas  $T_S$  son -34 °C para el dominio hFbp11-WW1, alrededor de -25 °C para hYap-WW1 (L30K) y hFbp11-WW2, y -9 °C para hNedd4-WW3. A medida que las condiciones experimentales se vuelven más desestabilizantes (más ácidas) para estos dominios, los valores de energía libre de

máxima estabilidad,  $\Delta G_{\text{máx}}(T_S)$ , disminuyen, mientras que los valores de  $T_S$  aumentan (Tablas 2.1 y 2.3). Este resultado indica que los dominios WW mantienen una conformación nativa con una estabilidad marginal, que es suficiente para poder llevar a cabo el reconocimiento de ligandos, en condiciones poco favorables. Un ejemplo ilustrativo de esta característica de los dominios WW en la bibliografía es el del mutante W17F del dominio hYap-WW1 que, aunque se encuentra seriamente desestabilizado y mayoritariamente desplegado en solución, al unir a un ligando natural aumenta su estabilidad y se produce un doble proceso acoplado de unión y plegamiento del dominio WW (Koepf et al. 1999a).

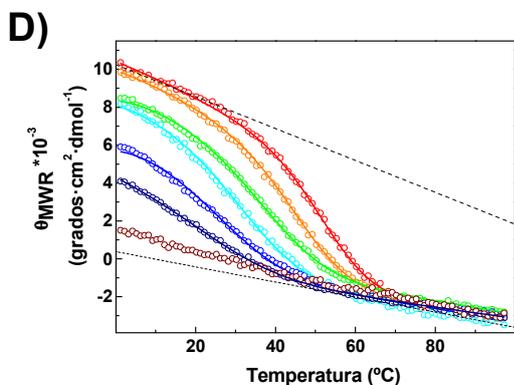
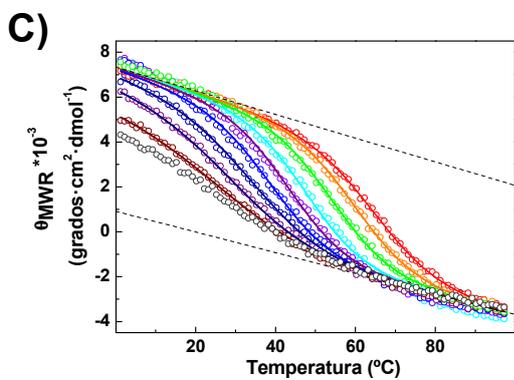
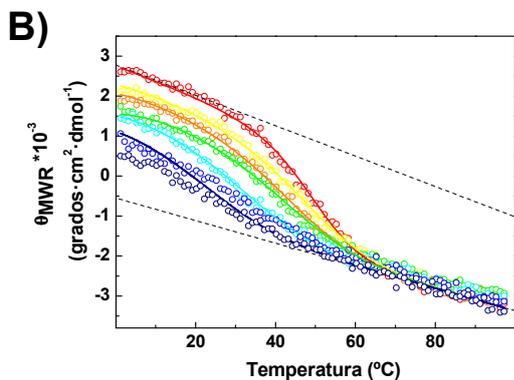
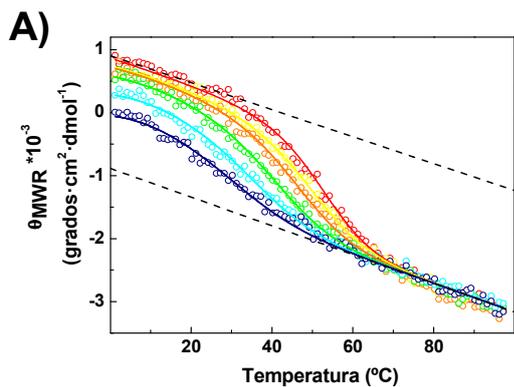
Como se ha mencionado para el dominio hNedd4-WW4, los valores de  $C_{p,N}(20\text{ °C})$  de estos dominios WW se encuentran muy próximos al límite superior de lo esperado para proteínas globulares (entre  $1.25$  y  $1.8\text{ J}\cdot\text{K}^{-1}\cdot\text{g}^{-1}$ ) (Makhatadze 1998) o por encima de éste en el caso de hNedd4-WW3 ( $2.05 \pm 0.2\text{ J}\cdot\text{K}^{-1}\cdot\text{g}^{-1}$ ), lo cual refleja una gran flexibilidad conformacional. De hecho, estos valores se correlacionan muy bien con la intensidad de la señal de CD normalizada a 230 nm que, como hemos comentado anteriormente, guarda relación directa con el empaquetamiento de restos aromáticos y, por tanto, inversa con la plasticidad de la proteína. Así, los dominios que tenían una mayor señal a 230 nm (Figura 2.9) son los que presentan valores más bajos de  $C_{p,N}(20\text{ °C})$  y el dominio hNedd4-WW3 que no presentaba este máximo es el que tiene el valor más alto.

Resumiendo, aunque el comportamiento descrito para el desplegamiento del dominio hNedd4-WW4 está al límite de lo que se considera un equilibrio de dos estados, los resultados para los otros dominios WW estudiados apuntan más claramente hacia la existencia de un equilibrio conformacional más complejo para los dominios WW en general.

### 2.3.3. Experimentos de doble perturbación

Con el fin de profundizar en el estudio de la naturaleza del equilibrio de plegamiento-desplegamiento de estos dominios, hemos decidido investigar si el acoplamiento de la desnaturalización térmica y química en el equilibrio revela la existencia de una barrera de energía entre los estados nativo y desplegado.

En la Figura 2.15 se muestra el ajuste múltiple de las curvas de doble perturbación al modelo de dos estados para los dominios hNedd4-WW3,



**Figura 2.15. Dependencia de la desnaturalización térmica seguida por CD (230 nm) en presencia de distintas concentraciones de urea para los dominios WW: A) hNedd4-WW3, B) hYap1-WW1(L30K), C) hFbp11-ww1 y D) hFbp11-WW2.** En círculos vacíos se han representado los datos experimentales en fosfato sódico 20mM pH 7.0, en presencia de urea, atendiendo al siguiente código de colores: 0 M urea en rojo, 0.5 M urea en amarillo, 1.0 M urea en naranja, 2.0 M urea en verde, 3.0 M urea en turquesa, 4.0 M urea en púrpura, 5.0 M urea en azul, 6.0 M urea en azul marino, 7.0 M urea en azul marrón oscuro y 8.0 M urea en gris oscuro. Las líneas continuas, que acompañan los datos experimentales, corresponden al mejor ajuste global según el modelo de dos estados (ecuación II.IX) compartiendo las líneas base del estado nativo y del estado desnaturalizado para cada conjunto de datos. Las líneas rectas discontinuas representan dichas líneas base obtenidas a partir de la estrategia de ajuste global para los estados nativo y desplegado en cada conjunto de experimentos de desnaturalización térmica. A la derecha, para cada dominio se han representado los valores de entalpía de desplegamiento a 25 °C obtenidos a partir del ajuste individual frente a la concentración de urea. La línea continua corresponde al ajuste lineal de los datos. Las barras representan la propagación de errores.

hYap-WW1(L30K), hFbp11-WW1 y hFbp11-WW2. El ajuste al modelo de dos estados reproduce adecuadamente las curvas experimentales y resulta en valores de  $T_m$  y  $\Delta H_m$  que presentan una muy buena correlación lineal, con la excepción del dominio hFbp11-WW1 caracterizado por un aumento de la señal nativa de CD a bajas temperaturas a concentraciones de urea entre 1 y 5 M, que el modelo de dos estados no puede reproducir. Consecuentemente, la correlación entre los valores de  $T_m$  y  $\Delta H_m$  obtenidos para este dominio no es lineal. Estas características podrían ser indicativas de la variación de las propiedades conformacionales con la urea que cambiaría la sensibilidad del estado nativo frente a la desnaturalización térmica, de acuerdo con un equilibrio de tipo “downhill”.

En los paneles a la derecha de la Figura 2.15, se han representado los valores de  $\Delta H_D(25\text{ °C})$  obtenidos a partir del ajuste individual de las curvas para cada dominio WW (considerando los valores de  $\Delta C_p$  para cada proteína recogidos en la Tabla 2.3) frente a la concentración de urea. Los valores de correlación lineal,  $r$ , para los correspondientes ajustes son 0.94, 0.88 y 0.78 para hNedd4-WW3, hYap-WW1 (L30K) y hFbp11-WW2, respectivamente, mientras que para el dominio hFbp11-WW1,  $r$  es de 0.17. Como se puede apreciar en la figura, el error asociado a la determinación de los valores de  $\Delta H_D(25\text{ °C})$  es elevado. Para los dominios hNedd4-WW3, hYap-WW1 (L30K) y hFbp11-WW2 se observa una marcada tendencia a un acoplamiento lineal de los valores de  $\Delta H_D(25\text{ °C})$  con la concentración de urea, aunque con desviaciones significativas para los dominios hNedd4-WW3 y hFbp11-WW2. No obstante, es importante señalar que, en todos los casos, se observa una pendiente positiva, opuesta a la descrita para el dominio hNedd4-WW4 y para otras proteínas globulares (Cobos et al. 2002; Felitsky & Record 2003), y carente de sentido físico, ya que implicaría valores mayores de

$\Delta H_D$  (25 °C) para las condiciones en las que el estado desplegado está más poblado, es decir, a mayores concentraciones de urea.

Además, si se comparan los valores de  $T_m$  a pH 7 obtenidos en los experimentos de doble perturbación con los de los experimentos de DSC (Tabla 2.3), para los dominios hNedd4-WW3 y hFbp11-WW1 se observan grandes diferencias. Estas diferencias de  $T_m$  determinadas mediante el ajuste al modelo de dos estados a partir de distintos observables son de un 6-14 %, lo cual es otra evidencia de que el equilibrio conformacional de estos dos dominios presenta una mayor complejidad. Los resultados de los experimentos de doble perturbación de los dominios hYap-WW1(L30K) y hFbp11-WW2 aunque presentan parámetros termodinámicos más parecidos en los distintos análisis tampoco están de acuerdo con los criterios establecidos para una proteína dos estados y más bien apuntan a un plegamiento sin barreras o con barreras de energía muy bajas separando los diferentes macro-estados.

En resumen, del conjunto total de dominios WW caracterizados en este capítulo, es el dominio hNedd4-WW4 el que exhibe un comportamiento más acorde con el modelo de dos estados a pesar de tener valores anómalos en algunos de los parámetros termodinámicos discutidos anteriormente [ $\Delta h$  (110 °C),  $T_s$ ,  $C_{p,N}$  (25 °C), etc].

### **2.4. Perspectivas y futuro en el estudio de plegamiento de dominios WW**

Los resultados presentados para los distintos dominios WW, están de acuerdo con la hipótesis de que los equilibrios de dos estados y “downhill” no constituyen dos clases separadas claramente sino que son los límites extremos de una serie de posibles comportamientos determinados por la altura de las barreras, la cual puede ser modulada por las condiciones experimentales. En este contexto, para completar esta caracterización, la primera caracterización en profundidad del desplegamiento de un conjunto de dominios naturales cuya estructura es básicamente una lámina  $\beta$ , desde nuestro grupo de investigación hemos planteado una doble colaboración con los grupos de los Drs. Víctor Muñoz (Centro de Investigaciones Biológicas, CSIC, Madrid) y José Manuel Sánchez-Ruiz (Departamento de Química-Física, Universidad de Granada, Granada). Actualmente, en el seno de esta colaboración se está llevando a cabo la

investigación de las cinéticas de desplegamiento de los dominios WW empleados en este trabajo mediante técnicas de plegamiento ultra-rápido de proteínas (Eaton et al. 2000), así como la caracterización de las barreras termodinámicas de desplegamiento implicadas mediante el análisis de los termogramas de DSC con el modelo de barrera variable (Munoz & Sanchez-Ruiz 2004).



## CAPÍTULO 3



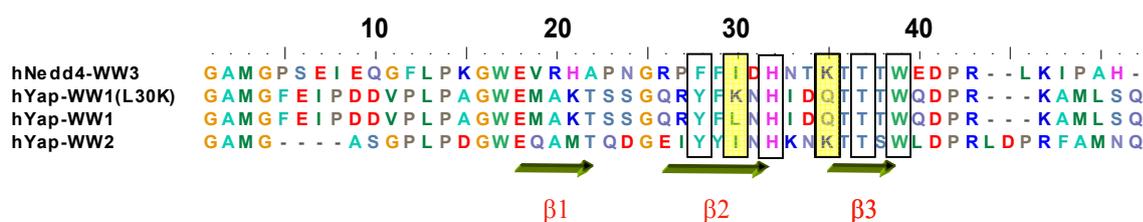
## **CAPÍTULO 3. Estudio de unión de dominios L víricos al tercer dominio WW de la ligasa de ubiquitina humana Nedd4**

### **3.1. Introducción**

Numerosos estudios sobre el ciclo de vida de los virus han puesto de manifiesto que las interacciones específicas entre los dominios L víricos del tipo PPxY y los dominios WW de las ligasas de ubiquitina de la familia Nedd4 son fundamentales para las fases finales de encapsulamiento y salida del virus de la célula huésped (Freed 2002; Park et al. 2005; Bieniasz 2006). Este es el caso de la interacción del tercer dominio WW de la ligasa de ubiquitina humana Nedd4 (hNedd4-WW3) con varias secuencias de algunas proteínas víricas, como son la proteína Gag del virus de la leucemia tipo I (HTLV1) (Blot et al. 2004) o la proteína VP40 (del inglés, "Virus Particle 40") del virus del Ébola (Harty et al. 2000; Timmins et al. 2003). Los dominios WW de las ligasas se caracterizan por presentar una elevada promiscuidad, ya que unen a una gran variedad de proteínas celulares (Hesselberth et al. 2006; Belgareh-Touze et al. 2008). A pesar de esto, los virus han adquirido un elevado grado de especialización a la hora de reconocer estas ligasas. En este contexto, el estudio de las interacciones entre los dominios L víricos y los dominios WW de ligasas tipo Nedd4 a nivel molecular presentan gran interés para el diseño racional de inhibidores específicos de estas interacciones que podrían constituir antivirales de amplio espectro. A día de hoy, todavía no se dispone de ningún estudio termodinámico que aporte información acerca de los determinantes moleculares de este tipo de interacciones.

En este capítulo presentamos un estudio termodinámico-estructural de las interacciones del tercer dominio WW de la ligasa de ubiquitina humana Nedd4 con un conjunto de ligandos peptídicos correspondientes a las secuencias de unión de dominios L de varios virus, con el fin de investigar los determinantes moleculares de la afinidad de unión en este sistema. Además, hemos abordado el estudio de la especificidad de unión de los dominios L frente a distintos dominios WW del grupo I (Figura 3.1), como son los dominios WW de la proteína humana hYap, hYap-WW1 y hYap-WW2, así como el mutante L30K del primer dominio WW, hYap-

WW1(L30K). Estos dominios WW se han escogido porque en las primeras caracterizaciones de dominios L del tipo PPxY se comprobó que interactuaban con el primer dominio WW de hYap (Garnier et al. 1996). En el caso de los dominios hYap-WW1 y hYap-WW1(L30K), además disponemos de información estructural tanto de distintos complejos con ligandos ricos en prolinas (Pires et al. 2001) como del dominio hYap-WW1 libre (Macias et al. 1996). Nuestro objetivo es evaluar si, tal como sugiere su independencia funcional de los dominios L (Parent et al. 1995; Freed 2002), la especificidad observada para los dominios L víricos radica en características intrínsecas de sus secuencias o, por el contrario, dicha especificidad responde a factores contextuales, como son la localización celular o el efecto sinérgico debido a interacciones establecidas con diferentes partes de la proteína o con distintas proteínas (Ladbury & Arold 2000).



**Figura 3.1. Diferencias entre los dominios WW hNedd4-WW3, hYap-WW1(L30K), hYap-WW1 y hYap-WW2.** Alineamiento múltiple de secuencias utilizando el algoritmo ClustalX. La estructura secundaria se ha representado con flechas rojas bajo los restos implicados. Los aminoácidos que conforman el sitio de unión se han recuadrado y aquellos que no están conservados aparecen sobre fondo amarillo.

### 3.2. Estudio de unión de distintos dominios L víricos al dominio hNedd4-WW3

Para alcanzar estos objetivos, hemos realizado la primera caracterización termodinámica exhaustiva mediante calorimetría isotérmica de titulación (ITC) de la unión de un conjunto de ligandos peptídicos a dominios WW: concretamente la unión de cuatro dominios L víricos a su diana celular, el tercer dominio WW de la ligasa de ubiquitina humana Nedd4, hNedd4-WW3.

#### 3.2.1. Descripción de los ligandos

Para investigar los determinantes moleculares de la interacción de los dominios L víricos con hNedd4-WW3, hemos seleccionado seis ligandos diferentes

derivados de cuatro proteínas víricas que interaccionan con ligasas de ubiquitina de la familia Nedd4. Las secuencias y los nombres de estos ligandos se recogen en la Tabla 3.1.

**Tabla 3.1. Selección de los ligandos utilizados para la caracterización termodinámica de la unión a un conjunto de dominios WW.**

Proteína vírica	Ligando	Secuencia*
<b>HTLV1</b> (Proteína Gag del virus de la leucemia humana de células T tipo I)	<b>HTLV1-a</b>	<u>4</u> · <b>PPPY</b> VEPTAP <sub>14</sub> ·
	<b>HTLV1-b</b>	· <sub>1</sub> SDPQI <b>PPPY</b> VEP <sub>14</sub> ·
<b>Ébola</b> (Proteína de asociación a membrana VP40 del virus del Ébola)	<b>Ébola-a</b>	· <sub>1</sub> ·ILPTA <b>PPPEY</b> ME <sub>9</sub> ·
	<b>Ébola-b</b>	· <sub>1</sub> ·ILPTA <b>PPPEY</b> <sub>7</sub> ·
<b>Marburgo</b> (Proteína de asociación a membrana VP40 del virus del Marburgo)	<b>Marburgo</b>	· <sub>1</sub> ·MQYLN <b>PPPY</b> ADH <sub>10</sub> ·
<b>Rabia</b> (Proteína de matriz M2 del virus de la Rabia)	<b>Rabia</b>	· <sub>1</sub> ·DLWL <b>PPPEY</b> VPL <sub>10</sub> ·
Proteína natural		
<b>p53bp2</b> (Proteína supresora de tumores tipo 2 de unión a p53)	<b>p53bp2</b>	· <sub>2</sub> ·EYPPYP <b>PPPY</b> PSG <sub>10</sub> ·
Motivo canónico de unión dominios WW de tipo I		
<b>PPPY</b> (Motivo de unión a dominios WW de la Clase I)	<b>PPPY</b>	<u>4</u> · <b>PPPY</b> <sub>7</sub> ·

\* Para identificar más fácilmente la posición del motivo PPxY en la secuencia de los ligandos éste se muestra en azul y la otra secuencia característica de dominios L víricos, PTAP, presente en algunos de estos ligandos, además se ha subrayado. La numeración de las secuencias de todos los ligandos se ha unificado como se indica en esta tabla y se diferenciado de las secuencias de los dominios mediante el símbolo prima. Se ha establecido como criterio general la numeración del ligando GTPPPPYTVG en el trabajo Pires et al. 2001.

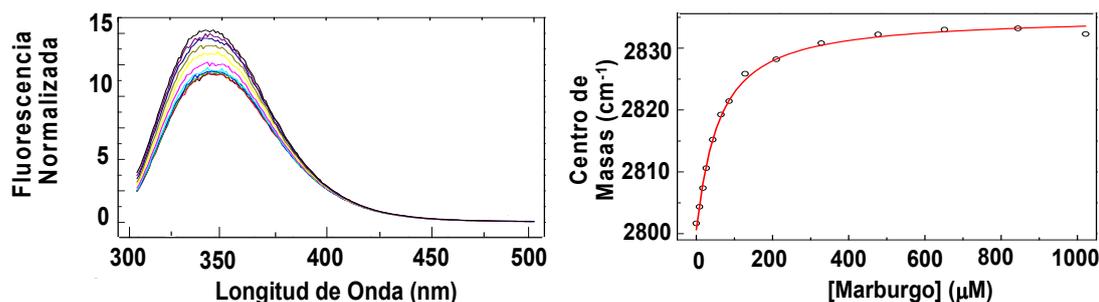
Las proteínas VP40 (virus del Ébola) y Gag (virus HTLV1) presentan dos dominios L: PTAP y PPxY. En ambos casos, se han seleccionado dos péptidos de diferente longitud con el motivo PPxY dispuesto en distintas posiciones. En el caso de la proteína Gag, se sabe que el ligando **HTLV1-a** (**PPPY**VEPTAP) interacciona con hNedd4-WW3 (Blot et al. 2004) mientras que el ligando **HTLV1-b** (SDPQI**PPPY**VEP) ha sido seleccionado de forma que el motivo PPPY quede centrado en su secuencia. En el caso de la proteína VP40 del virus del Ébola, se conoce la interacción del dominio hNedd4-WW3 con un fragmento de esta proteína (Harty et al. 2000; Timmins et al. 2003). A partir de este fragmento se han seleccionado dos ligandos de diferente longitud **Ébola-a** (ILPTA**PPPEY**ME) y **Ébola-b** (ILPTA**PPPEY**). La secuencia **Marburgo** (MQYLN**PPPY**ADH) deriva de la proteína VP40 del filovirus Marburgo, para la que se ha descrito su interacción con

varios de los dominios WW de Nedd4 (Urata & Yasuda 2010). Para la fosfoproteína de matriz M2 del virus de la Rabia se ha escogido el ligando **Rabia** (DLWLPPPEYVPL), cuya interacción con proteínas homólogas a Nedd4 se ha descrito en ratón y levadura (Harty et al. 1999; Harty et al. 2000). Además, hemos seleccionado otros dos ligandos que interactúan con hYap-WW1. Por un lado, el ligando **p53bp2** (EYPPYPPPYPSG), derivado de la proteína “p53 binding protein-2”, que es una diana natural de hYap-WW1 descubierta mediante la técnica de doble híbrido de levadura (“yeast two-hybrid screening”) (Españel & Sudol 2001). Se ha descrito que esta secuencia interactúa con hYap-WW1 con una constante de disociación de 5.9  $\mu\text{M}$  (Koepp et al. 1999a). Por otro lado, para evaluar la influencia, tanto en la afinidad como en la especificidad de unión de las secuencias que flanquean el motivo consenso PPXY, se ha seleccionado el ligando **PPPY** que es la secuencia más corta de unión a hYap-WW1 (**PPPY**). Además, se dispone de la estructura del complejo G**PPPY**/hYap-WW1 resuelta mediante RMN (PDB: 1K9Q) (Pires et al. 2001).

#### 3.2.2. Caracterización termodinámica de la unión

Con el fin de confirmar la unión y obtener una estimación inicial de las constantes de disociación que nos permita diseñar correctamente los experimentos calorimétricos, hemos llevado a cabo titulaciones preliminares mediante espectroscopía de fluorescencia de hNedd4-WW3 con todos los ligandos seleccionados, excepto con el ligando **Rabia**. Este ligando contiene un resto triptófano en su secuencia que no permite seguir fácilmente la amortiguación de la fluorescencia propia del dominio WW durante la titulación. En la Figura 3.2 se muestra, como ejemplo, la titulación del dominio hNedd4-WW3 con el ligando **Marburgo** a 25 °C en tampón fosfato sódico 20 mM pH 7.0. A este valor de pH el dominio hNedd4-WW3 presenta la mayor estabilidad conformacional (Capítulo 2). La interacción entre hNedd4-WW3 y el ligando **p53bp2** presenta la mayor afinidad y su constante de disociación es del mismo orden que la descrita para su diana natural, el dominio hYap-WW1 (Koepp et al. 1999a). Los ligandos **HTLV1-a**, **HTLV1-b**, **Ébola-a** y **Marburgo** presentan constantes de disociación menores, pero que aún están dentro del intervalo adecuado para su caracterización

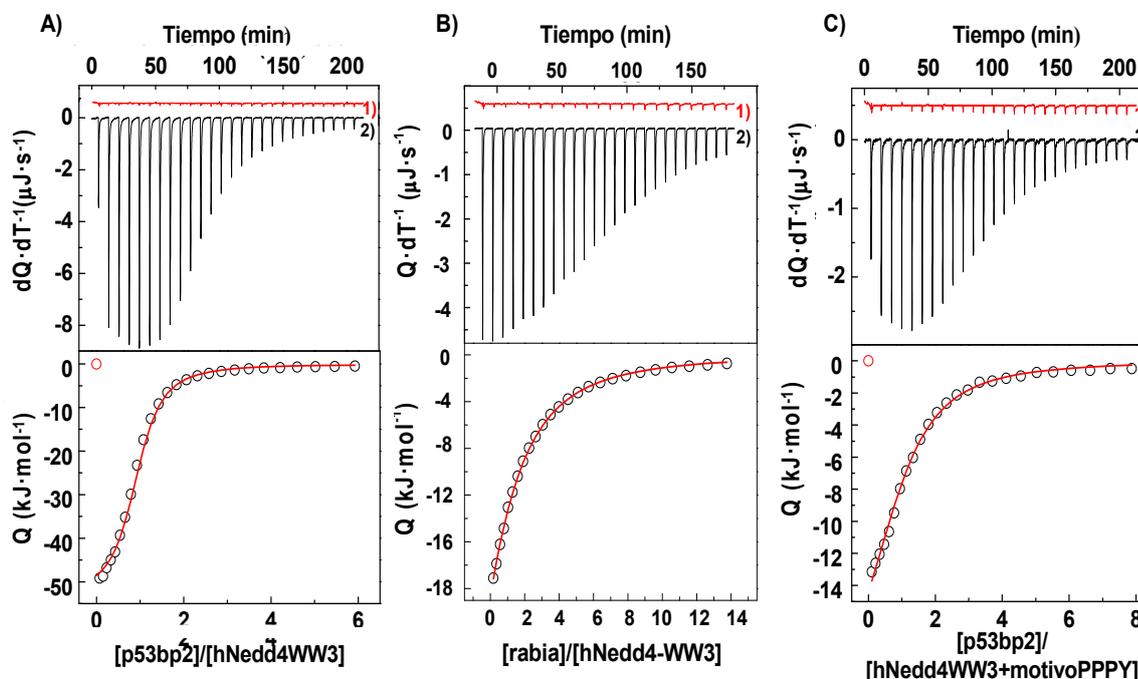
mediante experimentos directos de ITC. Las constantes de disociación de los ligandos **Ébola-b** y **PPPY** no se han podido determinar mediante fluorescencia debido a su baja afinidad, por lo que ha sido necesario recurrir a experimentos calorimétricos de desplazamiento utilizando el ligando **p53bp2**, para su caracterización termodinámica.



**Figura 3.2. Titulación del dominio hNedd4-WW3 con el ligando Marburgo seguida mediante espectroscopía de fluorescencia.** El experimento se ha realizado a 25 °C en tampón fosfato sódico 20 mM pH 7.0. En el panel izquierdo, se recogen los espectros de fluorescencia, normalizados por la concentración de proteína, tomados tras cada adición de ligando sobre el dominio hNedd4-WW3. En el panel derecho, se ilustran los valores del centro de masas (círculos abiertos) para cada uno de los espectros del panel superior frente a la concentración de ligando añadido. La línea continua corresponde al ajuste según el modelo de unión de un ligando a una proteína mediante fluorescencia (ecuación 6.5.3, Apartado 6.5.4 de Materiales y Métodos).

Todas las titulaciones calorimétricas se han realizado en las mismas condiciones utilizadas para los experimentos de fluorescencia. En los paneles superiores de la Figura 3.3 se muestran, a modo de ejemplo, las titulaciones del dominio hNedd4-WW3 con los ligandos **p53bp2** (panel A), **Rabia** (panel B) y **PPPY** (panel C). En la parte inferior de cada experimento se muestran las isotermas de unión correspondientes, que se han analizado según distintos modelos de unión: el modelo de “n” sitios de unión idénticos e independientes (ecuación X.III Apéndice III) para las titulaciones con los ligandos **p53bp2** y **Rabia** (Paneles A y B) y el modelo de unión en un experimento por desplazamiento (ecuación XIV.IV Apéndice IV) para el ligando **PPPY** (Panel C). Los valores de los parámetros termodinámicos obtenidos para las interacciones entre el dominio hNedd4-WW3 y los diferentes ligandos se recogen en la Tabla 3.2, excepto para la interacción con el ligando **Ébola-b**. En este caso, no se han podido analizar adecuadamente

ninguno de los experimentos realizados, probablemente debido a que la afinidad es demasiado baja.



**Figura 3.3. Experimentos de titulación del dominio hNedd4-WW3 mediante ITC a 25 °C en fosfato sódico 20 mM pH 7.0.** A) Experimento de titulación directa mediante la adición volúmenes constantes del ligando **p53bp2**. B) Experimento de titulación directa mediante la adición siguiendo un perfil de volúmenes del ligando **Rabia**. C) Experimento de titulación por desplazamiento del ligando **PPPY**. En cada caso, en la parte superior, se ha representado, en rojo 1), el termograma de dilución del ligando formado por los calores por unidad de tiempo liberados tras la inyección sobre tampón (desplazados en el eje para mayor claridad) y, en negro 2), el termograma de titulación del ligando en las mismas condiciones sobre el una disolución de proteína (en el experimento de titulación directa) o de una mezcla proteína-ligando de baja afinidad (en el experimento de competición). En la parte inferior de cada gráfica se muestran las isothermas de unión correspondientes obtenidas tras normalizar los calores de titulación y corregir por los calores de dilución. Los datos experimentales se representan como círculos, mientras que las líneas continuas corresponden a los mejores ajustes según el modelo de “n” sitios de unión idénticos e independientes (ecuación III.X, Apéndice III) para las titulaciones con los ligandos **p53bp2** y **Rabia** y según el modelo de unión en un experimento por desplazamiento (ecuación IV. XIV, Apéndice IV) para el ligando **PPPY**.

Las constantes de unión obtenidas para todos los ligandos seleccionados son de magnitud similar a las descritas para otras interacciones con dominios WW (Macias et al. 1996; Koepf et al. 1999a; Kanelis et al. 2001; Pires et al. 2001; Shi et al. 2002; Pires et al. 2005; Meiyappan et al. 2007; Morales et al. 2007).

El ligando **p53bp2** es el que presenta mayor afinidad de unión por hNedd4-WW3 ( $K_d=5 \mu\text{M}$ ) mientras que el ligando **PPPY** es el que presenta la afinidad

**Tabla 3.2. Parámetros termodinámicos de unión.** Se han obtenido a partir de la titulación en tampón fosfato sódico 20 mM a pH 7.0 de los dominios hNedd4-WW3, hYap-WW1(L30K), hYap-WW1 y hYap-WW2 con varios dominios L víricos y los ligandos **p53bp2** y **PPPY**

Ligando/ secuencia	Dominio	T <sup>a</sup> (°C)	<sup>&amp;</sup> K <sub>d</sub> (μM)	ΔG <sub>ap</sub> (kJ mol <sup>-1</sup> )	<sup>&amp;</sup> ΔH <sub>ap</sub> (kJ mol <sup>-1</sup> )	-TΔS <sub>ap</sub> (kJ mol <sup>-1</sup> )
EYPPY <b>PPPY</b> PSG	hNedd4-WW3	20	4.1	-30.8	-45.6	14.8
		25	5.3	-30.1	-53.2	23.1
		30	9.1	-29.2	-61.1	31.9
	hYap-WW1(L30K)	20	12.5	-27.5	-47.3	19.8
		25	20.4	-26.7	-53.2	26.5
		30	30.3	-26.2	-56.1	29.9
	hYap-WW1	20	1.1	-34.1	-54.8	20.7
		25	1.8	-32.8	-61.1	28.3
		30	3.0	-31.6	-66.4	34.8
	hYap-WW2	20	--	--	--	--
		25	12.0	-28.1	-56.5	28.4
		30	--	--	--	--
HTLV1-a <b>PPPY</b> VEPTAP	hNedd4-WW3	25	177.9	-21.4	-74.0	52.6
	hYap-WW1(L30K)	25	769.2	*-17.8	*-34.7	*16.9
	hYap-WW1	25	--	--	--	--
	hYap-WW2	25	--	--	--	--
HTLV1-b SDPQI <b>PPPY</b> VEP	hNedd4-WW3	25	61.0	-24.1	-68.2	44.1
	hYap-WW1(L30K)	25	270.3	*-20.4	*-35.6	*15.2
	hYap-WW1	25	#308	--	--	--
	hYap-WW2	25	#255	--	--	--
Ébola-a ILPTA <b>PPEY</b> ME	hNedd4-WW3	25	146.9	-21.9	-50.7	28.8
	hYap-WW1(L30K)	25	547.0	*-18.6	*-49.4	*30.8
	hYap-WW1	25	#754	--	--	--
	hYap-WW2	25	#555	--	--	--
Marburgo MQYLN <b>PPPY</b> ADH	hNedd4-WW3	25	51.2	-24.5	-73.2	48.7
	hYap-WW1(L30K)	25	#106	--	--	--
	hYap-WW1	25	#67	--	--	--
	hYap-WW2	25	#172	--	--	--
Rabia DLWL <b>PPPEY</b> VPL	hNedd4-WW3	25	50.5	-24.0	-65.8	41.8
	hYap-WW1(L30K)	25	218.3	-21.4	-63.7	42.3
	hYap-WW1	25	180.8	-20.9	-63.7	42.8
	hYap-WW2	25	--	--	--	--
<b>PPPY</b> <b>PPPY</b>	hNedd4-WW3	25	210.0	*-21.0	*-38.0	*17.0
	hYap-WW1(L30K)	25	622.6	*-18.3	*-40.1	*21.8
	hYap-WW1	25	320.7	*-20.0	*-39.3	*19.3
	hYap-WW2	25	--	--	--	--

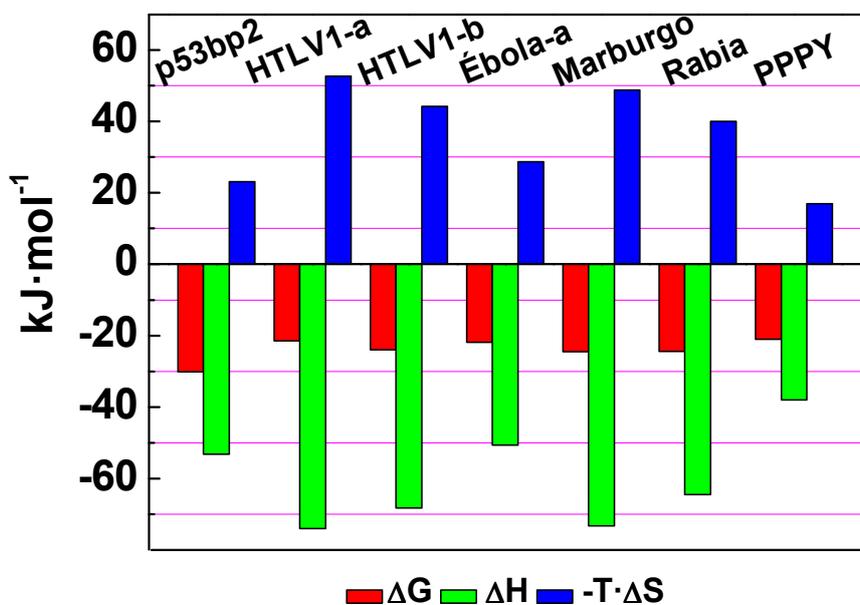
<sup>&</sup>Los errores en la determinación de los parámetros termodinámicos se estiman en torno al 5% para la entalpía de unión y al 10% para la constante de disociación. \*Parámetros termodinámicos determinados mediante experimentos de ITC por desplazamiento con el ligando **p53bp2**. #Constantes de disociación determinadas mediante experimentos de titulación seguidos por fluorescencia. Para el ligando **Ébola-b** no se ha podido determinar ninguno de los parámetros termodinámicos de unión

menor ( $K_d=210 \mu\text{M}$ ). La afinidad de unión de **p53bp2** a hNedd4-WW3 es muy similar a la descrita para su diana celular (hYap-WW1) (Koepp et al. 1999a), lo que indica la baja especificidad intrínseca de esta interacción. Estos resultados también confirman la unión de hNedd4-WW3 a la secuencia más corta que reconocen los dominios WW de tipo I, el ligando **PPPY**. Es interesante señalar que la afinidad de esta interacción es mayor que la determinada para un péptido similar (n-Octilo-GPPPY) con hYap-WW1 ( $K_d=700 \mu\text{M}$ ) en un estudio mutacional y de truncado de secuencias de unión (Pires et al. 2001). Los ligandos de origen vírico presentan una afinidad de unión a hNedd4-WW3 intermedia, con constantes que oscilan entre 50 y 178  $\mu\text{M}$  (-21 a -25  $\text{kJ}\cdot\text{mol}^{-1}$  en términos de energía libre de unión). Un estudio de la interacción mediante plasmón superficial entre fragmentos de la proteína VP40 del virus del Ébola y de la proteína Nedd4 determinó la formación de complejos supramoleculares con unas constantes de disociación de orden  $\mu\text{M}$ , no muy diferentes a la que se ha determinado en nuestro estudio con el ligando **Ébola-a** (Timmins et al. 2003). En este sentido, las afinidades de unión de los dominios L víricos determinadas para el dominio hNedd4-WW3, confirman que las secuencias de los dominios L extraídas del contexto de la proteína completa mantienen la capacidad de reconocer sus dianas naturales. Nuestros resultados ponen de manifiesto que los determinantes de la afinidad de unión de los dominios L están contenidos en su secuencia, en concordancia con los estudios mutacionales o de producción de quimeras con distintos dominios L que han propuesto que estos dominios funcionan como unidades relativamente autónomas (Parent et al. 1995; Freed 2002).

Las interacciones de hNedd4-WW3 con los dos dominios L que tienen el motivo PPxY en sus extremos, **Ébola-b** (ILPTA**PP**PEY) y **HTLV1-a** (**PP**PYVEPTAP), son más débiles que las interacciones de hNedd4-WW3 con los ligandos homólogos, **Ébola-a** y **HTLV1-b**, o con los otros ligandos, **Marburgo** y **Rabia**, todos con el motivo PPxY centrado. Así, estos resultados ponen de manifiesto que algunos de los determinantes moleculares de la unión de los dominios L al dominio hNedd4-WW3 se hallan en algunos de los restos que flanquean el motivo PPxY. El ligando **Ébola-b** presenta una afinidad por hNedd4-WW3 tan baja que ni siquiera se ha podido determinar su constante de disociación mediante titulaciones por

desplazamiento. El ligando **HTLV1-a** presenta una afinidad por hNedd4-WW3 del mismo orden que la obtenida en los experimentos de titulación por desplazamiento para el ligando **PPPY** (Tabla 3.2).

El análisis termodinámico realizado no sólo nos permite obtener información acerca de la afinidad de unión, sino que, además, hace posible investigar la naturaleza y magnitud de las fuerzas que contribuyen a la energía libre de unión. De esta manera, puede ayudarnos a conocer los determinantes moleculares de las interacciones entre el dominio hNedd4-WW3 y los dominios L. Para facilitar la comparación de la energética de las diferentes interacciones del dominio hNedd4-WW3, los parámetros termodinámicos se han representado como diagramas de barras en la Figura 3.4.



**Figura 3.4. Parámetros de unión del dominio hNedd4-WW3 al ligando p53bp2, a varios dominios L víricos y al ligando PPPY.** Representación en forma de barras de la energía libre de Gibbs (rojo) y de las contribuciones entálpica (verde) y entrópica (azul) de la energética de unión de los distintos ligandos con el dominio hNedd4-WW3. Todos los parámetros se han determinado mediante ITC a 25 °C y en fosfato sódico 20 mM, pH 7.0. Los errores en la determinación de los parámetros termodinámicos se estiman en torno al 5%.

Como se puede observar, todas las interacciones del dominio hNedd4-WW3 con ligandos ricos en prolina comparten un mismo perfil termodinámico, ya que están gobernadas por un valor de entalpía favorable (negativo), de unos -70 kJ·mol<sup>-1</sup>, al que se opone parcialmente una contribución entrópica desfavorable

(positiva), de unos  $45 \text{ kJ}\cdot\text{mol}^{-1}$ . Este patrón termodinámico es similar al descrito para la unión de dominios SH3 y ligandos poliprolina (Renzoni et al. 1996; Arold et al. 1998; Wang et al. 2001a; Ferreon & Hilser 2004; Palencia et al. 2004). Normalmente, una contribución entálpica favorable está relacionada con interacciones polares específicas, como la formación de puentes de hidrógeno, y con la desolvatación de restos polares (Hendsch et al. 1996). Las interacciones con los dominios WW se basan fundamentalmente en la inserción de dos restos prolina y una cadena lateral aromática (la tirosina del motivo PPxY) en dos bolsillos hidrofóbicos, por lo que sería razonable esperar que la interacción estuviera dirigida por la ganancia en entropía de solvatación debida a la unión de los restos hidrofóbicos y que la contribución entálpica no fuera tan determinante, pudiendo ser incluso desfavorable. En dominios SH3, esta discrepancia entre el perfil termodinámico esperado para una interacción fundamentalmente hidrofóbica y el patrón termodinámico obtenido experimentalmente se ha interpretado como una indicación de un mayor nivel de complejidad en estas interacciones al que contribuyen factores adicionales, como pueden ser la modulación del equilibrio conformacional del propio dominio SH3 o del ligando como consecuencia de la unión (Wang et al. 2001a; Ferreon & Hilser 2004) o bien, la contribución de moléculas de agua interfaciales que proporcionan un segundo nivel de interacción (Palencia et al. 2004; Palencia et al. 2010). A la vista de los resultados obtenidos, estos factores podrían ser también de relevancia para el reconocimiento de secuencias ricas en prolina por dominios WW. De hecho, es interesante hacer notar que la entalpía de unión del motivo **PPPY** aislado, fundamentalmente hidrofóbico, constituye entre un 55% y un 75% de la entalpía de unión de las secuencias completas, confirmando la importancia del motivo central en estas interacciones y apuntando hacia la existencia de otras contribuciones adicionales que expliquen dicho patrón termodinámico. En este sentido, la elevada flexibilidad conformacional de estos dominios, discutida en el Capítulo 2, y puesta de manifiesto en la baja cooperatividad de sus trazas calorimétricas, indica que la reorganización de la distribución conformacional del propio dominio WW con la unión puede ser un factor determinante del patrón termodinámico observado. De igual manera, en las pocas estructuras cristalográficas disponibles para dominios

WW se confirma la existencia de moléculas de agua interfaciales, que también podrían estar contribuyendo al carácter exotérmico de la interacción.

A pesar de compartir un mismo patrón termodinámico general, en la Figura 3.4 se observa que existen diferencias importantes en las contribuciones entálpicas y entrópicas entre los distintos ligandos. Sin embargo estas diferencias debido a efectos de compensación entalpía-entropía, estas diferencias no se ponen de manifiesto en los valores de las constantes de unión, que, como ya hemos comentado, son similares para todos los ligandos.

Como hemos indicado, el motivo central **PPPY** aislado contribuye un elevado porcentaje, no sólo de la energía libre de unión, sino también, sorprendentemente, de las contribuciones entrópicas y entálpicas. Los aminoácidos adyacentes a este motivo en los distintos ligandos establecen interacciones adicionales con el dominio WW que resultan en entalpías de unión entre un 25% y un 45% más favorables. Estas contribuciones adicionales son mínimas en los ligandos **p53bp2** y **Ébola-a**, unos 15.2 kJ·mol<sup>-1</sup> y 12.7 kJ·mol<sup>-1</sup>, respectivamente, mientras que en los demás dominios L víricos suponen entre 27.8 y 36.0 kJ·mol<sup>-1</sup> de la entalpía en la interacción.

La energética de unión del ligando **Ebola-a** es diferente a la interacción de los otros dominios L con hNedd4-WW3. Además de una entalpía de unión parecida a la del ligando **p53bp2**, la contribución entrópica también es unos 20 kJ·mol<sup>-1</sup> más baja y más favorable que en las interacciones los otros ligandos víricos. A pesar de que la compensación entre entalpía y entropía resulta en una afinidad similar a la de los otros dominios L, las diferencias en el balance energético del ligando **Ébola-a** indican un comportamiento diferente de hNedd4-WW3 en el reconocimiento de este ligando.

Entre las interacciones de los ligandos **HTLV1-a** (**PPPY**VEPTAP) y **HTLV1-b** (SDPQI**PPPY**VEP), la mayor diferencia se produce en la compensación entropía-entalpía. La interacción de ambos ligandos presenta un término entrópico unos 8.5 kJ·mol<sup>-1</sup> menos favorable para **HTLV1-a**. En ese caso, es lógico pensar que la mayor entropía para **HTLV1-a** se deba a la solvatación de su extremo C-terminal (motivo PTAP) que, probablemente, se encuentre expuesto al solvente en el complejo. Las interacciones de los ligandos **Marburgo** y **Rabia** presentan una

afinidad muy parecida a la de **HTLV1-b**. Aunque se observan algunas diferencias en las contribuciones termodinámicas de su unión, como son valores ligeramente mayores para la entalpía y entropía en el primero de los ligandos, ambos términos se compensan y sus afinidades son comparables.

En definitiva, el estudio de la termodinámica de unión de los dominios L al dominio hNedd4-WW3 revela que los ligandos seleccionados interaccionan de forma parecida, con algunas excepciones. Por un lado, los ligandos **Ébola-b** y **HTLV1-a** presentan una afinidad menor que sus ligandos homólogos con los motivos PPxY centrados en la secuencia, revelando que los restos que flanquean dicho motivo son necesarios para establecer interacciones adicionales que estabilizan los complejos. Por otro lado, el ligando **Ébola-a** presenta una afinidad por hNedd4-WW3 similar a la de los otros dominios L, pero las magnitudes de las fuerzas termodinámicas de su unión, indican que su modo de unión es distinto del de los otros dominios L.

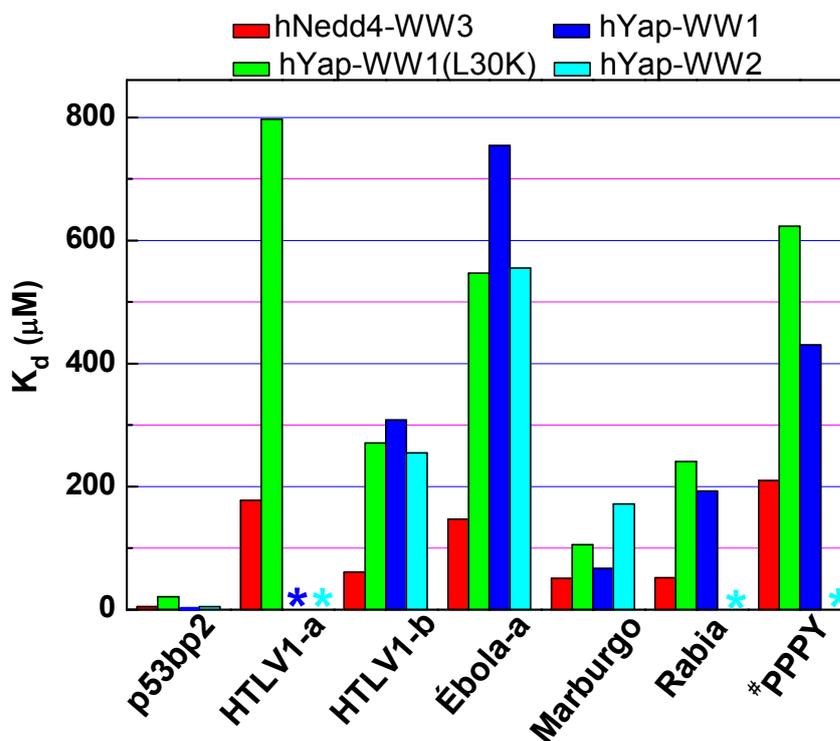
### 3.3. Especificidad de unión de los dominios L a diferentes dominios WW

En la Figura 3.1 se muestra un alineamiento múltiple de las secuencias de los tres dominios WW y el mutante empleados en este estudio: hNedd4-WW3, hYap-WW1, hYap-WW1(L30K) y hYap-WW2. Como se puede observar, además de las características que se mencionan en el Capítulo 1 (Apartado 1.3.1) para los dominios WW de clase I, sus secuencias presentan una gran similitud (superior al 60 %). La segunda y tercera hebras de los cuatro dominios WW, donde se encuentran los principales restos que definen el sitio de unión (Pires et al. 2001), están compuestas prácticamente por los mismos restos con excepción de las posiciones 28, 30 y 35. En la posición 28, el dominio hNedd4-WW3 presenta una fenilalanina en lugar de la tirosina de los otros dominios, que es una variación conservativa, una cadena lateral aromática por otra. En la posición 30, la mutación L30K introduce una carga positiva (lisina) en hYap-WW1(L30K) en lugar de la cadena alifática que se observa en los otros dominios. En la posición 35 del dominio hYap-WW1 y su mutante hYap-WW1(L30K) se encuentra un resto polar (glutamina) en lugar de la carga positiva (lisina) de los dominios hNedd4-WW3 y hYap-WW2.

Con el objeto de evaluar la relevancia de estas diferencias en la especificidad de unión de estos dominios, hemos analizado sus interacciones con los ligandos seleccionados en condiciones análogas a las descritas para el dominio hNedd4-WW3. Todos los datos obtenidos se han resumido en la Tabla 3.2, junto con los datos procedentes del análisis de las interacciones con hNedd4-WW3.

### 3.3.1. Análisis termodinámico de las interacciones de los dominios hNedd4-WW3, hYap-WW1, hYap-WW1(L30K) y hYap-WW2 con los dominios L víricos

En la Figura 3.5 se han representado en forma de diagramas de barras las constantes de disociación de los distintos ligandos y los dominios WW.



**Figura 3.5. Representación en forma de barras de las constantes de disociación ( $K_d$ ).** Los valores mostrados han sido determinados mediante espectroscopia de fluorescencia e ITC para la interacción de ligandos víricos, del ligando **p53bp2** y del ligando **PPPY** con diferentes los dominios WW: hNedd4-WW3 en rojo, mutante hYap-WW1(L30K) en verde, hYap-WW1 en azul oscuro y hYap-WW2 en azul claro. Los errores en la determinación de las  $K_d$  se estiman en torno al 10%. En el caso del ligando **Ébola-b**, la afinidad por los diferentes dominios WW es tan baja que no ha sido posible determinar experimentalmente el valor de  $K_d$  en ninguno de los casos. #Los valores de las  $K_d$  para el ligando **PPPY** se han determinado mediante experimentos de desplazamiento con el ligando **p53bp2** mediante ITC. \*Los asteriscos indican que estas  $K_d$  para los dominios hYap-WW1 y hYap-WW2 (en azul oscuro y azul claro, respectivamente) son tan elevadas que no se han podido determinar experimentalmente.

Al igual que sucedió para la interacción del ligando **Ébola-b** con hNedd4-WW3, tampoco se han podido determinar las constantes de disociación para este ligando con ninguno de los otros dominios WW y, por esa razón, no se han incluido en el diagrama de barras de la Figura 3.5. Otros complejos con interacciones de baja afinidad para las que no se ha podido determinar las constantes de disociación experimentalmente son los formados por el ligando **HTLV1-a** con hYap-WW1 y con hYap-WW2, el ligando **Rabia** con hYap-WW2 o el ligando **PPPY** con hYap-WW2. Es importante señalar que el valor de la constante de disociación del ligando **p53bp2** con hYap-WW1 ( $K_d=1.8 \mu\text{M}$ ) es ligeramente menor que el descrito previamente en la bibliografía ( $K_d=5.9 \mu\text{M}$ ), probablemente porque se determinó a un pH más bajo (pH 6.0) (Koepp et al. 1999a).

La afinidad de los dominios hNedd4-WW3, hYap-WW1(L30K) y hYap-WW1 por el ligando **PPPY** es baja, pero, por un lado, confirman esta secuencia como el motivo canónico de unión de los dominios WW de clase I y, por otro lado, pone de manifiesto la importancia que tienen los restos adyacentes a este motivo en las interacciones de los otros ligandos con los dominios WW.

El ligando **p53bp2** presenta afinidades elevadas por todos los dominios WW, de forma que todas las constantes de disociación determinadas para este ligando son inferiores a  $21 \mu\text{M}$  (Tabla 3.2). Por tanto, como ya hemos mencionado, se confirma que el ligando **p53bp2**, seleccionado como diana natural de hYap-WW1 (Españel & Sudol 2001), no muestra realmente una especificidad de unión significativa por este dominio.

La principal característica de los ligandos víricos es su especificidad por el dominio hNedd4-WW3 (Figura 3.5). Las interacciones de los otros tres dominios, hYap-WW1(L30K), hYap-WW1 y hYap-WW2, con los ligandos **HTLV1-a**, **HTLV1-b**, **Ébola-a** y **Rabia** tienen afinidades muy bajas, similares a las del ligando **PPPY**. El ligando **Marburgo** presenta una afinidad algo más elevada por los distintos dominios WW, es el que distingue menos, pero aún así muestra una preferencia por el dominio hNedd4-WW3. A pesar de que cada dominio WW presenta alguna peculiaridad en distintas posiciones del sitio de unión (Figura 3.1), sólo el dominio hNedd4-WW3 une específicamente a todos los ligandos víricos. Esto sugiere que, bien la principal diferencia de este dominio con respecto a los otros, es una

fenilalanina por una tirosina en la posición 28 (bolsillo xP), o bien, como ha sugerido el estudio la caracterización termodinámica de la unión del motivo **PPPY**, existen otras contribuciones adicionales a la unión, como podría ser alguna interacción entre el ligando y otras regiones próximas a los bolsillos xP y xY de hNedd4-WW3.

Al igual que sucedía en los procesos de unión de hNedd4-WW3, las interacciones de los ligandos ricos en prolinas con los dominios hYap-WW1, hYap-WW1(L30K) y hYap-WW2 también están gobernadas por un valor de entalpía favorable (negativo) a la que se opone parcialmente una contribución entrópica desfavorable (positiva) (Tabla 3.2). Esta similitud en la energética de unión nos lleva a pensar que se trata de un patrón termodinámico común para los dominios WW. Además, teniendo en cuenta que existen otros dominios modulares con una energética de unión similar a la descrita, como por ejemplo los dominios SH3 (Renzoni et al. 1996; Wang et al. 2001a; Ferreon & Hilser 2004; Palencia et al. 2004) o los dominios UEV (Palencia 2008; Schlundt et al. 2009; Im et al. 2010), este patrón termodinámico de unión podría ser común para los dominios modulares que reconocen secuencias ricas en prolinas.

Los ligandos **p53bp2** y **PPPY** no distinguen entre los diferentes dominios WW mientras que los dominios L víricos sí. De hecho, los valores de entalpía y entropía de unión de estos dos ligandos con los distintos dominios WW difieren menos de  $4 \text{ kJ}\cdot\text{mol}^{-1}$ , lo que confirma que son reconocidos inespecíficamente de forma similar por los cuatro dominios WW.

Debido a que la afinidad de los ligandos víricos por los dominios hYap-WW1, hYap-WW1(L30K) y hYap-WW2 es baja (Figura 3.5), sólo hemos podido determinar la energética de la unión de los dominios hYap-WW1 y hYap-WW1(L30K) con algunos dominios L. Los parámetros termodinámicos de estas interacciones obtenidos en titulaciones por desplazamiento son parecidos a los obtenidos para el ligando **PPPY** (Tabla 3.2). Este resultado pone de manifiesto que las interacciones adicionales responsables de la especificidad de unión de los dominios L por el dominio hNedd4-WW3 no se producen en los complejos formados por estos dominios L víricos con otros dominios WW.

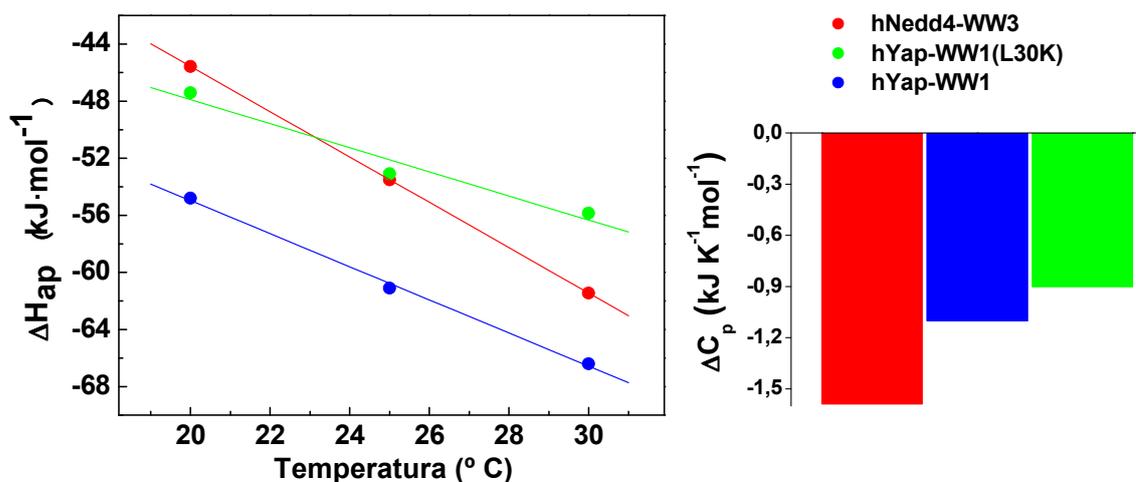
En definitiva, podemos confirmar que la afinidad de todos los dominios L por su diana natural, hNedd4-WW3, es superior a la afinidad por los demás dominios WW. Este resultado indica que las secuencias víricas seleccionadas presentan un cierto nivel de especificidad intrínseca por el dominio hNedd4-WW3.

#### 3.3.2. Determinación de la capacidad calorífica de unión mediante ITC

Hemos llevado a cabo la determinación de la capacidad calorífica de unión aparente de los dominios hNedd4-WW3, hYap-WW1 y hYap-WW1(L30K) con el ligando de mayor afinidad, el ligando **p53bp2**, para obtener una información más detallada de sus interacciones. Para ello hemos realizado experimentos de ITC a tres temperaturas diferentes (20, 25 y 30 °C) en tampón fosfato sódico 20 mM pH 7.0 y los resultados se han recogido en la Tabla 3.2.

En la Figura 3.6 se han representado los valores de entalpía de unión aparente de los tres dominios WW en función de la temperatura. En esta representación, se puede observar una dependencia lineal del valor de dicha magnitud con la temperatura para cada dominio WW. El valor de la pendiente de cada una de las rectas se calcula a partir del ajuste lineal de mínimos cuadrados y es una buena estimación de la capacidad calorífica de unión aparente en cada caso. Los resultados obtenidos para la capacidad calorífica de unión se muestran en el diagrama de barras de la Figura 3.6. Como se puede observar, existen diferencias significativas en el cambio de capacidad calorífica de unión entre los tres dominios: el dominio hNedd4-WW3 presenta el valor más negativo,  $\Delta C_{p,ap} = -1.59 \text{ kJ} \cdot (\text{K} \cdot \text{mol})^{-1}$ ; el mutante hYap-WW1(L30K) tiene un valor intermedio,  $\Delta C_{p,ap} = -1.10 \text{ kJ} \cdot (\text{K} \cdot \text{mol})^{-1}$ , y el dominio hYap-WW1 presenta el valor menos negativo,  $\Delta C_{p,ap} = -0.90 \text{ kJ} \cdot (\text{K} \cdot \text{mol})^{-1}$ .

La contribución más importante a los cambios de capacidad calorífica asociados a las interacciones proteína-ligando proviene de cambios en la solvatación de ambos cuando están libres con respecto al complejo que, a su vez, es proporcional a los cambios en las áreas accesibles al disolvente. Contribuciones adicionales, aunque generalmente de menor magnitud, pueden tener su origen en cambios conformacionales del dominio y del ligando o en el equilibrio de ionización acoplados a la unión (Cooper 2005).



**Figura 3.6.** Cálculo de la capacidad calorífica de unión a los dominios hNedd4-WW3, hYap-WW1(L30K) y hYap-WW1. A la izquierda se han representado las entalpías aparentes de unión obtenidas de los experimentos de calorimetría para el ligando p53bp2 con los dominios hNedd4-WW3 (color rojo), hYap-WW1(L30K) (color verde) y hYap-WW1 (color azul) en fosfato sódico 20 mM pH 7.0 a 20, 25 y 30  $^{\circ}\text{C}$ . De la pendiente de las rectas obtenemos el cambio de capacidad calorífica de unión aparente. En el panel derecho, se comparan los valores de capacidad calorífica de unión aparente para los tres dominios en el mismo código de colores.

Nuestros datos sugieren que el cambio de área accesible al disolvente debido a la interacción con el ligando p53bp2 es mayor para el dominio hNedd4-WW3 que para los dominios hYap-WW1 o hYap-WW1 (L30K). Este resultado, además, está de acuerdo con los estudios realizados con el dominio homólogo de hNedd4-WW3 en *Drosófila* (dNedd4-WW3), muy similar al dominio humano utilizado en este estudio, y un péptido rico en prolina de la proteína *Comm* (Kanelis et al. 2006). Para este complejo se ha obtenido un cambio de área accesible al solvente tras la unión del ligando que es considerablemente mayor ( $895 \text{ \AA}^2$ ) que el área de hYap-WW1 con distintos ligandos ricos en prolina, cuyos valores están entre  $620$  y  $730 \text{ \AA}^2$  (Pires et al. 2001). En el caso de la formación complejo *Comm*/dNedd4-WW3, que el cambio de área accesible al solvente sea mayor, se ha relacionado con una serie de contactos adicionales responsables de la elevada afinidad entre ambos y que se producen entre los extremos del ligando y regiones próximas a los bolsillos xP y xY del dominio WW (Kanelis et al. 2006). En nuestro caso, el patrón y la magnitud de las contribuciones termodinámicas para el reconocimiento del ligando **p53bp2** por los diferentes dominios WW es muy parecido y, como hemos visto, este ligando no presenta una preferencia de unión

clara por ninguno de ellos. Parece, pues, que el origen de las diferencias en capacidad calorífica de unión además de estar relacionado con el cambio de área accesible al solvente también podría guardar relación con otros procesos adicionales. Así, otros efectos como un equilibrio de ionización acoplado a la unión o un cambio conformacional del dominio podrá contribuir al cambio en el  $\Delta C_p$ . Este último efecto, además, podría relacionar se con la elevada flexibilidad conformacional de estos dominios WW, en cuyo caso, al igual que la capacidad calorífica de unión a **p53bp2**, la flexibilidad conformacional más elevada es para el dominio hNedd4-WW3 (Apartado 2.3, Capítulo 2)

### 3.4. Caracterización estructural mediante RMN

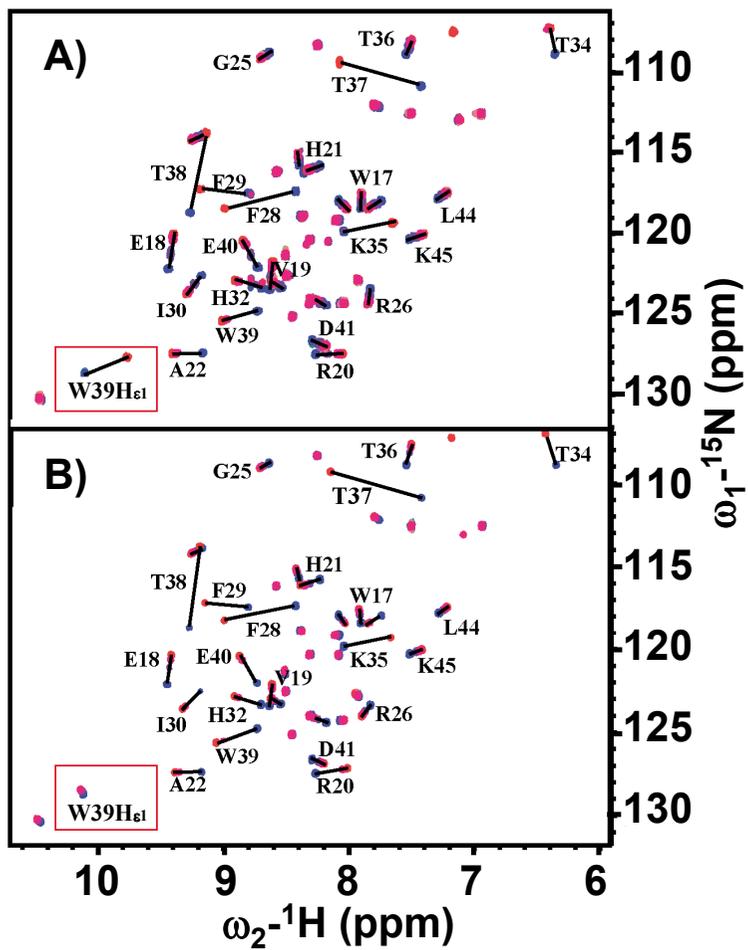
Para profundizar en la comprensión del reconocimiento de los dominios L por hNedd4-WW3, hemos resuelto las estructuras de los complejos formados por el dominio hNedd4-WW3 y los ligandos **HTLV1-b** y **Ebola-a**, que presentan diferencias significativas en sus parámetros termodinámicos de unión. Este trabajo se ha realizado en colaboración con el Dr. Francisco J. Blanco (Centro Nacional de Investigaciones Oncológicas (CNIO) de Madrid), y la Dra. Maria J. Macias (Instituto de Investigaciones Biomédicas (IRB) en el Parque Científico de Barcelona).

#### 3.4.1. Identificación de los residuos del dominio hNedd4-WW3 implicados en el reconocimiento de los dominios L víricos HTLV1-b y Ébola-a

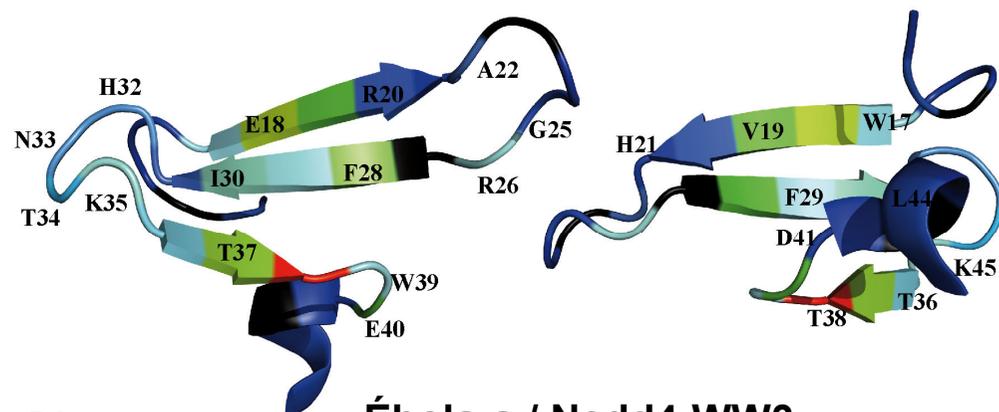
Con el objetivo de obtener información sobre qué residuos están implicados en la unión, hemos realizado sendas titulaciones del dominio hNedd4-WW3 con los péptidos **HTLV1-b** y **Ébola-a**. Las titulaciones se han realizado a 15 °C en tampón fosfato sódico 20mM pH 7.0, añadiendo cantidades conocidas de cada ligando. En el caso del ligando **HTLV1-b** la relación proteína:ligando alcanzada para saturar la formación de complejo es 1:3, mientras que para el **Ébola-a** se ha llegado a una relación 1:10, debido a la menor afinidad de unión. Los cambios en los desplazamientos químicos del dominio WW se han seguido mediante la adquisición de espectros  $^{15}\text{N}$ -HSQC (Apartado 6.7 del Capítulo 6), que muestran las frecuencias de cruce  $^1\text{H}$ - $^{15}\text{N}$  del esqueleto polipeptídico. Para poder asignar los espectros  $^{15}\text{N}$ -HSQC, se ha realizado la asignación secuencial de las frecuencias

en el conjunto de experimentos de triple resonancia CBCA(CO)NH y CBCANH (Apartado 6.7 del Capítulo 6).

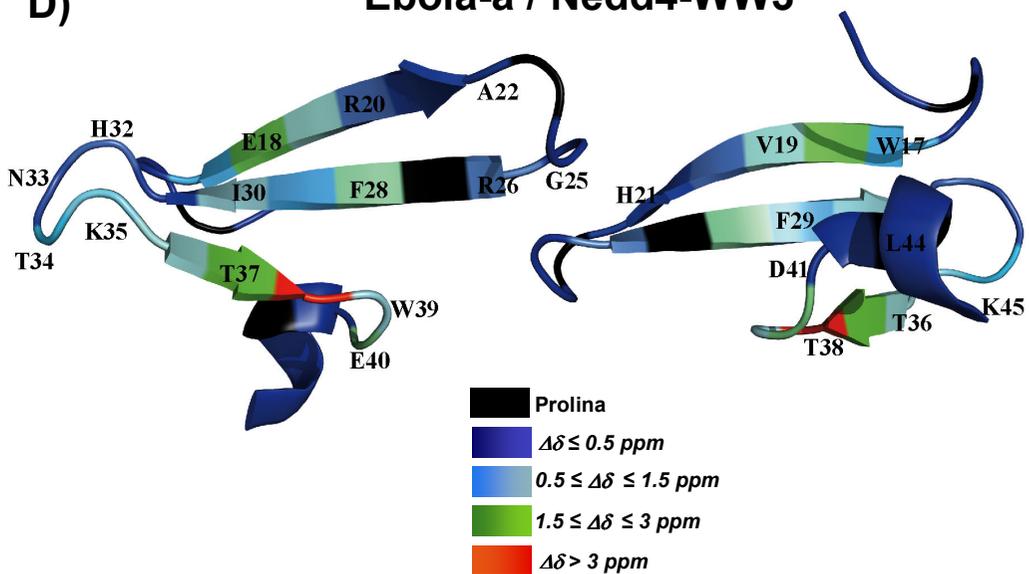
En la Figura 3.7 se muestra la superposición de los espectros  $^{15}\text{N}$ -HSQC del dominio hNedd4-WW3 con los ligandos **HTLV1-b** y **Ébola-a**, paneles A y B respectivamente. Es interesante señalar que los cambios en los desplazamientos químicos del dominio hNedd4-WW3 presentan un patrón muy parecido en las dos titulaciones. Esto nos indica que la interacción de hNedd4-WW3 con los dos ligandos se produce de manera similar. En los paneles C y D (Figura 3.7), se han representado dos vistas del dominio WW en las que se han coloreado los enlaces peptídicos en función de la intensidad del cambio en los desplazamientos químicos ( $\Delta\delta$ ) para cada complejo. En estas representaciones podemos observar los cambios que se producen en los desplazamientos químicos  $^1\text{H}$ - $^{15}\text{N}$  de la cadena principal en los restos que conforman los dos bolsillos del sitio de unión del dominio hNedd4-WW3 para cada complejo. En ambos casos podemos comprobar cambios en los desplazamientos químicos  $^1\text{H}$ - $^{15}\text{N}$  de la cadena principal en los restos que conforman el sitio canónico de unión (los bolsillos xP y xY) y algunos restos contiguos en la secuencia, como son: H32, T34, K35 en el segundo giro del dominio; F28, F29 y I30 en la hebra  $\beta$ 2; T36, T37 y T38 en la hebra  $\beta$ 3; y en los restos W39 y E40 del giro que hay tras la hebra  $\beta$ 3. Además, en ambas titulaciones se puede apreciar que existen otros restos que no constituyen los bolsillos xP y xY y que, sin embargo, son los que presentan un cambio en el desplazamiento químico mayor debido a la unión de los dos ligandos. Algunos de estos restos se encuentran en la hebra  $\beta$ 1 (E18, V19 y R20) y otros en el primer giro (R26). Los cambios en los desplazamientos químicos de estos restos en el entorno del sitio canónico de unión, podrían indicar que en dichas se regiones del dominio hNedd4-WW3 se producen algunos contactos adicionales con ambos ligandos víricos, como se ha sugerido a partir del análisis de la energética de unión.



**C) HTLV1-b / Nedd4-WW3**



**D) Ébola-a / Nedd4-WW3**



**Figura 3.7. Cambios en los desplazamientos químicos del dominio hNedd4-WW3 producidos por la titulación de los ligandos HTLV1-b y Ébola-a.** En los paneles A) y B), a la izquierda, se muestran la superposición de los espectros  $^{15}\text{N}$ -HSQC del dominio hNedd4-WW3 durante la titulación con los ligandos HTLV1-b y Ébola-a, respectivamente. Los picos de cruce  $^1\text{H}$ - $^{15}\text{N}$  de la cadena principal que han variado sus frecuencias durante las titulaciones se han rotulado con el número del residuo correspondiente. El código de colores empleado en cada espectro según la relación proteína:ligando aumenta es: azul, morado, rosa y rojo. En la titulación para con el ligando HTLV1-b (Panel A) los espectros recogidos se corresponden con las relaciones proteína:ligando 1:0.0, 1:0.5, 1:1.5 y 1:3.0. En la titulación con el ligando Ébola-a (Panel B) las relaciones proteína:ligando son 1:0.0, 1:2.5, 1:5 y 1:10.0. En los paneles C) y D), a la derecha, se ha representado dos vistas del dominio hNedd4-WW3, desde el sitio de unión y desde la parte posterior, coloreadas según la magnitud de los cambios en los desplazamientos químicos ( $\Delta\delta$ ) para la titulación del ligando HTLV1-b y Ébola-a, respectivamente. El código de colores empleado es una gradación de colores que van desde el azul al rojo para valores  $\Delta\delta$  entre 0.5 y 4 ppm (como muestra la leyenda). Las prolinas para las que no hay información están coloreadas en negro. Los cambios en los desplazamientos químicos se han calculado según la ecuación:  $\Delta\delta_{\text{ppm}} = \sqrt{(\Delta\delta_{\text{HN}})^2 + (\Delta\delta_{\text{N}}/6.51)^2}$ .

No obstante, a pesar de que los patrones de desplazamientos químicos son muy similares entre los dos complejos, existe una diferencia muy significativa, señalada mediante un cuadrado rojo en la Figura 3.7, asociada al protón del grupo indol de la cadena lateral del resto W39 ( $\text{W39H}_{\varepsilon 1}$ ). Esta es una de las cadenas laterales que componen el bolsillo xP donde se intercalan las prolinas de los ligandos. En los espectros de la titulación con el péptido **Ébola-a** se aprecia que el desplazamiento químico de este protón prácticamente no se ve afectado ( $\Delta\delta < 0.5$ ), mientras que en la titulación con el ligando **HTLV1-b**, el valor de  $\Delta\delta$  para este protón es bastante elevado ( $\Delta\delta \sim 2$ ). En otras titulaciones de dominios WW, se ha utilizado la variación del desplazamiento químico de este protón para seguir la unión entre los dominios WW y sus respectivos ligandos (Koepf et al. 1999a; Ramirez-Espain et al. 2007). Esta diferencia entre las 2 titulaciones con los ligandos **HTLV1-b** y **Ébola-a**, sugiere que las prolinas de cada ligando interactúan de forma distinta en el bolsillo xP del dominio hNedd4-WW3.

### 3.4.2. Asignación de los espectros de RMN del dominio hNedd4-WW3

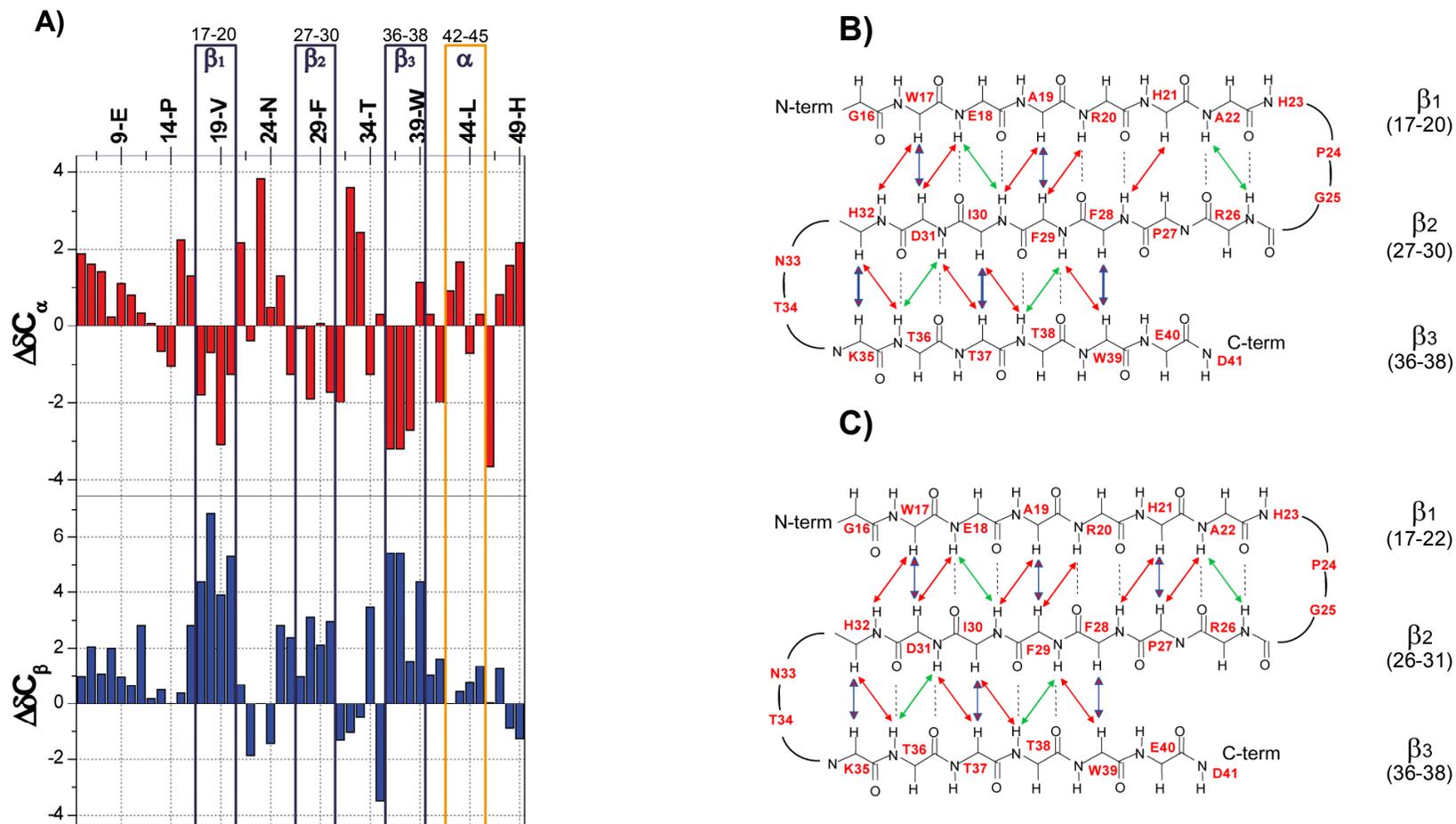
La asignación del complejo formado con el ligando **HTLV1-b** se ha realizado como se describe en el Apartado 6.7.2 del Capítulo 6 (Materiales y Métodos). Una vez completada la asignación del complejo con el ligando **HTLV1-b**, se ha utilizado la información obtenida en este primer complejo para realizar la asignación del complejo con el ligando **Ébola-a**, ya que los valores de  $\Delta\delta$  para el dominio hNedd4-WW3 varían poco de un complejo al otro. Esta similitud ha

permitido superponer la asignación del dominio WW con el ligando **HTLV1-b** sobre los espectros bidimensionales del mismo dominio con el ligando **Ébola-a** y corregir manualmente las asignaciones cuyas frecuencias estén ligeramente desviadas. Finalmente, se ha completado la asignación de los núcleos  $^1\text{H}$  del ligando **Ébola-a** y los NOEs intermoleculares entre éste y el dominio WW.

### 3.4.3. Estructura secundaria de los complejos

La estructura secundaria del dominio en ambos complejos se ha elucidado a partir de los desplazamientos conformacionales de los carbonos  $^{13}\text{C}_\alpha$  y  $^{13}\text{C}_\beta$  (Wishart & Sykes 1994), así como de los patrones de NOEs secuenciales (entre núcleos de restos  $i$  e  $i+1$ ) y de medio alcance (entre núcleos de restos  $i$  e  $i+2$ ,  $i+3$  ó  $i+4$ ) (Wüthrich 1986). En el Panel A de la Figura 3.8 se muestra el patrón de desplazamientos químicos de los  $\text{C}_\alpha$  y  $\text{C}_\beta$  para las 3 hebras  $\beta$  de hNedd4-WW3 formando un complejo con **HTLV1-b**. Este patrón comprende las posiciones 17-20, 27-30 y 36-38 del dominio. Por otro lado, en los paneles B y C, se representan la disposiciones de las correlaciones NOE de largo alcance entre carbonos  $\alpha$  ( $d_{\alpha\alpha(i,j)}$ , en color morado), entre carbonos  $\alpha$  y protones amina ( $d_{\alpha\text{N}(i,j)}$ , en rojo) y entre protones amina ( $d_{\text{NN}(i,j)}$ , en verde) de las distintas hebras  $\beta$  del dominio hNedd4-WW3 en cada complejo. El patrón de NOEs corresponde al de una lámina  $\beta$  antiparalela, aunque se observan pequeñas diferencias entre los dos complejos. Las hebras  $\beta_1$  y  $\beta_2$  para el complejo con el ligando **Ébola-a** son más largas, ya que se pueden observar los NOEs  $d_{\alpha\alpha(21,27)}$  y  $d_{\alpha\text{N}(22,27)}$  (Figura 3.8C) que no aparecen en el otro complejo (Figura 3.8B). Con ambos resultados, los desplazamientos químicos de  $\text{C}_\alpha$  y  $\text{C}_\beta$  y la correlación de NOEs a larga distancia, hemos definido las diferentes restricciones para la lámina  $\beta$  del dominio WW en cada complejo. Así, finalmente para el complejo con el ligando **HTLV1-b**, la hebra  $\beta_1$  va de la posición 17 a 21, la hebra  $\beta_2$  de la posición 27 a 30 y la hebra  $\beta_3$  de 36 a 38, mientras que para el complejo con el ligando **Ébola-a**, la hebra  $\beta_1$  va de la posición 17 a 22, la hebra  $\beta_2$  de la posición 26 a 30 y la hebra  $\beta_3$  de 36 a 38.

Además, al igual que en la estructura del dominio homólogo WW3 de la isoforma de ratón mNedd4-2 (PDB 1WR7), se ha observado un patrón de NOEs de



**Figura 3.8. Restricciones del dominio hNedd4-WW3 utilizadas en el cálculo de las estructuras de sus complejos por RMN. A)** Relación de desplazamientos químicos con la conformación de los núcleos  $^{13}\text{C}_\alpha$  ( $\Delta\delta\text{C}_\alpha = \delta\text{C}_{\alpha, \text{observado}} - \delta\text{C}_{\alpha, \text{desplegado}}$ , en ppm) en color rojo y  $^{13}\text{C}_\beta$  ( $\Delta\delta\text{C}_\beta = \delta\text{C}_{\beta, \text{observado}} - \delta\text{C}_{\beta, \text{desplegado}}$ , en ppm) en color azul, en función de la secuencia de la proteína, a pH 7.0 y 15°C. Los valores de  $\delta\text{C}_{\alpha, \text{desplegado}}$  y  $\delta\text{C}_{\beta, \text{desplegado}}$ , se tomaron de (Wishart and Sykes 1994). Las restricciones de ángulos diedros impuestas a partir de la estructura secundaria durante el cálculo, aparecen en recuadros azul oscuro para los restos dispuestos en hebras  $\beta$  y naranjas para los restos en hélice  $\alpha$ . Diagramas esquemáticos de los posibles enlaces de hidrógeno entre las hebras  $\beta$  antiparalelas encontrados durante la asignación **B)** del complejo HTLV1-b / hNedd4-WW3 y **C)** del complejo Ébola-a / hNedd4-WW3. Una flecha entre protones indica una correlación NOE de largo alcance  $d_{\alpha\alpha(i,j)}$  (en morado),  $d_{\alpha N(i,j)}$  (en rojo) y  $d_{NN(i,j)}$  (en verde). Las líneas discontinuas representan los enlaces de hidrógeno formados entre las hebras que se han tenido en cuenta para el cálculo de estructura.

medio alcance de una pequeña hélice  $\alpha$  en los dos complejos entre los restos 43-45.

Los elementos de estructura secundaria definidos a partir de este análisis se utilizaron para el cálculo de la estructura tridimensional de los dos complejos.

#### 3.4.4. Asignación de los ligandos en los espectros de RMN

En cada uno de los complejos formados por el dominio hNedd4-WW3 se han identificado unos 50 NOEs intermoleculares. Estas señales NOE se interpretan como restricciones de distancia entre el dominio WW y el ligando. La cantidad de NOEs intermoleculares asignados para ambos complejos es similar a la encontrada para otras estructuras bien definidas de complejos de dominios WW (Pires et al. 2001). En la Tabla 3.3 se comparan los principales NOEs intermoleculares encontrados para cada uno de los complejos. Los restos del dominio y del ligando se han denominado según sus secuencias en la Figura 3.1 y la Tabla 3.1, respectivamente. Además, se ha empleado un símbolo prima tras el número que indica la posición de cada resto en las secuencias de los ligandos para distinguirlos de los aminoácidos del dominio.

En el bolsillo xP, los NOEs encontrados relacionan los mismos restos del dominio WW con aminoácidos equivalentes en cada uno de los ligandos. La mayoría son contactos entre las dos prolinas centrales de los ligandos, P4' y P5', con las cadenas aromáticas de F28 y W39 del dominio WW. Aunque en los dos complejos se han asignado varios contactos entre el indol del W39 y los restos en las posiciones 2' y 3' de los dos ligandos, cabe destacar que en el complejo con el ligando **HTLV1-b** el protón H $\epsilon_1$  del W39 presenta un mayor número de NOEs con los restos Q2' e I3' (incluso con la P4'). Estas señales indican que la proximidad de estos aminoácidos del ligando **HTLV1-b** puede ser la responsable del cambio del entorno químico para la cadena lateral del W39, observado durante la titulación de este ligando (Figura 3.7).

En el bolsillo xY se aprecian más diferencias entre los dos complejos. Aunque los restos I30, H32, K35 y T37 del dominio hNedd4-WW3 presentan NOEs con la Y7' en los dos complejos, los otros NOEs asignados implican átomos diferentes en cada complejo. Además, el resto Y7' del ligando **Ébola-a** presenta

algunos NOEs con el resto T36 que no se observan en el complejo **HTLV1-b**. Todas estas diferencias indican que la cadena lateral de la Y7' de cada ligando presenta una orientación distinta dentro de este bolsillo xY.

**Tabla 3.3. Interacciones intermoleculares (NOEs) observadas en los complejos formados por el dominio hNedd4-WW3 con los ligandos peptídicos HTLV1-b y Ébola-a**

hNedd4-WW3		HTLV1-b SDPQI <b>PPPY</b> VEP	Ébola-a ILPTA <b>PPEY</b> ME
Restos	Protón	Restos	Protón
W39	H $\epsilon_1$	Q2' ó T2'	H $\beta_1$ , H $\beta_2$
	H $\epsilon_1$	I3' ó A3'	H $\alpha$ , H $\beta$ , #H $\gamma$
	H $\delta_1$		H $\alpha$
	H $\zeta_2$		--
	H $\epsilon_1$		H $\alpha$ , #H $\beta$
	H $\delta_1$	P4'	H $\delta_1$ , H $\delta_2$
H $\zeta_2$	H $\delta_1$ , H $\delta_2$		
F28	#H $\delta$	P5'	H $\delta_1$ , H $\delta_2$ , H $\beta_1$ , H $\beta_2$ , H $\gamma_1$ , H $\gamma_2$
	#H $\epsilon$		H $\delta_1$ , H $\delta_2$ , H $\beta_1$ , H $\beta_2$ , H $\gamma_1$ , H $\gamma_2$
I30	#H $\gamma_1$	Y7'	#H $\delta$
	#H $\gamma_2$		#H $\delta$ , #H $\epsilon$
HN	--		
H32	H $\alpha$		#H $\epsilon$
	#H $\epsilon$		#H $\epsilon$
K35	H $\alpha$		#H $\epsilon$
	H $\beta_1$		#H $\epsilon$ , #H $\delta$
	H $\beta_2$		#H $\epsilon$ , #H $\delta$
	H $\delta_1$		#H $\epsilon$
	H $\delta_2$		#H $\epsilon$
	H $\gamma_1$	#H $\epsilon$	
T36	HN	--	
	H $\alpha$	--	
T37	#H $\gamma_2$	HN, H $\alpha$ , H $\beta_1$ , H $\beta_2$ , #H $\delta$ , #H $\epsilon$	
R20	H $\epsilon$	E9'	H $\gamma_2$ , H $\gamma_3$

#Designa la señal de desplazamiento químico para un conjunto de protones que se consideran químicamente equivalentes.

Los datos recogidos en Tabla 3.3 indican que la cadena lateral de la R20 del dominio WW interacciona con la cadena lateral del resto E9' en los dos complejos. Esta interacción electrostática se refleja en importantes cambios en el desplazamiento químico del aminoácido R20 en las titulaciones con ambos ligandos (Figura 3.7).

En definitiva, podemos concluir que los datos procedentes de la asignación de los NOEs intermoleculares, al igual que la energética de unión o el cambio de los desplazamientos químicos durante la titulación, sugieren la existencia de diferencias significativas en el modo de unión de los ligandos **HTLV1-b** y **Ebola-a** al dominio hNedd4-WW3.

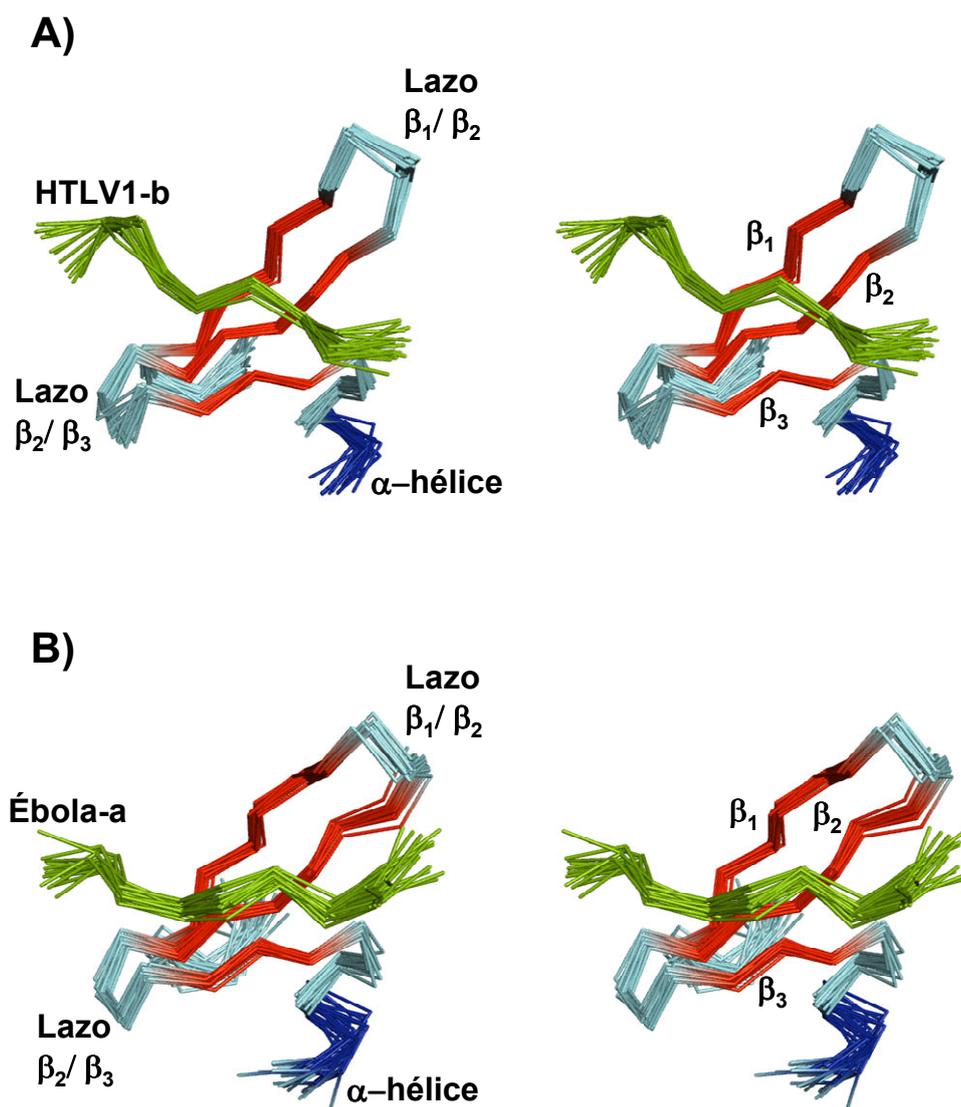
#### 3.4.5. Cálculo y evaluación de la calidad de las estructuras

Las estructuras tridimensionales en disolución de los dos complejos del dominio hNedd4-WW3 se han calculado tal y como se describe en el Apartado 6.7.3 de Materiales y Métodos. Se han obtenido 597 y 513 restricciones de distancia entre pares de núcleos a partir de los picos de cruce presentes en los espectros 2D  $^1\text{H}$ -NOESY registrados en 10%  $\text{D}_2\text{O}$ :90%  $\text{H}_2\text{O}$  para los complejos con los ligandos **HTLV1-b** y **Ébola-a**, respectivamente. El volumen de los picos de cruce de los espectros 2D  $^1\text{H}$ -NOESY se ha medido utilizando la subrutina de integración máxima del programa XEASY (Bartels et al. 1995). También se han utilizado restricciones para los ángulos diedros  $\phi$  y  $\psi$  de la cadena principal obtenidas a partir del análisis de los desplazamientos químicos, así como restricciones para los enlaces de hidrógeno obtenidas del análisis de los patrones de NOEs en la estructura secundaria. El número de ángulos diedros restringidos en el cálculo de cada estructura ha sido 40 para el complejo del **HTLV1-b** y 50 para el complejo **Ébola-a**. El número de restricciones para los enlaces de hidrógeno ha sido de 16 para los dos complejos.

Para realizar el cálculo se ha utilizado el programa informático *Structcalc* desarrollado en el laboratorio de María J. Macías (<http://maciasnmr.net/>). Este programa otorga mayor importancia al valor de los NOEs inequívocos y utiliza todas las restricciones experimentales, los puentes de hidrógeno y las restricciones de los ángulos diedros, en un proceso de refinamiento en agua. Así, para cada estructura se ha calculado un conjunto de 120 modelos en dos fases: primero simulando la dinámica de torsión de los ángulos en vacío y después simulando un solvente acuoso. De los 120 modelos calculados se han seleccionado los 20 de menor energía que satisfacen adecuadamente las restricciones experimentales. En

la Figura 3.9 se muestran las representaciones estereoscópicas de los conjuntos de 20 modelos para cada complejo.

Para validar las estructuras determinadas se ha comprobado que las estructuras cumplen distintos criterios de calidad. La Tabla 3.4 recoge la estadística de los principales parámetros estructurales de las 20 estructuras de menor energía en los dos complejos.



**Figura 3.9. Representaciones de las 20 estructuras de menor energía resueltas mediante RMN de los complejos formados por hNedd4-WW3. Representación estereográfica de la superposición de los átomos del esqueleto de las veinte estructuras finales: A) del complejo con el ligando HTLV1-b y B) del complejo con el ligando Ébola-a. Los ligandos se muestran en verde y las regiones correspondientes a las hebras  $\beta$  del dominio WW se indican en rojo y a la  $\alpha$ -hélice del extremo C-terminal en azul.**

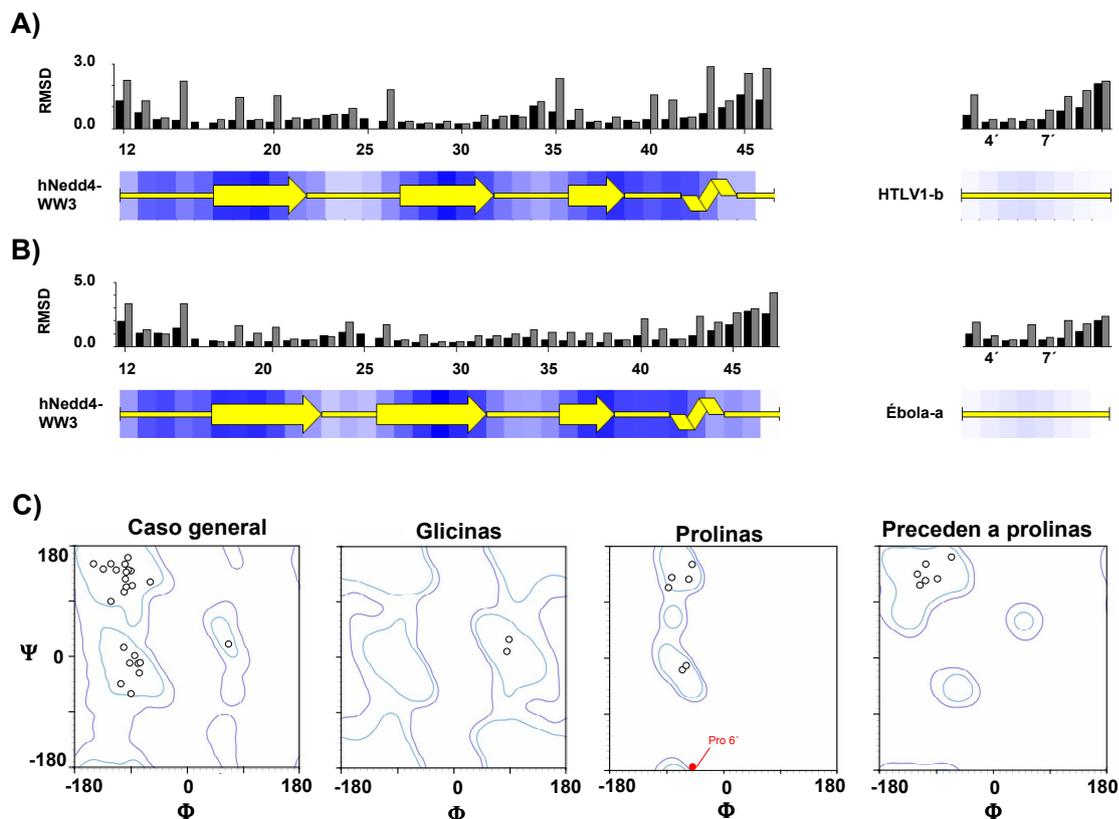
Tabla 3.4. Estadística de los principales parámetros estructurales de los conjuntos de 20 estructuras de los complejos HTLV1-b/hNedd4- WW3 y Ébola-a/hNedd4- WW3

	HTLV1-b	Ébola-a
<b>Restricciones usadas para el cálculo &lt;SA&gt;<sup>a</sup></b>		
Intra-residuo	0	0
Secuencial (  i-j  = 1)	259	236
Medio alcance ( 1 <  i-j  < 4)	90	72
Largo alcance (  i-j  > 4)	264	209
Inter-molecular (  i-j  > 4)	50	52
No ambiguos	All	All
Ambiguos	0	0
Ángulos diedros	50	40
Enlaces de hidrógeno	16	16
Todos	663	569
Promedio de restricciones por residuo (43)	15.4	13.2
<b>Desviaciones para datos experimentales (Å)<sup>b</sup></b>		
NOE:	0.01014 ± 3.2 x10 <sup>-4</sup>	0.004357 ± 3.8 x10 <sup>-4</sup>
Enlaces (Å)	0.00833 ± 1.9 x10 <sup>-4</sup>	0.004202 ± 1.6 x10 <sup>-4</sup>
Ángulos (°)	0.831 ± 0.028	0.5905 ± 0.02
<b>R.M.S.D. Precisión de las coordenadas (Å)<sup>c</sup></b>		
Átomos N, C <sub>α</sub> y *CO de la estructura secundaria	0.12	0.27
Átomos pesados de la estructura secundaria	0.51	0.68
Átomos N, C <sub>α</sub> y *CO de la cadena principal	0.60	0.83
Átomos pesados en todos los restos	1.11	1.31
<b>Energía potencial por CNS (kcal mol<sup>-1</sup>)</b>		
Energía total <sup>d</sup>	-1260 ± 48	-1535 ± 30
Electrostatica	-1633 ± 54	-1748 ± 39
van der Waals	-122 ± 12	-135 ± 13
Enlaces	51 ± 2	13 ± 1.0
Angulos	142 ± 10	71.5 ± 5.7
<b>Calidad estructural (%residuo): representación de Ramachandran</b>		
Regiones más favorables (%)	91.2	84.7
Regiones permitidas adicionalmente (%)	7.3	14.1
Regiones permitidas generosamente (%)	1.5	0.9
Regiones no permitidas (%)	0	0.3

<sup>a</sup><SA> se refiere al conjunto de 20 estructuras con menor energía de cada complejo. <sup>b</sup>Ninguna de las estructuras calculadas incumple las restricciones de distancia con valores superiores a 0.3 Å. <sup>c</sup>R.M.S.D. entre el conjunto de estructuras <SA> y la estructura de menor energía. <sup>d</sup>E<sub>L-J</sub> es la energía de van der Waals Lennard-Jones calculada usando los parámetros CHARMMMPARMALLH6. La función E<sub>L-J</sub> no está incluida en el cálculo de la estructura.

Las desviaciones del conjunto de estructuras se cuantifican mediante la desviación cuadrática media (RMSD, del inglés *Root Mean Square Deviation*) entre las coordenadas atómicas del conjunto promediado de las 20 estructuras con respecto a la de menor energía. En los Paneles A y B de la Figura 3.10 se han representado los diagramas de barras de la RMSD. En términos generales, los elementos de estructura secundaria del dominio WW, particularmente las hebras  $\beta$ , se encuentran mejor definidos en los dos complejos que el resto de elementos tanto del dominio como del ligando. El complejo del dominio hNedd4-WW3 con el ligando **HTLV1-b** (Figura 3.10A) está mejor definido que el complejo con el ligando **Ébola-a** (Figura 3.10B). Como se muestra en la Tabla 3.4, las RMSD promedio de los átomos pesados (promedio de 20 estructuras) son de 1.11 Å para el complejo con **HTLV1-b** y 1.31 Å para el complejo con **Ébola-a** y las RMSD promedio de las cadena polipeptídica principal son 0.60 Å para el complejo con el ligando **HTLV1-b** y 0.83 Å para con el complejo con ligando **Ébola-a**. En cualquier caso, los valores para los dos complejos se sitúan dentro del intervalo habitual para este tipo de cálculos estructurales:  $\text{RMSD} \leq \sim 0.8 \text{ \AA}$  para la cadena principal y  $\text{RMSD} \leq \sim 1.5 \text{ \AA}$  para todos los átomos pesados (Candel 2008a).

En el Panel C de la Figura 3.10 se muestra, a modo de ejemplo, el mapa de Ramachandran de la estructura de menor energía del complejo de hNedd4-WW3 con el ligando **Ébola-a**. En esta representación el mapa de Ramachandran se ha dividido en cuatro categorías según los ángulos estadísticamente permitidos: uno general y otros tres específicos para las glicinas, las prolina o los restos que preceden a una prolina en la secuencia (Laskowski et al. 1996; Lovell et al. 2003). En la Tabla 3.4 se resumen los datos estadísticos obtenidos a partir de la representación de Ramachandran de las 20 estructuras de cada complejo. Como se puede observar, la mayoría de los restos se encuentran en regiones favorecidas (91.2% para el **HTLV1-b** y 84.7% para el **Ébola-a**) o permitidas adicionalmente (7.3% para el **HTLV1-b** y 14.1% para el **Ébola-a**), mientras que sólo el 0.3% de los aminoácidos del complejo con el ligando **Ébola-a** se encuentra en regiones no permitidas.

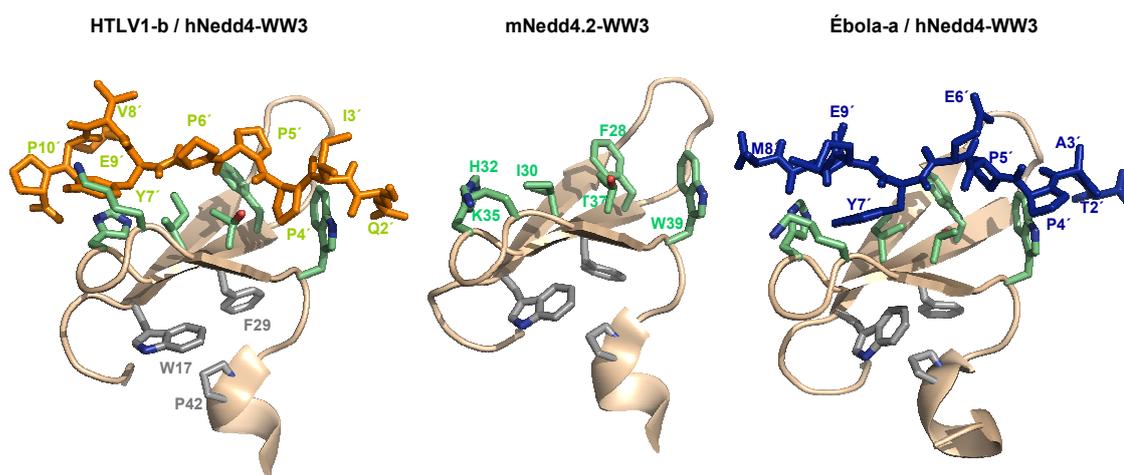


**Figura 3.10. Evaluación de las estructuras determinadas por RMN de los complejos de hNedd4-WW3.** En los diagramas de barras de los Paneles A) y B) se representan los valores de RMSD por residuo para las 20 estructuras de los complejos de hNedd4-WW3 con el ligando HTLV1-b y Ébola-a, respectivamente. Las barras en negro ilustran la desviación para la cadena principal de cada resto y las barras en gris la desviación correspondiente a cadena lateral. La representación esquemática con una gradación de colores de blanco a azul es una estimación estadística de la exposición al solvente para cada resto. El Panel C) muestra un ejemplo de la representación de Ramachandran para la estructura de menor energía del complejo del dominio hNedd4-WW3 con el ligando HTLV1-b (Lovell et al. 2003). Los aminoácidos se han dividido en cuatro categorías según los ángulos  $\Psi$  y  $\Phi$  permitidos estadísticamente (regiones delimitadas en azul): una general, otra para glicinas, otra para prolinas y la última para aquellos restos que preceden a las prolinas en la secuencia. El círculo rojo representa una prolina del ligando que se encuentra en una región generosamente permitida en este ejemplo, mientras que los aminoácidos en zonas más favorables se representan como círculos negros.

### 3.5. Descripción las estructuras de los complejos del dominio hNedd4-WW3

En la Figura 3.11 se muestran las estructuras de los complejos del dominio hNedd4-WW3 con los ligandos HTLV1-b y Ébola-a y la estructura del dominio homólogo en ratón, mNedd4.2-WW3, en su forma libre (código PDB 1WR7). La secuencia de mNedd4.2-WW3 es muy parecida a la secuencia de hNedd4-WW3, ya que sólo presenta 3 aminoácidos distintos que están localizados fuera del sitio

de unión. En las tres estructuras, los dominios WW presentan el plegamiento característico de estos módulos de reconocimiento, una lámina  $\beta$  torcida formada por tres hebras  $\beta$  antiparalelas. Dichas hebras  $\beta$  se encuentran estabilizadas por varios puentes de hidrógeno entre ellas y por el empaquetamiento (conocido como “broche hidrofóbico”) de las cadenas laterales del W17, de la F29 y de la P42 (Shaw et al. 2005). En la figura también se muestran los ligandos de cada complejo y los restos que conforman el sitio de unión típico de los dominios de clase I, es decir, las cadenas laterales de F28 y de W39 que forman el bolsillo xP; las cadenas laterales de I30, H32 y K35 que forman la cavidad xY y el resto T37 que separa ambos bolsillos. Salvando la mayor longitud de las hebras  $\beta$ 1 y  $\beta$ 2 en el complejo con el ligando **Ébola-a**, la disposición del esqueleto polipeptídico de los dominios WW en las tres estructuras es similar.



**Figura 3.11.** Representación en forma de cintas del complejo HTLV1-b/hNedd4-WW3 (a la izquierda), del dominio homólogo de ratón mNedd4.2-WW3 en estado libre (en el centro) y complejo Ébola-a/hNedd4-WW3 (a la derecha). En las tres estructuras, los restos responsables del plegamiento se muestran como varillas de color gris y los restos responsables de la interacción, como varillas verde claro. En los complejos, los ligandos se representan como varillas, naranjas, para el HTLV1-b, y azules, para el Ébola-a. El código PDB de la estructura del dominio mNedd4.2-WW3 es 1WR7.

Además, hemos realizado el alineamiento estructural teniendo en cuenta únicamente la lámina  $\beta$  (residuos 16-42) de las tres estructuras. De este modo, se ha calculado el valor de la RMSD entre los carbonos  $\alpha$ , comprobando que la desviación media de los carbonos  $\alpha$  es pequeña, unos 0.27 Å entre los dos complejos de hNedd4-WW3 y unos 0.20 Å para ambos complejos con el dominio

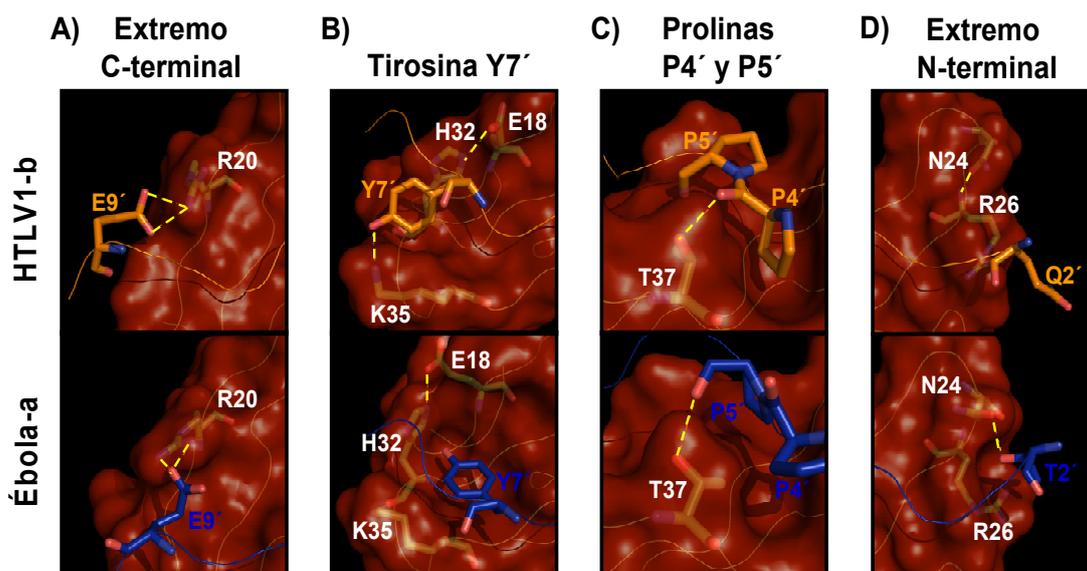
libre. Sin embargo, el valor de RMSD correspondiente a los carbonos  $\alpha$  de los aminoácidos en el giro entre las hebras  $\beta_1$  y  $\beta_2$  en los dos complejos se eleva a unos 0.76 Å. Los datos promediados para todas las estructuras no indican diferencias conformacionales importantes en el dominio hNedd4-WW3. En cambio, las desviaciones locales entre los dos complejos señalan diferencias en uno de los lazos del dominio que pueden tener relación con el modo de reconociendo de los ligandos. Además de las cadenas principales de las láminas  $\beta$  de las tres estructuras alineadas estructuralmente en la Figura 3.11 se han ilustrado varias cadenas laterales. De este modo, podemos comprobar que existen diferencias importantes en los complejos para la orientación de algunas de las cadenas laterales del sitio de unión (como son F28, I30, H30, K35 ó T37), así como para la orientación de los ligandos. Aunque ambos ligandos presentan una estructura tipo PPII en torno al motivo PPxY en los dos complejos, se puede apreciar que las cadenas laterales de estos restos no se encuentran en posiciones equivalentes en uno y otro complejo.

#### **3.5.1. Análisis del modo de unión de los dominios L víricos al dominio hNedd4-WW3**

En la Figura 3.12 se muestra una representación más detallada del modo de unión de los ligandos **HTLV1-b** y **Ébola-a** al dominio hNedd4-WW3. En esta figura, se han representado cuatro paneles ampliando distintas zonas de los ligandos en cada complejo: el extremo C-terminal, la tirosina Y7', las prolinas P4' y P5' del motivo de unión y el extremo N-terminal.

En el extremo C-terminal de ambos ligandos se produce una interacción de tipo electrostático entre el resto R20 de la primera hebra del dominio hNedd4-WW3 y los restos E9' presentes en ambos ligandos. En ambos casos se han observado señales NOE entre estos aminoácidos, pero en cada complejo se producen entre átomos diferentes (ver Tabla 3.3). Estas diferencias se ven reflejadas en la diferente disposición de la cadena lateral del resto E9' de cada ligando, como se puede apreciar en el Panel A de la Figura 3.12. Es importante mencionar que esta interacción electrostática puede ser uno de los determinantes de la especificidad de unión de los dominios L por el dominio hNedd4-WW3. En este sentido, la

mayoría de los dominios L víricos utilizados en el estudio (Tabla 3.1), presentan un resto ácido en la posición 9' y, además, el dominio hNedd4-WW3 es el único de los dominios WW estudiados que presenta un resto con carga positiva en la posición 20 o en otra próxima (Figura 3.1). De hecho, la eliminación de este resto cargado en el ligando **Ébola-b** disminuye drásticamente la afinidad de unión a hNedd4-WW3 con respecto a **Ébola-a**.



**Figura 3.12. Diferencias entre los complejos formados por hNedd4-WW3 con los ligandos HTLV1-b y Ébola-a.** **A)** Interacción del extremo C-terminal de ambos péptidos (resto E9'). **B)** Comparación de la tirosina Y7' de ambos ligandos en la cavidad xY de dominio hNedd4-WW3. **C)** Comparación de las prolinas P4' y P5' en la cavidad xP. **D)** Interacción en la región del extremo N-terminal (resto en la posición 2') de los dos ligandos con el dominio WW. La superficie del dominio WW se ha representado transparente y de color marrón dejando ver los restos implicados en la interacción como varillas de color verde. Los restos de los ligandos se muestran en representación de varillas, de color naranja, para el HTLV1-b, y de color azul, para el Ébola-a.

El bolsillo xY donde encaja la cadena lateral de la tirosina Y7' del ligando **HTLV1-b** parece más superficial que en el otro complejo. En el Panel B de la Figura 3.12, se puede observar cómo la cadena lateral de H32 está desplazada dejando más espacio para la Y7'. Además, se puede comprobar como la interacción del ligando **HTLV1-b** se ve estabilizada por un enlace de hidrógeno entre el grupo hidroxilo de Y7' y la cadena lateral de K35. La cadena lateral del resto Y7' del ligando **Ébola-a**, por el contrario, se encuentra más enterrada en el bolsillo xY. En este caso, la disposición de las cadenas laterales del dominio WW es mucho más parecida a la de la forma libre del dominio mNedd4.2-WW3 (en la

Figura 3.11). Estas características estructurales están en buena concordancia con las diferencias en la energética de unión observadas entre estos dos ligandos (Figura 3.4). Así, por ejemplo, aunqu al encontrarse el ligando **Ebola-a** más enterrado, la penalización entrópica desde el punto de vista de la entropía conformacional sería mayor, las contribuciones a la entropía de desolvatación de este ligando serían más importantes y más favorables.

La cavidad xP del dominio hNedd4-WW3 es más estrecha en el complejo con el ligando **Ébola-a**. Como se puede observar en el Panel C de la Figura 3.12, el anillo aromático de W39 se inclina ligeramente sobre la cavidad xP y ocupa parte del espacio necesario para que se entierren mejor las prolinas P4' y P5' del **Ébola-a**. En el complejo con el ligando **HTLV1-b**, el bolsillo xP es más abierto y las prolinas P4' y P5' de este ligando están en una posición menos superficial, provocando una modificación más acusada del desplazamiento químico del H<sub>ε1</sub> en el indol del W39 en la titulación de este ligando (Figura 3.7). Las diferencias en la geometría del sitio de unión del dominio hNedd4-WW3 además afectan a la formación del puente de hidrógeno entre el carbonilo de la P5' del ligando y el grupo hidroxilo de la T37, que es una de las características de las interacciones de los dominios WW (Ball et al. 2005). En el complejo formado con el ligando **Ébola-a**, este puente de hidrógeno podría verse comprometido debido a que la orientación y la distancia (unos 4 Å) entre los átomos implicados no son las más adecuadas. En el complejo con el ligando **HTLV1-b**, se produce un enlace de hidrógeno equivalente entre el grupo hidroxilo de T37 y el carbonilo de la P4' de este ligando. Sin embargo, en este caso, un ángulo menos forzado y una menor proximidad de los átomos (aproximadamente 3 Å) sugieren que el enlace es más estable. Así, si el complejo con el ligando **HTLV1-b** formase más puentes de hidrógeno (entre ellos, el formado por Y7' con K35 y este entre P4' y T37) que el complejo con el **Ébola-a**, esto conllevaría una contribución entálpica más favorable para este complejo y una contribución la entropía conformacional más desfavorable, situación que concuerda con los parámetros termodinámicos de unión determinados (Figura 3.4).

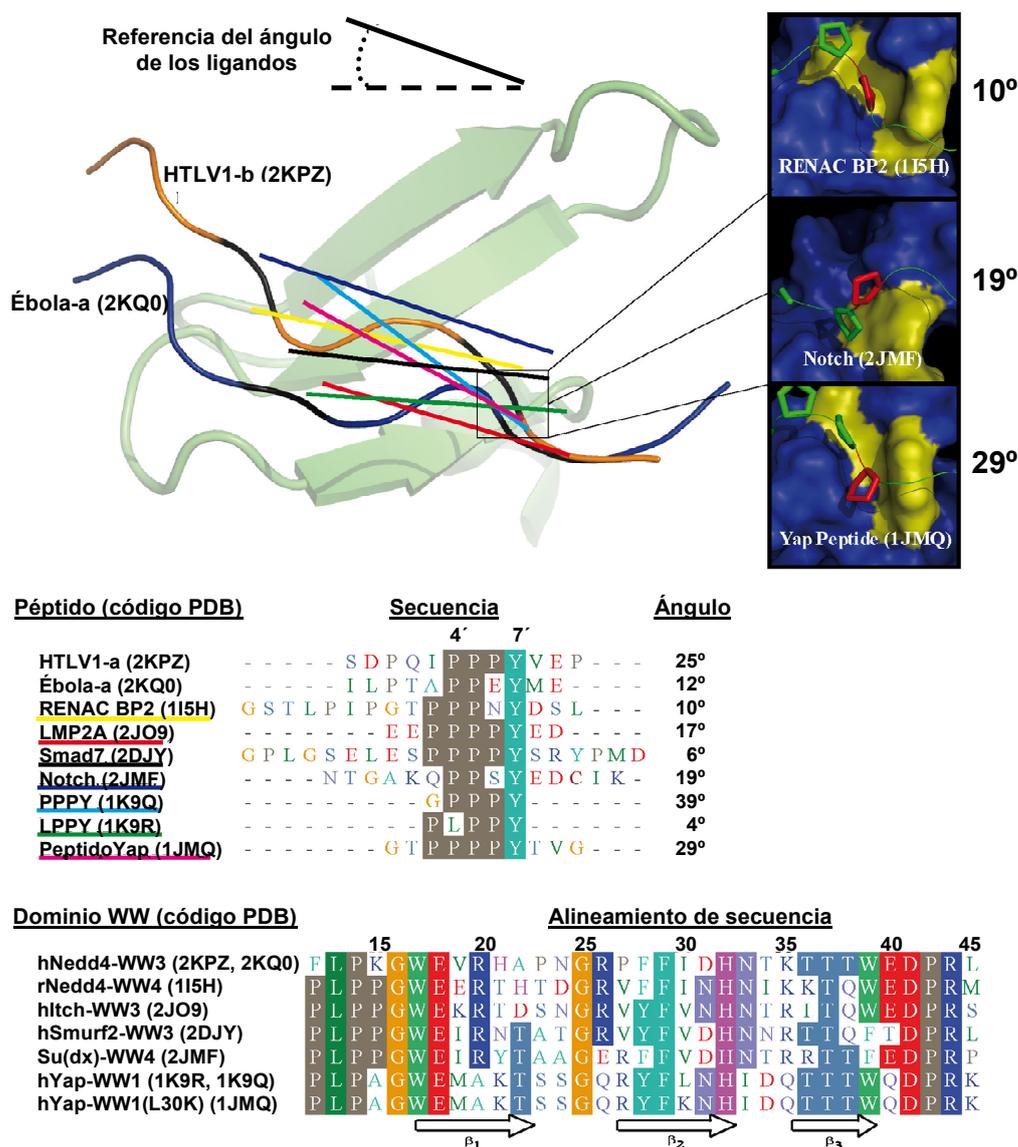
En el extremo N-terminal del complejo formado con el ligando **Ébola-a** se establece un puente de hidrógeno entre la cadena lateral del resto N24 del dominio

y el grupo hidroxilo de la cadena lateral de la T2' del ligando, como se aprecia en el Panel C de la Figura 3.12. En los extremos N-terminal, se han asignado señales NOEs de los restos I3' y Q2' de **HTLV1-b** y A3' y T2' de **Ébola-a** con el W39 del dominio (Tabla 3.3). Aunque estas señales NOEs son similares en los dos complejos, sólo en el caso del complejo con el **Ébola-a** el giro entre las hebras  $\beta 1$  y  $\beta 2$  adopta la conformación adecuada para producir esta interacción. Esta interacción adicional también se ha descrito para el dominio homólogo de mosca (Kanelis et al. 2006). Tanto en el dominio hNedd4-WW3 como en su homólogo de mosca, los lazos entre las hebras  $\beta 1$  y  $\beta 2$  presentan una secuencia A22, P23 y N24, que favorece que se produzcan los contactos adicionales con los ligandos. La P23 restringe las conformaciones posibles para N24, mientras que el resto A22 (cuya cadena lateral es muy pequeña) permite que el giro entre las hebras  $\beta 1$  y  $\beta 2$  adopte la conformación adecuada, en el caso del **Ébola-a**, para unir a la cadena lateral del resto T2' en una especie de cavidad adicional. A su vez, esta conformación del lazo de hNedd4-WW3 promueve la inclinación del grupo indol de W39 sobre la cavidad xP en el complejo que se forma con el ligando **Ébola-a**.

En resumen, el análisis detallado de las estructuras de los dos complejos del dominio hNedd4-WW3 con los ligandos víricos **HTLV1-b** y **Ébola-a** ha confirmado que estos ligandos son reconocidos de un modo diferente en cada caso, conforme como también sugiere la caracterización termodinámica de dichas interacciones. Además, la caracterización estructural ha puesto de manifiesto una interacción electrostática entre el extremo C-terminal de los dominios L víricos y una región próxima al sitio canónico de unión del dominio hNedd4-WW3, que parece ser responsable de la especificidad de unión entre ellos.

### **3.5.2. Comparación de los complejos formados por hNedd4-WW3 con otras estructuras**

La geometría del sitio de unión y el giro entre las hebras  $\beta 1$  y  $\beta 2$  del dominio hNedd4-WW3 son diferentes para la unión de los ligandos **HTLV1-b** y **Ébola-a**. Para saber si estos modos de unión son similares a los descritos para otros complejos de dominios WW, en la Figura 3.13 se comparan varios complejos de dominios WW de clase I. Se ha realizado un alineamiento estructural de los dominios WW de un



**Figura 3.13. Alineamiento estructural de varios complejos de dominios WW clasificados en el grupo I.** El alineamiento se ha realizado en base a las láminas  $\beta$  de los nueve complejos. A modo de ejemplo, se ha representado uno de los dominios WW en forma de cinta de color gris y un esquema de referencia de cómo se han determinado los ángulos para los ligandos. Los ligandos del HTLV1-b y Ébola-a también se presentan con forma de cintas en naranja y azul, respectivamente. El resto de ligandos se han representado como líneas rectas de distintos colores que describen el eje longitudinal de cada ligando con estructura PPII. Las líneas rectas van desde el carbono  $\alpha$  del resto P4' ó L4', hasta el carbono  $\alpha$  de la tirosina Y7' de cada ligando y son: amarillo para el ligando RENAC BP2 (PDB 115H), rojo para LMP2A (PDB 2JO9), negro para Smad7 (PDB 2DJY), azul marino para Notch (PDB 2JMF), turquesa para PPPY (PDB 1K9R), marrón para LPPY (PDB 1K9R) y púrpura para el péptido Yap (PDB 1JMQ). En los recuadros de la derecha se han ilustrado varios detalles de la geometría de los bolsillo xP en los complejos RENEAC BP2/rNedd4-WW4, Notch/Su(dx)-WW4 y péptido-Yap/hYap-WW1. En el centro, se muestran los alineamientos de secuencias de los ligandos y el ángulo que forman en los complejos. En la parte inferior, se presenta el alineamiento de secuencias de los dominios WW, indicando los correspondientes elementos de estructura secundaria.

conjunto de nueve complejos depositados en la base de datos PDB. En la figura se muestra un sólo dominio WW como ejemplo de dicho alineamiento y dos tipos de representaciones diferentes para los ligandos. Por un lado, se muestra una representación en forma de cintas para los ligandos **HTLV1-b** (naranja) y **Ébola-a** (azul), y por otro lado, se muestran varias líneas rectas correspondientes a los ejes longitudinales de cada ligando entre los enlaces peptídicos de las posiciones 4' (prolina o leucina) y 7' (tirosina). Además, en la figura también se han incluido los alineamientos de secuencias de los dominios WW y de sus ligandos.

El alineamiento estructural de los dominios WW no refleja importantes desviaciones de sus carbonos  $\alpha$  (el valor RMSD es inferior a 1Å). Sin embargo, en la Figura 3.13, se aprecia claramente que cada ligando presenta una orientación diferente sobre la superficie del dominio WW. Para ilustrar las diferentes orientaciones de los ligandos se ha calculado un ángulo de referencia para cada ligando como se indica en la propia figura. Los valores de estos ángulos varían para los distintos complejos, de modo que los ligandos se pueden distribuir en dos grupos según su modo de unión: ligandos similares al **Ébola-a** con ángulos del orden de los 12° (o inferiores) y ligandos similares a **HTLV1-b** con ángulos de valores igual a 20° o superiores.

En el detalle de la Figura 3.13 para el bolsillo xP del complejo RENAC-BP2/rNedd4-WW4, caracterizado por un ángulo de 10°, se puede apreciar que los complejos con ángulos bajos para la orientación de sus ligandos suelen presentar una cavidad xP estrecha, como también se ha descrito para el complejo de hNedd4-WW3 con el **Ébola-a** (Panel C, Figura 3.12). Por otro lado, los ligandos que presentan ángulos iguales o mayores 20° interaccionan en una cavidad xP mucho más abierta, como se muestra en el detalle de la Figura 3.13 para el péptido de Yap (29°) o como se ha mostrado para el complejo de hNedd4-WW3 con el ligando **HTLV1-b** (Panel C, Figura 3.12). Otro caso son aquellos complejos caracterizados por un bolsillo xP con un resto fenilalanina en la posición 39 en lugar de un triptófano; estos dominio WW presentan un bolsillo menos profundo por lo que la disposición del ligando depende de los contactos de éste con otras zonas del dominio. Los dos ejemplos de este tipo en la Figura 3.13 presentan

ángulos muy distintos: 6° para el complejo Smad7/hSmurf2-WW3 y 19° para el complejo Notch/Su(dx)-WW4.

Nuestros resultados para el reconocimiento de dos ligandos víricos con pequeñas diferencias en sus secuencias (**HTLV1-b** y **Ébola-a**) por el dominio hNedd4-WW3 han puesto de manifiesto que el patrón termodinámico puede ser muy diferente y que el dominio tiene la capacidad de adoptar pequeños cambios de conformación para facilitar el reconocimiento de los ligandos. También hemos comprobado que dicha variabilidad en el modo de reconocer estos ligandos está de acuerdo con las conformaciones adoptadas por otros dominios WW de clase I. Esta adaptabilidad no sólo se produce en las regiones más flexibles, como son los giros entre las hebras, sino que también se extiende a las cadenas laterales que conforman el sitio de unión. En este contexto, la elevada plasticidad y flexibilidad conformacional de estos dominios, puesta de manifiesto en los estudios de estabilidad (Capítulo 2), adquieren una clara relevancia funcional.

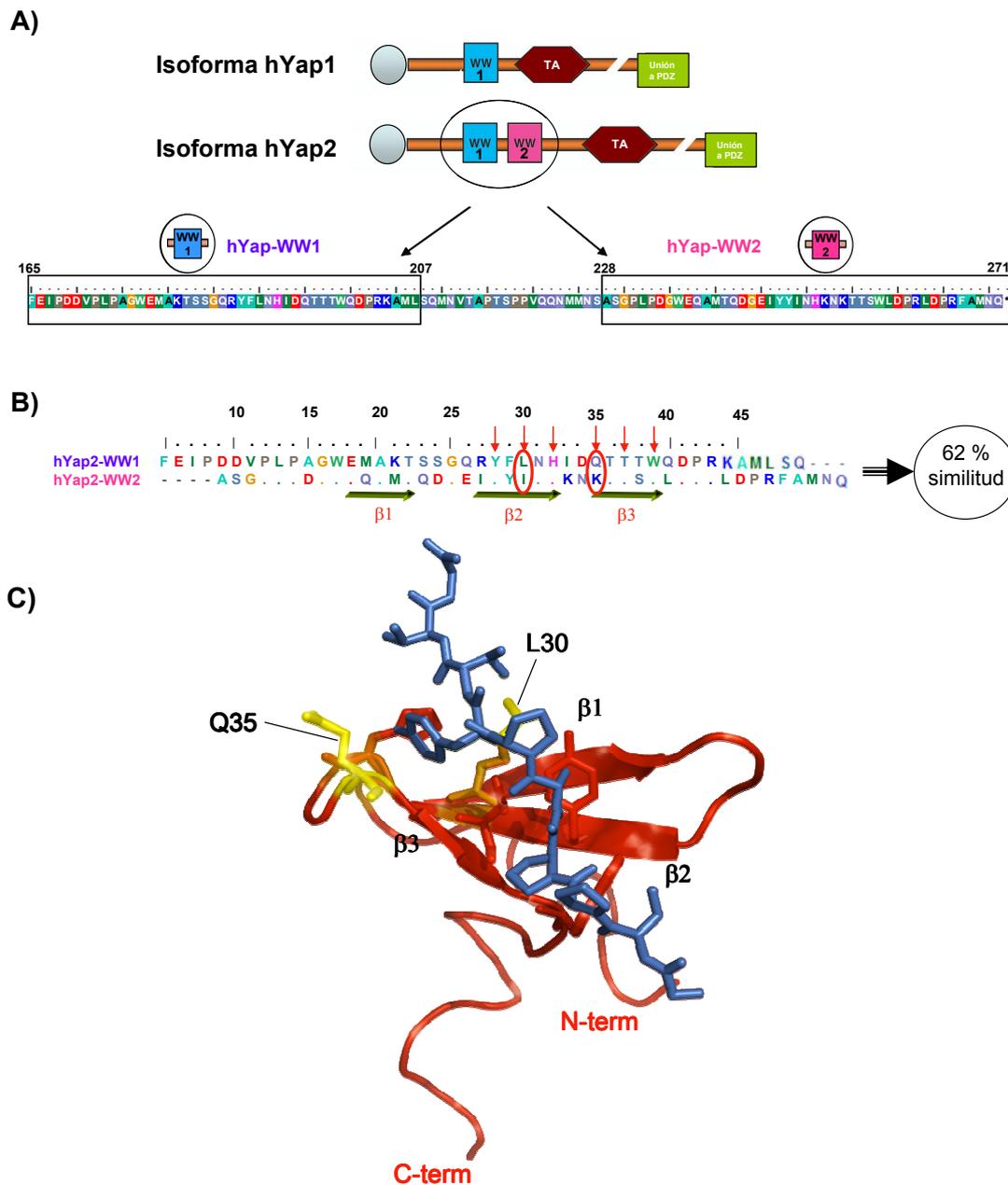
## CAPÍTULO 4



## **CAPÍTULO 4. Reconocimiento molecular mediado por el tándem de dominios WW de la proteína humana hYap**

### **4.1. Introducción**

Como hemos visto anteriormente, existen dos isoformas mayoritarias de hYap (Figura 4.1A). La forma hYap1 sólo presenta un dominio WW (hYap-WW1) y la forma hYap2 contiene dos dominios WW (el tándem hYap-WW1-WW2). Por otro lado, se ha demostrado que la isoforma hYap2 tiene una mayor actividad transcritora de genes que la isoforma hYap1 (Komuro et al. 2003). Estas dos características sugieren que la presencia de un dominio WW adicional, el dominio hYap-WW2, tiene un papel fundamental en la función de la isoforma hYap2. En el Panel B de la Figura 4.1 se ha ilustrado el alineamiento de las secuencias de ambos dominios WW en el que se observa que la similitud de secuencia es del 62 % aproximadamente. El dominio hYap-WW1 une a ligandos que contienen el motivo PPxY (Koepf et al. 1999b; Pires et al. 2001) y ha sido clasificado en el grupo I de especificidad. Para el dominio hYap-WW2 aún no se ha descrito ningún ligando, pero la elevada similitud con hYap-WW1 sugiere que también pertenece al grupo I. Los restos más característicos en el sitio de unión de esta clase de especificidad están perfectamente conservados en ambos dominios (P14, W17, Y28 ó F28 y P42) (Macias et al. 2000). No obstante, encontramos dos diferencias importantes en las posiciones 30 y 35 correspondientes al bolsillo xY (representadas en amarillo en el Panel C de la Figura 4.1). El aminoácido en la posición 30 de ambos dominios (una leucina en hYap-WW1 y una isoleucina en hYap-WW2) tiene propiedades similares. En cambio, la diferencia en posición 35 no es conservativa, ya que el dominio hYap-WW2 presenta una carga positiva (lisina) donde hay un resto neutro (glutamina) en hYAP-WW1.



**Figura 4.1. Representación esquemática de la isoforma hYap2 y de sus dominios WW.** **A)** Arriba, la representación de los dominios en que se configura la arquitectura de las isoformas de hYap. Abajo, la secuencia del tandem de dominios WW en la que los recuadros negros delimitan cada uno de los dominios individuales. **B)** Alineamiento de secuencias de ambos dominios WW individuales de Yap. Los puntos de colores en la secuencia de hYap-WW2 corresponden a aquellos restos que coinciden en idéntica posición con los del dominio hYap-WW1. **C)** Estructura resuelta por RMN del mutante L30K de hYap-WW1 formando complejo con el péptido GPPPPYTVG (PDB: 1JMQ). El esqueleto polipeptídico del dominio WW se ha representado como cintas en color rojo y el péptido como varillas de color turquesa. Las cadenas laterales de los restos implicados en el sitio de unión que son idénticas entre los dominios WW de hYap se han representado como varillas rojas y las que son diferentes, Q35 y L30, se han representado como varillas amarillas.

Hay dos posibles explicaciones para las diferencias funcionales observadas entre las isoformas hYap1 y hYap2: bien que el segundo dominio WW confiera una especificidad de unión diferente a hYap2, o bien que dicho dominio WW extra sea el responsable de efectos contextuales similares a los descritos en la bibliografía para otras parejas de dominios WW, como efectos sinérgicos en la estabilidad conformacional o en el reconocimiento de los ligandos (Wiesner et al. 2002; Fedoroff et al. 2004; Jennings et al. 2007; Huang et al. 2009; Chong et al. 2010). En resumen, las principales preguntas que planteamos son: ¿Por qué existen dos isoformas de hYap? ¿Qué implicaciones funcionales tiene la existencia de los dos dominios WW en hYap2? Con el fin de dar respuesta a estas cuestiones, hemos abordado, por una parte, la caracterización termodinámica del reconocimiento de los ligandos naturales derivados de las proteínas Lats1, Lats2 y Ptch1 y, por otra, del equilibrio conformacional de los dominios WW individuales (hYap-WW1 y hYap-WW2), así como del tándem hYap-WW1-WW2.

La diversidad de funciones determinantes para la vida de la célula atribuidas a hYap (proteína humana asociada a Yes) ha puesto de manifiesto la importancia que tiene comprender los fundamentos de su regulación. Actualmente la regulación de hYap se asocia a su retención en el citosol, mayoritariamente mediada por su interacción con la proteína 14-3-3. Cuando hYap está unida a 14-3-3 no entra en el núcleo y no se produce la expresión de genes. Se ha comprobado que esta interacción es dependiente de la fosforilación de la serina 127 de hYap por medio de quinasas de serina/treonina, y sólo cuando se produce esta fosforilación, hYap une a 14-3-3 (Zhao et al. 2007; Hao et al. 2008). Por otra parte, como se ha descrito en el capítulo de Introducción, la fosforilación de otras serinas de hYap se ha propuesto como mecanismo de modulación de algunas de sus funciones (Matallanas et al. 2007). La interacción de los dominios WW de hYap con la quinasa de serina/treonina Lats1 (Oka et al. 2008) parece ser uno de los factores determinantes en la fosforilación de las diferentes serinas. Otros posibles factores pueden ser las interacciones de los dominios WW de hYap con otras proteínas como la quinasa de serina/treonina Lats2 o la proteína citosólica Ptch1. En el caso de la proteína Lats2 se ha propuesto que actúa de forma análoga a Lats1 por tratarse de una proteína homóloga (Hao et al. 2008; Alderton

2010). En cambio, la unión de Ptch1 con hYap parece ser similar a la interacción de hYap con la proteína 14-3-3, pues impide que hYap entre en el núcleo y, por tanto, también regula negativamente la expresión de genes (Alderton 2010). Aunque la interacción de Ptch1 con alguno de los dominios WW de hYap no ha sido descrita, el grupo de investigación del Dr. Marius Sudol (Escuela Monte Sinaí de Medicina, Nueva York, USA) ha encontrado evidencias de una modulación recíproca de estas dos proteínas. En la Tabla 4.1 se han recogido cinco secuencias ricas en prolina pertenecientes a las proteínas Lats1, Lats2 y Ptch1, sugeridas por el Dr. Marius Sudol como posibles ligandos de los dominios WW de hYap en base al análisis de las secuencias de las proteínas completas, aunque no han sido confirmadas experimentalmente como ligandos reales de estos dominios. En Lats2 se ha escogido una secuencia, mientras que en Lats1 y Ptch1 se han elegido dos posibles ligandos. En ninguna de estas dos proteínas, Lats1 ó Ptch1, se han podido considerar ligandos más largos que contengan los dos motivos PPxY por estar muy separados en sus secuencias.

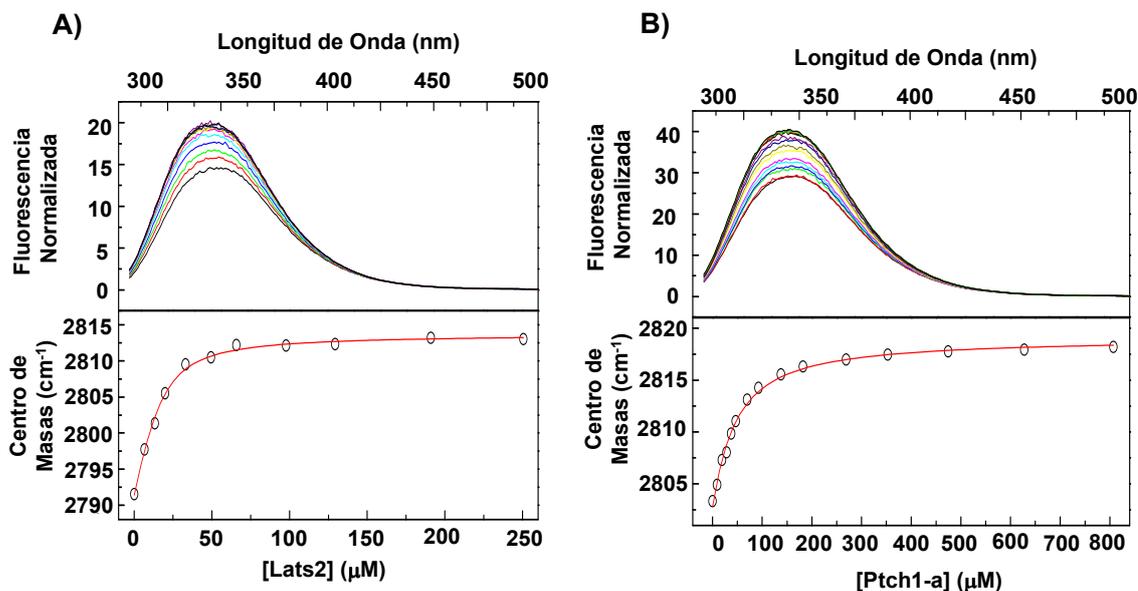
**Tabla 4.1. Selección de los ligandos derivados de varias proteínas naturales.** Todos los ligandos se han utilizado para la caracterización termodinámica de las interacciones de los dominios WW de la proteína hYap.

Ruta de señalización "Hippo"	Ligando	Secuencia*
<b>Lats1</b> Quinasa de treonina/serina "Large Tumor Suppressor Homolog 1"	<b>Lats1-a</b>	<sup>369</sup> NRQP <b>PPPY</b> PLTA <sup>371</sup>
	<b>Lats1-b</b>	<sup>552</sup> YQGP <b>PPPY</b> PKHL <sup>564</sup>
<b>Lats2</b> Quinasa de treonina/serina "Large Tumor Suppressor Homolog 2"	<b>Lats2</b>	<sup>511</sup> RRC <b>PPPY</b> PKHL <sup>523</sup>
Ruta de señalización "Sonic Hedgehog"		
<b>Ptch1</b> Receptor glicosilado de membrana "Protein Patched Homolog 1"	<b>Ptch1-a</b>	<sup>572</sup> RYSP <b>PPPY</b> SSH <sup>584</sup>
	<b>Ptch1-b</b>	<sup>1243</sup> EGLW <b>PPPY</b> RPRR <sup>1255</sup>

\* Para identificar más fácilmente la posición del motivo PPxY en la secuencia de los ligandos éste se muestra en azul y el superíndice indica la posición de estos ligandos dentro de las secuencias de las proteínas naturales completas

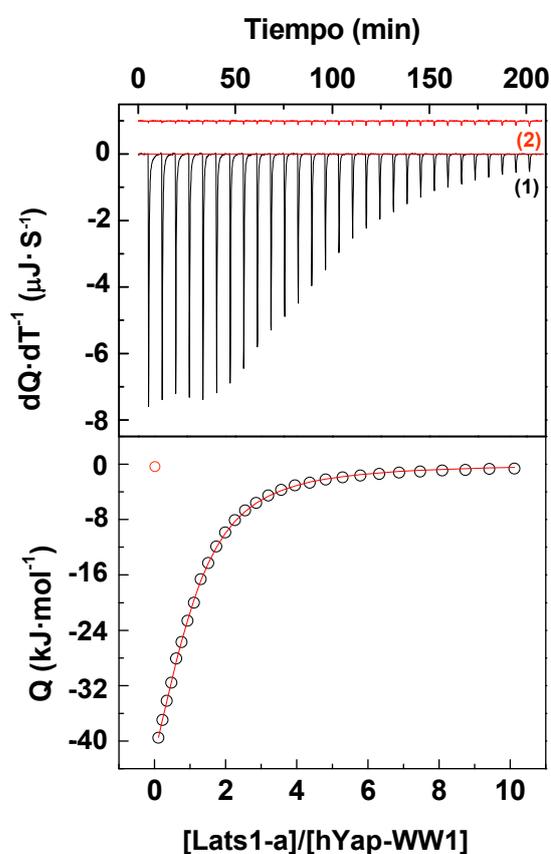
#### 4.2. Caracterización termodinámica de la unión de ligandos naturales a los dominios WW de hYap

Dado que los ligandos seleccionados aún no han sido confirmados como tales, en primer lugar hemos comprobado si los ligandos derivados de las proteínas Lats1, Lats2 y Ptch1 son reconocidos por los dominios WW de hYap. Con este fin, se ha realizado un análisis de la unión de estos ligandos a cada dominio WW aislado de hYap mediante espectroscopía de fluorescencia. Las únicas excepciones han sido las titulaciones del ligando **Ptch1-b**, que no han podido seguirse mediante esta técnica debido a la presencia de un triptófano en la secuencia del ligando. La cadena lateral de este resto de **Ptch1-b** impide distinguir la amortiguación de la fluorescencia de los dominios WW durante la titulación con este ligando. En la Figura 4.2 se muestran, a modo de ejemplo, las titulaciones del dominio hYap-WW1 con el ligando **Lats2** y el del dominio hYap-WW2 con el ligando **Ptch1-a**, ambos a 25 °C y en tampón fosfato sódico 20 mM pH 7.0.



**Figura 4.2.** Experimentos de titulación de los dominios WW de la proteína hYap con algunos de los ligandos naturales seguidos mediante espectroscopía de fluorescencia. **A)** Adición del ligando **Lats2** sobre el dominio hYap-WW1. **B)** Titulación de hYap-WW2 con el ligando **Ptch1-a**. En los paneles superiores, se han representado los espectros de fluorescencia tomados tras cada inyección de ligando a 25 °C y en fosfato sódico 20 mM pH 7.0. Todos los espectros se han normalizado por la concentración de proteína. En los paneles inferiores, se han representado los valores del centro de masas (círculos abiertos) para cada uno de los espectros del panel superior. Las líneas continuas corresponden a los respectivos ajustes según el modelo de unión seguido mediante fluorescencia (ecuación 6.5.3, Apartado 6.5.4).

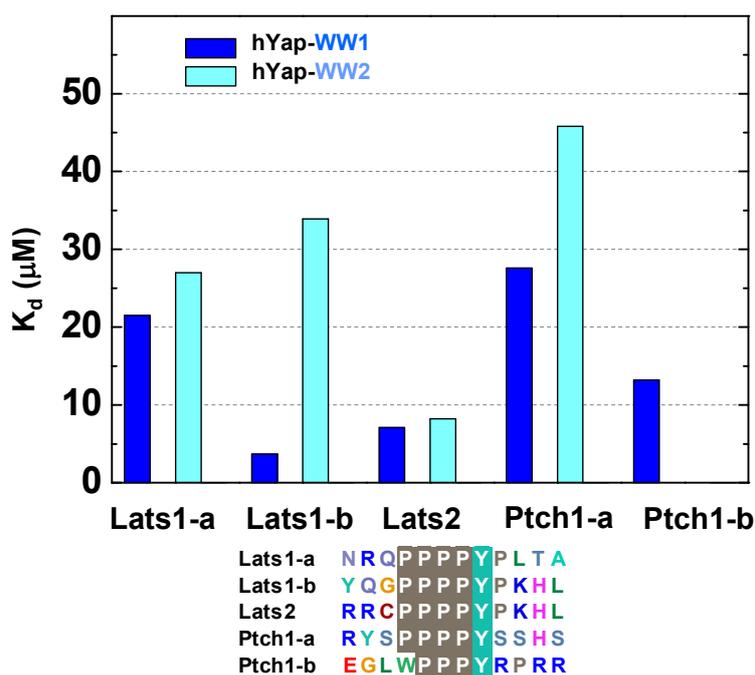
Los experimentos de fluorescencia han confirmado que ambos dominios WW de la proteína hYap interaccionan con los ligandos **Lats1-a**, **Lats1-b**, **Lats2**, y **Ptch1-a**. Además, estos experimentos nos han permitido comprobar que la magnitud de las constantes de disociación es adecuada para llevar a cabo titulaciones directas mediante ITC. De este modo, estas constantes se han utilizado para el diseño de las titulaciones calorimétricas en idénticas condiciones. En el panel superior de la Figura 4.3, se muestra el experimento de titulación del dominio hYap-WW1 con el ligando **Lats1-a** utilizando un perfil de volúmenes de



**Figura 4.3. Ejemplo de un experimento de titulación del dominio hYap-WW1 mediante ITC.** En el panel superior, se presenta el termograma de la titulación a 25 °C en fosfato sódico 20 mM, pH 7.0 con el ligando **Lats1-a** (1), en el que se muestran los calores por unidad de tiempo liberados tras la inyección de ligando sobre una disolución de proteína, y el termograma de dilución del ligando (2) en las mismas condiciones sobre el tampón (en rojo y desplazado en el eje para mayor claridad). En el panel inferior, se observa la isoterma de unión correspondiente a la normalización de los calores derivados de cada inyección, corregidos teniendo en cuenta los calores debidos a la dilución del ligando. Los datos experimentales se representan como círculos, mientras que la línea continua corresponde al mejor ajuste según el modelo de unión de un ligando a una macromolécula con “n” sitios de unión idénticos e independientes (ecuación III.X del Apéndice III).

inyección creciente (entre 4 y 22  $\mu\text{L}$ ) para definir mejor la isoterma de unión, que se muestra en el panel inferior junto con el mejor ajuste al modelo de unión a sitios idénticos e independientes, de acuerdo con la ecuación III.X del Apéndice III.

Los valores de las constantes de disociación resultantes del análisis de estos experimentos de ITC se han representado como un diagrama de barras en la Figura 4.4.



**Figura 4.4. Afinidad de unión de los diferentes ligandos a los dominios individuales de hYap.** Representación en forma de barras de las constantes de disociación para la unión de los ligandos **Lats1-a**, **Lats1-b**, **Lats2**, **Ptch1-a** y **Ptch1-b** a los dominios hYap-WW1 y hYap-WW2. En color azul oscuro se han representado las constantes de disociación para el dominio hYap-WW1 y en color azul claro para el dominio hYap-WW2. Todas las constantes se han obtenido mediante distintos experimentos de ITC en tampón fosfato sódico 20 mM pH 7.0 a 25 °C (en presencia de  $\beta$ -mercaptoetanol 20 mM en aquellos experimentos con el ligando **Lats2**). Los errores asociados a las constantes de disociación se estiman en torno al 10 %. Bajo el diagrama, se ha representado un alineamiento de las secuencias de los ligandos empleados en el estudio.

Estas constantes de disociación confirman que todas las secuencias seleccionadas, incluida **Ptch1-b**, son ligandos de al menos uno de los dominios WW de hYap. Además, todas las constantes son del mismo orden de magnitud descrito para otros dominios WW y sus ligandos naturales (Pires et al. 2001; Kanelis et al. 2006; Meiyappan et al. 2007; Morales et al. 2007; Ramirez-Espain et al. 2007). La afinidad de los distintos ligandos por cada dominio WW no es siempre

exactamente igual, a veces hay diferencias. Por ejemplo, las constantes de disociación de **Lats2** y de **Ptch1-a** por el dominio hYap-WW2 son 8.2  $\mu\text{M}$  y 45.8  $\mu\text{M}$ , respectivamente.

Al menos uno de los ligandos de las proteínas Lats1 y Ptch1, en concreto **Lats1-b** y **Ptch1-b**, muestran una afinidad de unión bastante elevada por el dominio hYap-WW1 ( $K_d < 15 \mu\text{M}$ ), lo que garantiza la interacción en el contexto de las proteínas completas. Sin embargo, los otros dos ligandos, **Lats1-a** y **Ptch1-a**, unen al dominio hYap-WW1 con afinidades menores, lo cual sugiere la existencia de una clara preferencia de unión.

En cuanto a la especificidad de unión entre los dos dominios hYap-WW1 y hYap-WW2, podemos afirmar que para algunos ligandos estudiados existen diferencias en la afinidad de unión a uno u otro dominio. Los ligandos **Lats1-a** y **Lats2** presentan afinidades muy parecidas por ambos dominios WW de hYap. Pero los ligandos **Ptch1-a**, **Ptch1-b** y **Lats1-b** muestran una mayor afinidad de unión por el dominio hYap-WW1, como podemos comprobar en la Figura 4.4. El ejemplo más importante de la especificidad de unión, es el ligando **Ptch1-b** que, presenta una de las constantes de disociación más pequeñas para el dominio hYap-WW1, mientras que su interacción con el segundo dominio WW es varios ordenes de magnitud más débil, tanto que no se ha podido analizar correctamente ningún experimento de unión. En el caso de **Ptch1-a** y **Lats1-b**, las diferencias de afinidad por uno u otro dominio WW en términos de energía libre de Gibbs son pequeñas pero detectables, están entorno a 5  $\text{kJ}\cdot\text{mol}^{-1}$ . Estos resultados sugieren que uno de los mecanismos mediante los cuales la presencia de un segundo dominio WW en la isoforma 2 de hYap es la introducción de nuevas especificidades de unión que pudieran contribuir a modular su actividad.

Además, si nos fijamos en las secuencias de los ligandos atendiendo a la preferencia de unión por los dominios WW de hYap (Figura 4.4), encontramos algunas características comunes. Así, por ejemplo, los dos ligandos con menor afinidad por hYap-WW2, **Ptch1-b** y **Ptch1-a**, no tienen el motivo de unión PPxY flanqueado por restos prolina como sucede con los demás ligandos, mientras que los ligandos **Lats1-a** y **Lats2**, que no distinguen entre los dos dominios WW, contienen un resto arginina en la segunda posición (R2) de sus secuencias.

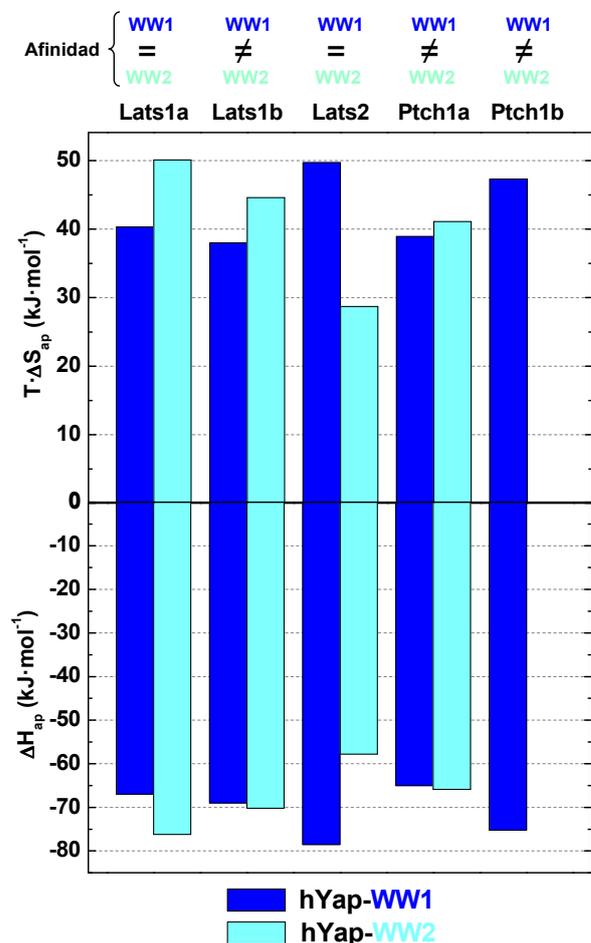
En resumen, nuestro estudio confirma que las cinco secuencias seleccionadas son ligandos de los dominios WW de hYap. La mayoría muestran una mayor afinidad por el dominio hYap-WW1. También se ha puesto de manifiesto que pequeñas diferencias en sus secuencias, inducen importantes variaciones en la afinidad por el mismo dominio WW. Además, estos resultados indican que la presencia del segundo dominio hYap-WW2 modifica la especificidad de la isoforma hYap2 con respecto a hYap1, ya que, al presentar dos dominios WW con capacidad de reconocer secuencias diferentes, presenta un espectro mayor de posibles proteínas con las que interaccionar.

#### **4.2.1. Disección de las contribuciones termodinámicas a la afinidad de unión**

El estudio de las interacciones entre los dominios WW de hYap y varios ligandos naturales revela una mayor preferencia de algunos de los ligandos por el dominio hYap-WW1 (Figura 4.4). Esta situación puede tener relación con las diferencias observadas en la actividad de las dos isoformas de hYap. Los dos dominios WW de hYap presentan una elevada similitud de secuencia con muy pocas diferencias en el sitio de unión (Figura 4.1). Como se puede observar en la Tabla 4.1, los ligandos también son parecidos, ya que todos presentan el mismo motivo PPPY centrado y bastantes restos básicos en los extremos de sus secuencias. La caracterización calorimétrica del equilibrio de unión de estos ligandos nos proporciona información sobre la naturaleza de las fuerzas que dirigen la interacción y determinan la especificidad de unión.

En la Figura 4.5 se han representado las magnitudes termodinámicas que gobiernan cada una de las interacciones (cambios en entalpía y entropía aparentes) como un diagrama de barras. Como se puede observar, y como se ha descrito para otros dominios WW (Koepp et al. 1999b; Pires et al. 2001; Meiyappan et al. 2007; Morales et al. 2007; Ramirez-Espain et al. 2007; Huang et al. 2009), la unión de ambos dominios WW de hYap con los ligandos naturales **Lats1-a**, **Lats1-b**, **Lats2**, **Ptch1-a** y **Ptch1-b** es un proceso cuyo patrón termodinámico se caracteriza por valores de entalpía favorables, superiores a  $-60 \text{ kJ}\cdot\text{mol}^{-1}$ , parcialmente compensados por una contribución entrópica desfavorable entre los  $+30$  y  $+50 \text{ kJ}\cdot\text{mol}^{-1}$ . Estas fuerzas presentan el mismo signo y una magnitud del

mismo orden que las obtenidas para la unión de ligandos de origen vírico al tercer dominio WW de hNedd4 en el Capítulo 3 (Figura 3.7).



**Figura 4.5. Disección de la energética de unión de los ligandos Lats1-a, Lats1-b, Lats2, Ptch1-a y Ptch1-b a los dominios WW de hYap.** En la parte superior, se muestra de forma esquemática si la afinidad por el ligando es igual para los dos dominios. En el medio, se representan los términos entrópicos ( $T \cdot \Delta S_{ap}$ ) y, en la parte inferior, se representan los términos entálpicos ( $\Delta H_{ap}$ ) correspondientes a cada interacción. Ambos términos se han ilustrado como barras de color azul oscuro y de color azul claro para el dominio hYap-WW1 y para el dominio hYap-WW2, respectivamente. Todos los parámetros termodinámicos de unión se han obtenido a 25 °C mediante experimentos de ITC en fosfato sódico 20 mM pH 7.0 y conllevan un error estimado en un 5%.

Los ligandos que presentan afinidades muy parecidas por los dos dominios de hYap (Figura 4.4), **Lats1-a** y **Lats2**, muestran valores muy diferentes para los términos entálpico y entrópico dependiendo del dominio WW al que estén unidos (Figura 4.5). En ambos casos, los parámetros termodinámicos de la unión de cada ligando con los dos dominios se compensan y el resultado final es una afinidad muy parecida por los dos dominios WW. Al igual que se ha discutido en relación a

las diferencias en la energética de unión de algunos ligandos de origen vírico (como **HTLV1-b** y **Ébola-a**) con el dominio hNedd4-WW3 (Apartado 3.2.2 del Capítulo 3), las diferencias en las contribuciones energéticas entre los dominios WW de hYap ponen de manifiesto que el modo de reconocimiento de los ligandos **Lats1-a** ó **Lats2** es diferente en cada caso, de hecho, sugieren que las pequeñas diferencias en el sitio de unión entre los dos dominios WW de hYap (Figura 3.1) podrían ser responsables de que se produzcan distintas interacciones en la formación de complejos con hYap-WW1 o hYap-WW2.

Los dos ligandos **Lats1-b** y **Ptch1-a**, que han mostrado cierta preferencia de unión por el dominio hYap-WW1, presentan unos cambios de entalpía de unión a uno u otro dominio muy parecidos mientras que los térmicos entrópicos son algo diferentes, ligeramente más favorables para la unión de los dos ligandos con hYap-WW1 (Figura 4.5).

El resultado más importante en cuanto a especificidad de unión es el del ligando **Ptch1-b** que sólo interacciona con dominio hYap-WW1 y, además, lo hace con uno de los cambios de entalpía de unión más elevados (Figura 4.5). La especificidad y los valores de entalpía de unión lo que sugieren es que las interacciones se hallan bastante optimizadas en este complejo. En su secuencia encontramos algunas características que lo diferencian de los otros ligandos: tres restos básicos casi consecutivos en el extremo C-terminal (**EGLWPPPYRPRR**) y un resto triptófano precediendo al motivo PPxY.

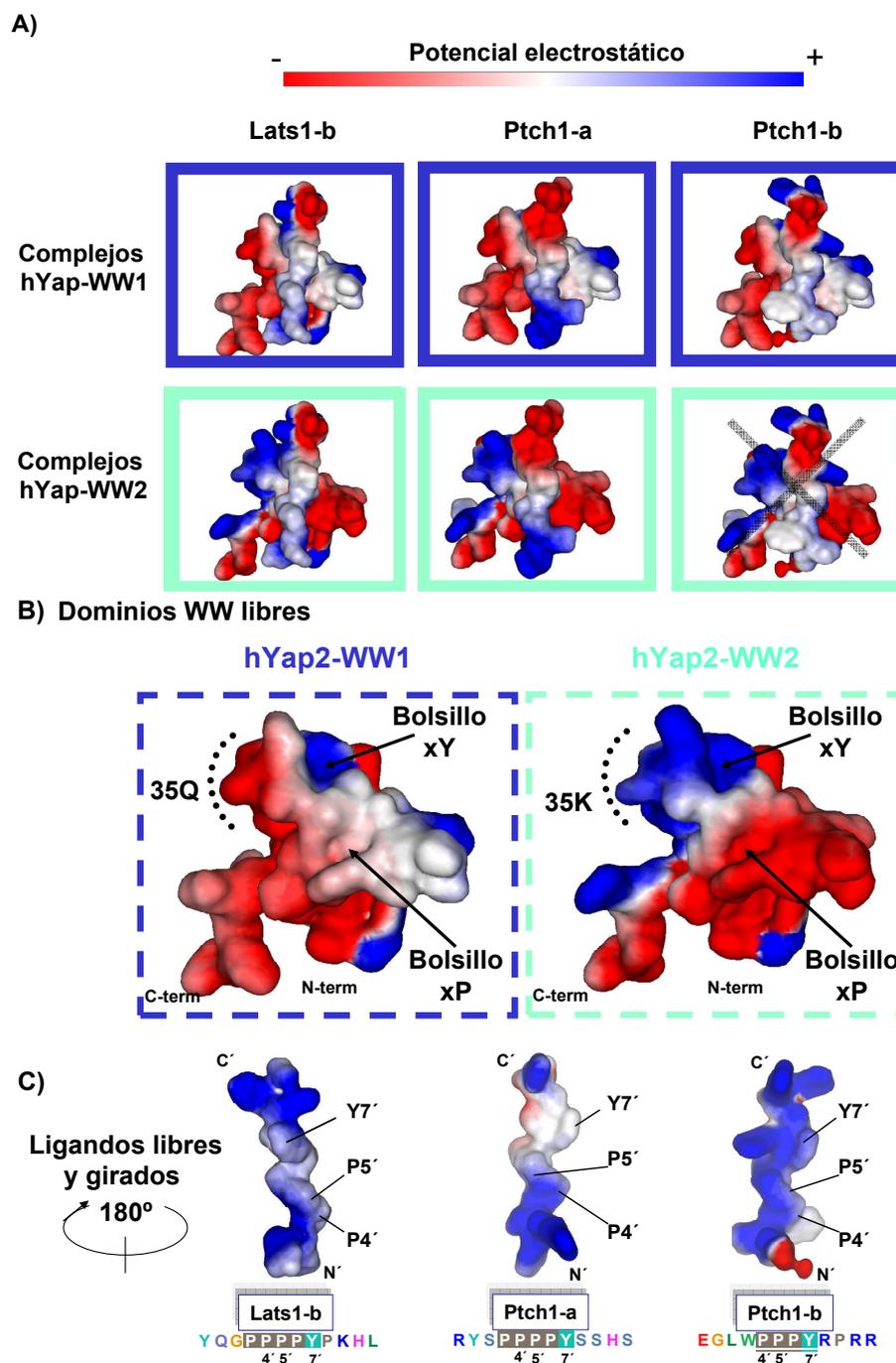
En definitiva, el estudio comparativo de la termodinámica de unión de un conjunto de ligandos a los dos dominios WW de hYap revela importantes diferencias en la forma en que los ligandos son reconocidos por uno y otro dominio. Estos resultados claramente señalan que las diferencias en el sitio de unión entre los dos dominios WW, son suficientes para conferir una especificidad diferente a ambos dominios.

#### **4.2.2. Potencial electrostático y especificidad de unión**

Como hemos comentado, la diferencia de afinidad entre los dominios WW por **Lats1-b** y **Ptch1-a** tiene su origen en una contribución entrópica más favorable. Teniendo en cuenta que la principal diferencia en el sitio de unión de

ambos dominios implica la sustitución de un resto no cargado en hYap-WW1 por una carga positiva en hYap-WW2 (Figura 4.1) y que los ligandos que presentan un elevado número de restos cargados (Tabla 4.1), es nuestra hipótesis que las diferencias entrópicas podrían estar relacionadas con interacciones de tipo electrostático. Para investigar esta posibilidad, hemos calculado el potencial electrostático para un conjunto de modelos estructurales de los complejos formados por ambos dominios WW y dichos ligandos con diferentes especificidades de unión. Los modelos estructurales se han generado mediante sustitución de las cadenas laterales (manteniendo las coordenadas del esqueleto polipeptídico) a partir de uno de los complejos más parecidos que hay disponibles en el PDB. Este complejo se ha determinado por RMN y está formado por el mutante L30K de hYap-WW1 y el ligando GTPPPYTVG (código PDB 1JMQ). Los cálculos del potencial electrostático se han realizado para cada complejo, para cada ligando y para cada dominio WW conforme se detalla en el Capítulo 6 de Materiales y Métodos (Apartado 6.8).

En el Panel A de la Figura 4.6, se muestra el potencial electrostático superficial calculado para los modelos de los complejos entre los dos dominios WW de hYap y los ligandos **Lats1-b**, **Ptch1-a** y **Ptch1-b**. En el Panel B de la Figura 4.6 se muestran los potenciales electrostáticos de los dominios WW de hYap libres. En esta representación podemos observar que, a pesar de su elevada homología de secuencia, ambos dominios presentan importantes diferencias en cuanto al potencial electrostático superficial del sitio de unión. La mayor parte de la superficie de unión del dominio hYap-WW1 es neutra, salvo una zona en la parte posterior del bolsillo xY que muestra un potencial positivo. En cambio, el dominio hYap-WW2 muestra una mayor polarización de su sitio de unión. Toda la región del bolsillo xY y del lazo entre las hebras  $\beta 2$  y  $\beta 3$ , incluyendo la lisina K35, presenta un potencial electrostático positivo, mientras que el otro bolsillo (xP) y la región contigua, correspondiente al otro lazo del dominio, muestra un potencial electrostático totalmente negativo. En el Panel C de la Figura 4.6, se han representado los ligandos girados  $180^\circ$  sobre su eje longitudinal para dejar ver la superficie orientada hacia los dominios WW en la interfase de unión. Como es de



**Figura 4.6. Cálculo del potencial electrostático superficial:** **A)** para los modelos basados en los complejos formados por los dominios WW y los ligandos **Lats1-b**, **Ptch1-a** y **Ptch1-b**, **B)** para los modelos de los dominios WW libres (hYap-WW1 y hYap-WW2) y **C)** para los modelos de los ligandos libres. Sobre los dominios WW se han indicado los bolsillos xP y xY. Los ligandos se han girado 180° para mostrar la interfase de unión y se ha denotado la posición en sus secuencias de acuerdo con la nomenclatura adoptada para los ligandos en el capítulo 3. Las superficies con un potencial positivo o negativo se han coloreado en azul o rojo respectivamente, de acuerdo con la barra del Panel A. Todos los modelos se han generado mediante la sustitución de cadenas laterales a partir del complejo formado por el dominio hYap-WW1 (L30K) y el ligando GTPPPYTVG (PDB 1JMQ). Los cálculos de potencial electrostático se hicieron con la aplicación Delphi (dentro de Discovery Studio) por separado para cada dominio, ligando o complejo

esperar, los ligandos libres presentan un potencial electrostático muy positivo debido a los restos básicos que presentan en sus secuencias.

Los potenciales sugieren una interacción electrostática entre el extremo C-terminal de **Lats1-b** y el bolsillo xY del dominio hYap-WW1. Aunque el potencial positivo de este extremo en el caso del ligando **Ptch1-a** es menor, éste también podría interaccionar con la parte posterior del bolsillo xY de hYap-WW1. En cambio, en el dominio de hYap-WW2 el potencial es positivo en todo el bolsillo xY, por lo que la interacción entre esta parte del dominio y el extremo C-terminal de cualquiera de estos ligandos será de naturaleza repulsiva y, por tanto, desfavorable. Estas interacciones electrostáticas entre **Lats1-b** o **Ptch1-a** y el dominio hYap-WW1 están de acuerdo con una contribución entrópica más favorable en comparación con las correspondientes interacciones de ambos ligandos con el dominio hYap-WW2.

El extremo C-terminal del ligando **Ptch1-b**, al igual que sucede con los ligandos **Lats1-b** y **Ptch1-a**, también puede producir una interacción electrostática favorable en su unión a hYap-WW1, pero no en el complejo con hYap-WW2. Además, esta interacción entre **Ptch1-b** y el dominio hYap-WW2 estará también desfavorecida por interacciones electrostáticas repulsivas entre el extremo N-terminal del ligando y la región posterior del bolsillo xP del dominio, ambos con un potencial considerablemente negativo.

En resumen, el estudio de la unión de un conjunto de ligandos naturales de hYap ha revelado que algunos de estos ligandos interaccionan más favorablemente con el dominio hYap-WW1. En estos casos, se ha determinado que distintas contribuciones a la entropía son el origen de la especificidad de unión y se ha sugerido que pueden estar relacionadas con algunas interacciones de tipo electrostático. La determinación de las superficies de potencial electrostático en modelos estructurales confirman esta posibilidad e indican que estas interacciones se establecen entre distintas regiones de la estructura tridimensional dependiendo del dominio WW y del ligando.

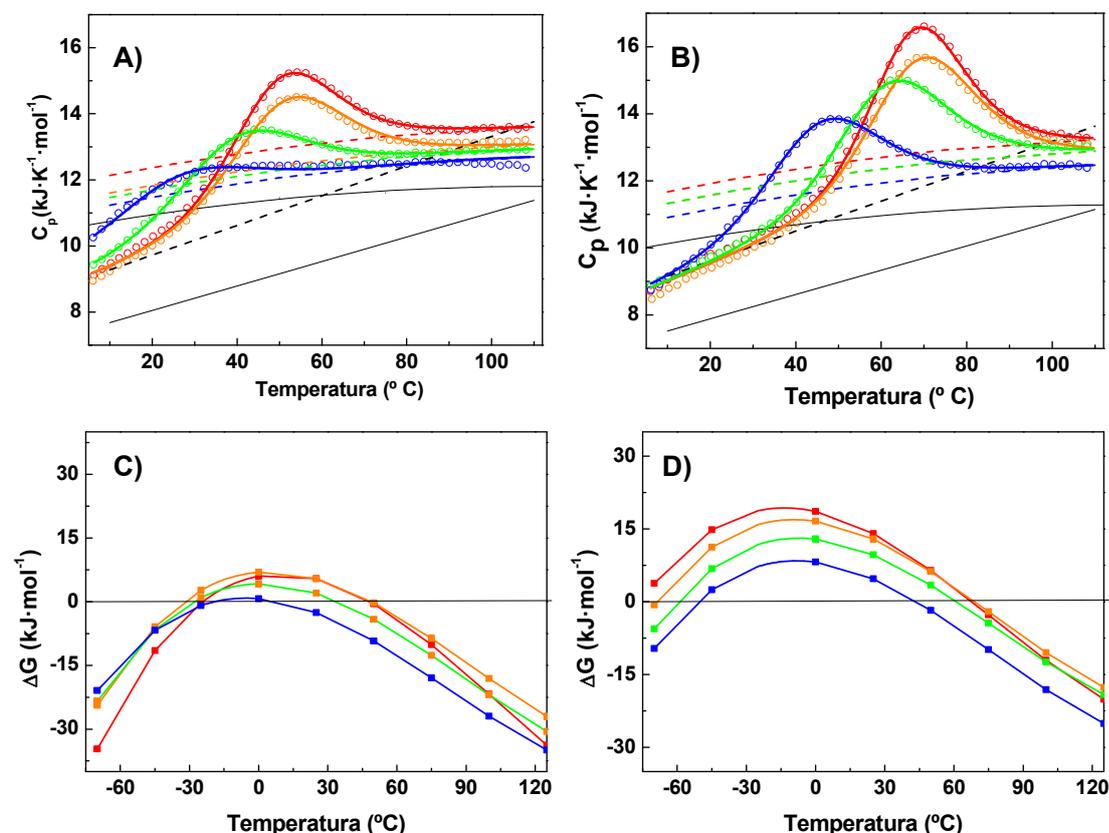
### **4.3. Estudio del equilibrio conformacional de los dominios WW de la isoforma hYap2**

Hemos visto que el segundo dominio WW dota de una mayor especificidad de unión a la isoforma hYap2, pero el estudio de la cooperatividad entre los dominios WW aporta otro tipo de información que también es necesaria para entender las implicaciones biológicas del tándem de dominios WW. El estudio de las interacciones entre los dos dominios WW en la isoforma hYap2 nos permitiría evaluar si existe un comportamiento cooperativo que produzca efectos contextuales en el reconocimiento de los ligandos por el tándem. Para ello, hemos realizado la caracterización del equilibrio conformacional de los dos dominios WW individuales (hYap-WW1 y hYap-WW2) así como de estos dominios WW dispuestos en tándem.

#### **4.3.1. Caracterización del equilibrio conformacional de los dominios individuales hYap-WW1 y hYap-WW2**

Para caracterizar el equilibrio conformacional de los dominios WW de hYap, se han elegido condiciones similares a las escogidas para los estudios de estabilidad con otros dominios WW realizados en el Capítulo 2. En la Figura 4.7 se ha representado la desnaturalización térmica de cada dominio WW aislado mediante DSC a baja fuerza iónica y dentro del intervalo de pH entre 7.0 y 3.0. Una vez comprobado que la reversibilidad del plegamiento es superior al 80 %, se ha procedido al análisis de las trazas calorimétricas de acuerdo con el modelo de desplegamiento más sencillo, el modelo de dos estados (Apéndice I).

Las curvas se analizaron tanto individual (datos no mostrados) como globalmente (Paneles A y B de la Figura 4.7) obteniéndose muy buena calidad en los respectivos ajustes, con valores de  $r$  y  $r^2$  entre 0.99 y 0.999. Las funciones de capacidad calorífica correspondientes al estado nativo y desnaturalizado obtenidas en los ajustes se cruzan bajo la transición endotérmica principal, del mismo modo que sucede en los análisis realizados para otros dominios WW en el Capítulo 2 (Figuras 2.4 y 2.10). Los valores de los distintos análisis globales de los dominios WW de hYap se han recogido en la Tabla 4.2.



**Figura 4.7.** Dependencia de la capacidad calorífica con la temperatura medida por DSC y curvas de estabilidad en diferentes condiciones de pH de los dominios WW individuales de hYap. Experimentos de DSC **A)** del dominio hYap-WW1 y **B)** del dominio hYap-WW2. Los datos experimentales se han representado como círculos vacíos con diferente color según la condición de pH que representan: rojo para fosfato sódico 20mM pH 7.0, naranja para acetato sódico 20mM pH 5.0, verde para acetato sódico 20mM pH 4.0 y azul para glicina 20mM pH 3.0. La línea continua en el mismo código de colores representa el mejor ajuste global según el modelo de dos estados (ecuación I.X en el Apéndice I). Las líneas discontinuas corresponden a las funciones  $C_{p,N}(T)$  en negro y  $C_{p,D}(T)$  en el color asignado a cada pH. Las líneas rectas continuas en negro bajo las transiciones representan la función teórica  $C_{p,N}(T)$  correspondiente a una estimación basada en el peso molecular de cada dominio (Freire 1995). Las líneas curvas en negro representan la función teórica  $C_{p,D}(T)$  calculadas mediante la adición de las contribuciones individuales de cada aminoácido en la secuencia de cada dominio WW (Privalov & Makhatadze 1990). En los paneles inferiores, se muestran las curvas de estabilidad frente a temperatura para **C)** el dominio hYap-WW1 y para **D)** el dominio hYap-WW2. Cada condición de pH se ha representado manteniendo el código de colores de los paneles superiores.

Como se puede observar en las curvas de estabilidad representadas en los Paneles C y D de la Figura 4.7, ambos dominios son poco cooperativos y tienen una estabilidad térmica muy baja en las condiciones estudiadas. Los dominios hYap-WW1 y hYap-WW2 se encuentran plegados entre pH 7.0 y pH 4.0 a baja temperatura (5-20  $^{\circ}\text{C}$ ). El dominio hYap-WW2 presenta una estabilidad mayor que hYap-WW1 en todas las condiciones de pH. La diferencia de estabilidad con el pH

entre ellos es tal que el dominio hYap-WW2 a pH 3.0 presenta una energía libre superior a los valores de la curva de mayor estabilidad del dominio hYap-WW1 (pH 7.0). Además, los parámetros termodinámicos obtenidos para hYap-WW1 son muy parecidos a los que previamente se han determinado para este dominio en la bibliografía (Koepf et al. 1999b; Crane et al. 2000; Ferguson et al. 2001a) y también a los obtenidos para su mutante hYap-WW1(L30K) en el Capítulo 2 (Tabla 2.3). Al igual que sucede con el mutante L30K, la función de energía libre de este dominio no cambia mucho de pH 7.0 a pH 4.0. Sin embargo, a pH 3.0 es negativa en todo el rango de temperaturas, lo que significa que su equilibrio conformacional se halla totalmente desplazado hacia el estado desnaturalizado.

**Tabla 4.2. Parámetros termodinámicos de la desnaturalización térmica seguida por DSC de los dominios individuales de hYap en diferentes condiciones de pH.<sup>a</sup> Todos los parámetros se han obtenido del ajuste global de acuerdo con el modelo de dos estados de los distintos experimentos.**

pH	hYap-WW1			hYap-WW2		
	$\Delta H_{m,exp}$ (kJ·mol <sup>-1</sup> )	$T_m$ (°C)	$F_p$	$\Delta H_m$ (kJ·mol <sup>-1</sup> )	$T_m$ (°C)	$F_p$
7.0	101	48.5	1.0	128	67.8	0.85
5.0	93	49.1	0.5	117	69.0	0.5
4.0	72	34.7	0.3	105	61.3	0.5
3.0	32	9.4	0.1	90	43.9	0.0
	$\Delta C_{p,prot}^b = 1.0 \text{ kJ}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$			$\Delta C_{p,prot}^b = 1.0 \text{ kJ}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$		
<b>Líneas base<sup>c</sup></b>						
$C_{p,N}$	$9.7 - 0.044\cdot(T-T_r)$			$9.6 - 0.044\cdot(T-T_r)$		
$C_{p,D}$	$11.4 + 0.023\cdot(T-T_r) - 10.2\cdot 10^{-5}\cdot(T-T_r)^2$			$11.2 + 0.024\cdot(T-T_r) - 10.3\cdot 10^{-5}\cdot(T-T_r)^2$		

<sup>a</sup>Los errores se han estimado en un valor del 5% para los valores de  $T_m$ , del 10% para los valores de  $\Delta H_m$  y del 20% para los valores de  $F_p$ . <sup>b</sup>El cambio de capacidad calorífica de protonación,  $\Delta C_{p,prot}$ , se ha obtenido a partir de las diferencias en los valores de  $C_p$  a 100°C entre las curvas de DSC a pH 3.0 y a pH 7.0. <sup>c</sup>Los valores de los coeficientes de las líneas bases compartidas entre los diferentes experimentos se definen de acuerdo con el modelo de dos estados (Apéndice I) donde,  $T_r$  es la temperatura de referencia (25 °C). Los valores subrayados se corresponden con los estimados a partir de la contribución individual de cada aminoácido (Privalov & Makhatadze 1990).

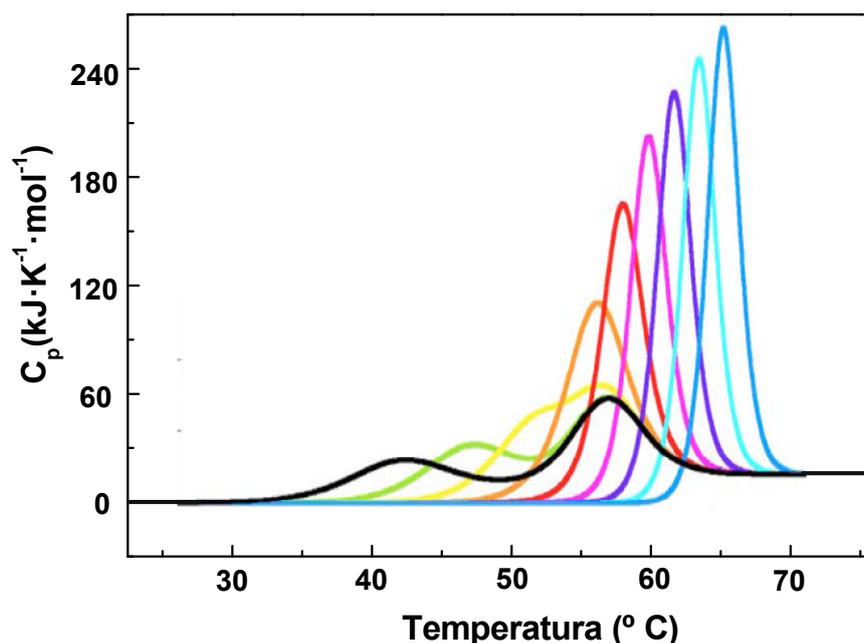
Comparando los parámetros del desplegamiento de los dos dominios WW de hYap con la caracterización del equilibrio conformacional de otros dominios WW realizada en el Capítulo 2, se observa que los parámetros de convergencia de todos los dominios WW son muy parecidos. Los dos dominios de hYap también presentan valores de entalpía específica a 110 °C inferiores a los 54 J·g<sup>-1</sup> típicos de las proteínas globulares. Las curvas de estabilidad de hYap-WW1 y hYap-WW2 además presentan unos máximos a temperaturas anormalmente bajos ( $T_s$  en torno a -2 °C y a -15 °C, respectivamente). En las trazas de DSC también se observan

variaciones de la función de capacidad calorífica molar parcial del estado desnaturalizado con el pH,  $C_{p,D}(T)$ , debidas a diferentes grados de protonación en el estado desnaturalizado (Paneles A y B, Figura 4.7). Además, las funciones de capacidad calorífica molar parcial del estado nativo con la temperatura,  $C_{p,N}(T)$ , obtenidas del ajuste global de los datos experimentales de cada dominio WW se desvían de sus funciones teóricas calculadas en base al peso molecular (Freire 1995), al igual que se ha discutido para los otros dominios WW. Así, los valores de  $C_{p,N}$  (20°C) para hYap-WW1 y para hYap-WW2 (1.80 y 1.82 J·K<sup>-1</sup>·mol<sup>-1</sup>, respectivamente) también superan los 1.73 J·K<sup>-1</sup>·mol<sup>-1</sup> propuestos como valor universal para las proteínas globulares (Privalov & Khechinashvili 1974; Makhatadze 1998).

### 4.3.2. Caracterización del equilibrio conformacional del tándem de dominios WW de hYap

La forma de las transiciones calorimétricas determinadas mediante DSC de una proteína con diferentes dominios puede proporcionar información sobre las interacciones cooperativas entre los dominios. Esta técnica calorimétrica, además de medir directamente las propiedades termodinámicas del proceso de desnaturalización, proporciona una conexión directa entre los datos experimentales y la función de partición de dicho proceso (Freire 1995).

En la Figura 4.8 se muestran, a modo de ejemplo, una serie de simulaciones en las que se ilustra cómo varía el perfil calorimétrico para un tándem de dominios en función del grado de cooperatividad entre ellos (Luque et al. 2002). Estas simulaciones se han realizado teniendo en cuenta dos dominios con diferente estabilidad estructural. Las distintas conformaciones accesibles para el tándem están enumeradas en la Tabla 4.3. De este modo, su función de partición se puede expresar como  $Q=1+\chi\cdot K_1+\chi\cdot K_2+\chi\cdot K_1\cdot K_2$ , donde  $K_1$  y  $K_2$  son las constantes de desplegamiento de cada dominio por separado y  $\chi$  es un parámetro de cooperatividad que hace referencia a la energía de interacción entre los dos dominios. De esta forma, si no existe interacción entre ambos, el parámetro de cooperatividad será  $\chi = 1$ , y la función de partición se reducirá a la expresión

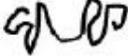


**Figura 4.8.** Simulaciones de perfiles de capacidad calorífica de un tándem de dominios considerando diferentes grados de cooperatividad entre éstos. Se ha definido un parámetro  $\chi$  que tiene en cuenta la interacción entre ambos dominios y por tanto, la función de partición se escribe como  $1 + \chi \cdot K_1 + \chi \cdot K_2 + \chi \cdot K_1 \cdot K_2$ . Los parámetros de estabilidad intrínsecos de cada dominio utilizados en la simulación son  $\Delta H_1 = 418 \text{ kJ} \cdot \text{mol}^{-1}$ ,  $T_{m1} = 67 \text{ }^\circ\text{C}$  y  $\Delta C_{p1} = 8.3 \text{ kJ} \cdot \text{K}^{-1} \cdot \text{mol}^{-1}$  para el primer dominio y  $\Delta H_2 = 251 \text{ kJ} \cdot \text{mol}^{-1}$ ,  $T_{m2} = 42 \text{ }^\circ\text{C}$  y  $\Delta C_{p2} = 8.3 \text{ kJ} \cdot \text{K}^{-1} \cdot \text{mol}^{-1}$  para el segundo dominio. En distintos colores se han representado las curvas esperadas para diferentes grados de cooperatividad: en negro para  $\chi = 1$ , en verde para  $\chi = 1 \cdot 10^{-1}$ , en amarillo para  $\chi = 1 \cdot 10^{-2}$ , en naranja para  $\chi = 1 \cdot 10^{-3}$ , en rojo para  $\chi = 1 \cdot 10^{-4}$ , en magenta para  $\chi = 1 \cdot 10^{-5}$ , en púrpura para  $\chi = 1 \cdot 10^{-6}$ , en azul claro para  $\chi = 1 \cdot 10^{-7}$  y en azul oscuro para  $\chi = 1 \cdot 10^{-8}$ . Esta figura ha sido tomada de Luque et al. 2002.

$Q = (1 + K_1) \cdot (1 + K_2)$ . Esta expresión de la función de partición representa el desplegamiento de dos dominios independientes (situación que se representa en color negro en la Figura 4.8) y el termograma será la suma de las dos transiciones correspondientes a los dominios individuales. En el supuesto de que se produzcan fuertes interacciones interdominio y la cooperatividad sea máxima, entonces el parámetro de cooperatividad,  $\chi$ , tendrá valores muy bajos, próximos a cero. Esta situación está representada por las curvas en colores que van desde el rojo al azul oscuro en la Figura 4.8. En todas estas curvas se observa cómo el tándem se comporta como una única unidad cooperativa, en la que la desnaturalización presenta una única transición estrecha que puede ser descrita por el modelo de dos estados. De hecho, la función de partición se puede escribir como  $Q = 1 + K_{1-2}$ , donde  $K_{1-2}$  es una constante resultado de la combinación de las constantes de desplegamiento de los dos dominios individuales y dependiente del parámetro de

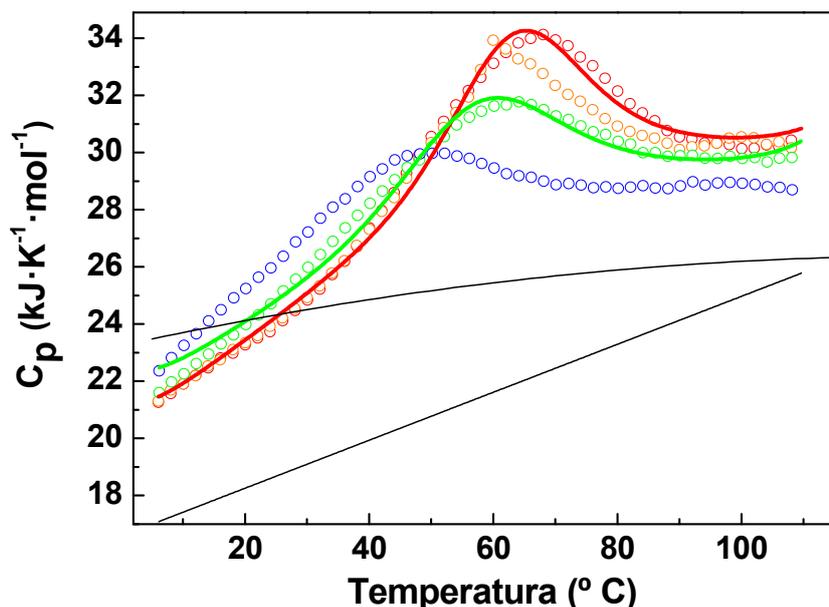
cooperatividad,  $K_{1-2} = \chi \cdot K_1 + \chi \cdot K_2 + \chi \cdot K_1 \cdot K_2$ . A medida que la interacción entre los dos dominios es mayor, se estabilizan mutuamente,  $\chi$  se aproxima a cero y, por tanto,  $K_{1-2}$  tiene un valor más bajo. Esto hace que sólo se observe una transición cada vez más estrecha con valores de  $T_m$  más elevados que las transiciones de los dominios individuales (líneas magenta, azul claro y azul oscuro en la Figura 4.8). Estas características de las curvas de DSC se explican como un aumento de la estabilidad del estado nativo producido por un mayor número de interacciones cooperativas entre los dominios.

Tabla 4.3. Enumeración de los estados accesibles para el desplegamiento de un tándem de dominios.

	Estado	Energía libre ( $\Delta G$ )	Peso estadístico
0		0	1
1		$\Delta G_2 + \Delta g_{int}$	$\chi \cdot K_2$
2		$\Delta G_1 + \Delta g_{int}$	$\chi \cdot K_1$
3		$\Delta G_1 + \Delta G_2 + \Delta g_{int}$	$\chi \cdot K_1 \cdot K_2$

Para estudiar la cooperatividad entre los dos dominios WW de hYap, además de la caracterización del equilibrio conformacional de los dominios individuales hYap-WW1 y hYap-WW2, se ha registrado el perfil de capacidad calorífica molar parcial del tándem,  $C_p$ , en función de la temperatura en idénticas condiciones a las utilizadas en los experimentos de los dominios individuales. En la Figura 4.9 se muestran los termogramas correspondientes al primer barrido de temperatura a pH 7.0, 5.0, 4.0 y 3.0. En todas las condiciones, las transiciones del tándem también presentan una elevada reversibilidad, igual que lo hacen los dominios WW individuales.

Los perfiles de desnaturalización térmica del tándem (Figura 4.9) presentan una forma similar a los de los dominios individuales (Figura 4.7), con una única transición, que decrece en intensidad y cuyo máximo se desplaza a temperaturas menores a medida que el pH se hace más ácido. Estas curvas a distintos valores de pH coinciden en sus valores de  $C_p$  a bajas temperaturas (aproximadamente  $22 \text{ kJ} \cdot \text{K}^{-1} \cdot \text{mol}^{-1}$ ), en lo que corresponde a la capacidad calorífica del estado nativo. En



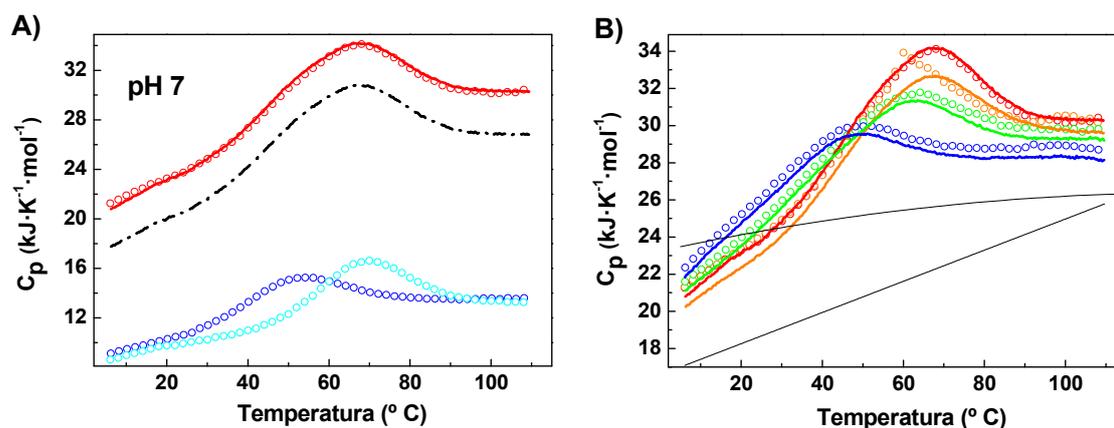
**Figura 4.9.** Dependencia de la capacidad calorífica con la temperatura medida experimentalmente mediante DSC en diferentes condiciones de pH para el tándem hYap-WW1-WW2. Los datos experimentales se han representado como círculos vacíos en diferente color según la condición de pH que representan: rojo para fosfato sódico 20mM pH 7.0, naranja para acetato sódico 20mM pH 5.0, verde para acetato sódico 20mM pH 4.0 y azul para glicina 20mM pH 3.0. Las líneas continuas en rojo y en verde corresponden al mejor ajuste individual según el modelo de dos estados (ecuación I.X del Apéndice I) de los datos experimentales a pH 7.0 y pH 4.0, respectivamente. La línea recta continua en negro bajo los datos experimentales representa la función teórica  $C_{p,N}(T)$  correspondiente a una estimación basada en el peso molecular del tándem (Freire 1995), mientras que la línea continua con cierta curvatura en negro representa la función teórica  $C_{p,D}(T)$  calculada mediante la adición de las contribuciones individuales de cada aminoácido (Privalov & Makhatadze 1990).

general, en el tándem, las funciones  $C_{p,N}(T)$  y  $C_{p,D}(T)$  están peor definidas que en los dominios individuales porque la transición es más ancha y abarca casi todo el intervalo de temperaturas estudiado (aproximadamente desde los 30 °C hasta los 80 °C). La función de  $C_{p,N}(T)$  del tándem es lineal y, al igual que ocurría con los dominios individuales, se caracteriza por una elevada pendiente. La función  $C_{p,D}(T)$  ha tenido que ser definida para cada pH, al igual que sucede con los dominios independientes. Posiblemente, su variación se debe a distintas contribuciones de la protonación de las cadenas laterales durante el desplegamiento. De hecho, el valor estimado para el cambio de la capacidad calorífica de protonación a 100 °C,  $\Delta C_{p,prot}(100\text{ °C})$ , es de unos de 2  $\text{kJ}\cdot\text{K}^{-1}\cdot\text{mol}^{-1}$ , y está en concordancia con la adición de los valores estimados para los dominios individuales. Además, se observa que la traza calorimétrica recogida a pH 5.0 presenta un artefacto

reproducibile en torno a 60 °C, que podría deberse a eventos de agregación parcial originados por la proximidad de este pH al punto isoeléctrico del tándem ( $pI_{\text{teórico}} = 4.88$ ).

Para investigar el nivel de cooperatividad entre los dominios en el tándem, en primer lugar, se ha realizado el ajuste individual de las curvas de DSC de acuerdo con el modelo de dos estados (Apéndice I). Sin embargo, este modelo no reproduce adecuadamente ninguno de los perfiles de  $C_p$  del tándem de dominios WW de hYap. En la Figura 4.9, se han representado los mejores resultados obtenidos, correspondientes a las condiciones de pH 7.0 y pH 4.0 y, como se puede apreciar, la desviación del ajuste con el modelo de dos estados de los datos experimentales se produce en todo el intervalo de temperatura. Este resultado es indicativo de que el desplegamiento del tándem de estos dominios WW no se produce como una única unidad cooperativa. No obstante, este resultado tampoco descarta la existencia de un cierto grado de cooperatividad entre los dominios WW, o entre ellos y la secuencia que los conecta, que pueda inducir a una estabilización entre dominios, a una orientación espacial como la observada en el tándem de dominios WW de Prp40 (Wiesner et al. 2002) o a otros efectos contextuales, como por ejemplo efectos autorregulatorios (Sudol et al. 2005; Jennings et al. 2007).

A raíz de estos resultados, en segundo lugar, hemos considerado el caso opuesto, es decir, hemos supuesto que no hay ningún efecto cooperativo entre ellos. En este contexto, el valor de  $\chi$  es 1 y su función de partición es  $Q=(1+K_1)\cdot(1+K_2)$ , o lo que es lo mismo, la capacidad calorífica molar parcial del tándem es equivalente a la adición de las respectivas capacidades caloríficas de los dominios individuales ( $C_{p,\text{tándem}}=C_{p,\text{WW1}}+C_{p,\text{WW2}}$ ). Para comprobar si éste es nuestro caso hemos comparado los distintos perfiles de desnaturalización por DSC en la Figura 4.10. En el Panel A, se han representado las curvas de DSC a pH 7.0 de los dominios individuales hYap-WW1 y hYap-WW2 (círculos en colores azul oscuro y azul claro, respectivamente). Comprobamos que ambas curvas solapan casi totalmente debido a que los dominios WW individuales presentan transiciones muy anchas. En esta misma figura, la línea discontinua en color negro representa la suma de los perfiles de  $C_p$  de ambos dominios, y como se puede apreciar,



**Figura 4.10. Comparación de los datos experimentales de DSC del tándem hYap-WW1-WW2 con las curvas de DSC de los dominios WW individuales. A)** Adición de la dependencia con la temperatura de las capacidades caloríficas de los dominios WW individuales en tampón fosfato sódico pH 7.0. Los datos experimentales normalizados de los dominios WW individuales se muestran como círculos de color azul oscuro para el dominio hYap-WW1 y de color azul claro para el dominio hYap-WW2. La línea discontinua negra (---) representa la curva resultante de la adición lineal de las curvas para los dominios individuales. La línea continua de color rojo resulta de la adición a la curva anterior de la contribución a la capacidad calorífica de cada aminoácido de la secuencia de conexión entre los dos dominios (Privalov & Makhatadze 1990). **B)** Comparación de los datos experimentales de DSC del tándem hYap-WW1-WW2 con las curvas de DSC de los dominios WW individuales en diferentes condiciones de pH. Los círculos corresponden a los datos experimentales empleando el mismo código de colores que en la Figura 4.9. Las líneas representan la adición lineal de las curvas experimentales de los dominios individuales más la contribución de la secuencia que los conecta para cada condición experimental. La líneas continuas en negro bajo los datos experimentales representan la función teórica  $C_{p,N}(T)$  y la función teórica  $C_{p,D}(T)$  calculadas como se detalla en la Figura 4.9.

aunque no describe el perfil de DSC del tándem (círculos en color rojo) presenta exactamente su misma forma. Si además, a la suma de los perfiles de DSC de los dominios hYap-WW1 y hYap-WW2 se añade la contribución a la  $C_p$  de la secuencia de conexión entre los dos dominios de acuerdo con los valores tabulados por cada aminoácido (Privalov & Makhatadze 1990), entonces la curva resultante (la línea continua de color rojo) describe perfectamente los datos experimentales. En el Panel B de la Figura 4.10, se ha realizado el mismo procedimiento para las curvas obtenidas a otros valores de pH y se han obtenido resultados similares. Los perfiles de DSC del tándem en todas las condiciones de pH se describen adecuadamente por la adición de las contribuciones de los dominios individuales más la de la secuencia que los conecta, salvando ligeras desviaciones para las curvas a pH 4 y pH 3. Las desviaciones podrían tener relación con las pequeñas contribuciones a la  $C_p$  debidas a la protonación de algunos residuos en la secuencia de conexión. Estas contribuciones de la

protonación normalmente son de pequeña magnitud respecto al cambio de  $C_p$  en el desplegamiento y no se suelen tener en cuenta. En la tabulación de los valores de  $C_p$  por resto, tampoco se consideraron (Makhatadze & Privalov 1990; Privalov & Makhatadze 1990).

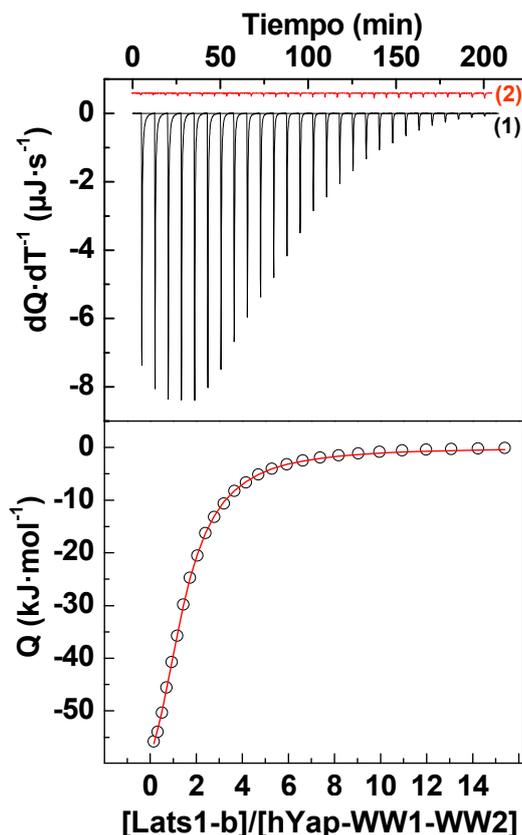
Al comprobar que la relación  $C_{p,tándem} = C_{p,WW1} + C_{p,WW2} + C_{p,conector}$  describe con precisión los perfiles de  $C_p$  obtenidos experimentalmente para el tándem a distintos valores de pH, podemos considerar que  $\chi$  vale uno y, por tanto, que ambos dominios se comportan de manera totalmente independiente durante el desplegamiento. En resumen, el tándem de dominios WW de hYap, no presenta ningún tipo de cooperatividad conformacional entre los dos dominios, ni tampoco entre ellos y la secuencia que los conecta.

#### **4.4. Reconocimiento de ligandos naturales por el tándem de dominios WW de hYap**

Hemos planteado previamente los posibles mecanismos mediante los cuales la presencia del segundo dominio en la isoforma hYap2 puede dotar a la proteína de nuevas funcionalidades. En este capítulo, se ha investigado la unión de los dominios WW individuales de esta isoforma a una serie de ligandos naturales de hYap y se ha caracterizado el equilibrio conformacional de los dominios individuales y del tándem. Los resultados obtenidos indican, por un lado, que los dominios hYap-WW1 y hYap-WW2 presentan una especificidad de unión diferente y, por otro lado, que no existe ninguna cooperatividad conformacional entre los dos dominios en tándem durante el desplegamiento. Es interesante señalar que, en ocasiones, un mecanismo de modulación mediante una estabilización por la formación de complejos supra-macromoleculares entre los ligandos y diferentes módulos de la misma proteína también pueden tener efectos importantes en el reconocimiento molecular de los dominios WW (Huang et al. 2000; Sudol et al. 2005; Chong et al. 2010). En el caso del tándem de dominios WW de hYap, este tipo de interacciones también puede proporcionar una funcionalidad diferente a la isoforma hYap2. Con el fin de investigar este comportamiento para el tándem en la isoforma hYap2, se han realizado los experimentos de titulación del tándem hYap-WW1-WW2 con los cinco ligandos

naturales **Lats1-a**, **Lats1-b**, **Lats2**, **Ptch1-a** y **Ptch1-b** a 25 °C y en tampón fosfato sódico 20 mM, pH 7.0.

En la Figura 4.11, a modo de ejemplo, se representa la titulación calorimétrica del tándem hYap-WW1-WW2 con el ligando **Lats1-b**.



**Figura 4.11.** Titulación del tándem hYap-WW1-WW2 con el el ligando **Lats1-b**. En el panel superior, se presenta el termograma de la titulación del tándem hYap-WW1-WW2 a 25 °C en fosfato sódico 20 mM, pH 7.0 del ligando **Lats1-b** (1), en el que se muestran los calores por unidad de tiempo liberados tras la inyección de ligando sobre una disolución de proteína, y el termograma de dilución del ligando (2) en las mismas condiciones sobre el tampón (en rojo y desplazado en el eje para mayor claridad). En el panel inferior, se observa la isoterma de unión correspondiente a la normalización de los calores derivados de cada inyección, corregidos teniendo en cuenta los calores debidos a la dilución del ligando. Los datos experimentales se representan como círculos, mientras que la línea continua corresponde al mejor ajuste según el modelo de unión de un ligando a una macromolécula con dos clases de sitios diferentes e independientes (ecuación IV.XI del Apéndice IV).

A raíz de los resultados obtenidos anteriormente para el estudio de los dominios individuales, cabe esperar que el tándem reconozca estos ligandos, aunque pueden existir algunas variaciones en el modo de reconocimiento que afecten a la afinidad de unión. Las titulaciones del tándem con aquellos ligandos que unen a ambos dominios individuales, pero que muestran cierta preferencia de

unión por uno de ellos, **Lats1-b** y **Ptch1-a**, se han analizado de acuerdo con el modelo que considera dos clases de sitios distintos e independientes (SDI, Apéndice IV), mientras que para el resto de titulaciones del tándem se ha utilizado el modelo que tiene en cuenta sitios de unión idénticos e independientes (SII, Apéndice III).

Los parámetros termodinámicos obtenidos a partir de estos análisis de los experimentos de titulación, tanto del tándem como de los dominios WW individuales, se han recogido en la Tabla 4.4.

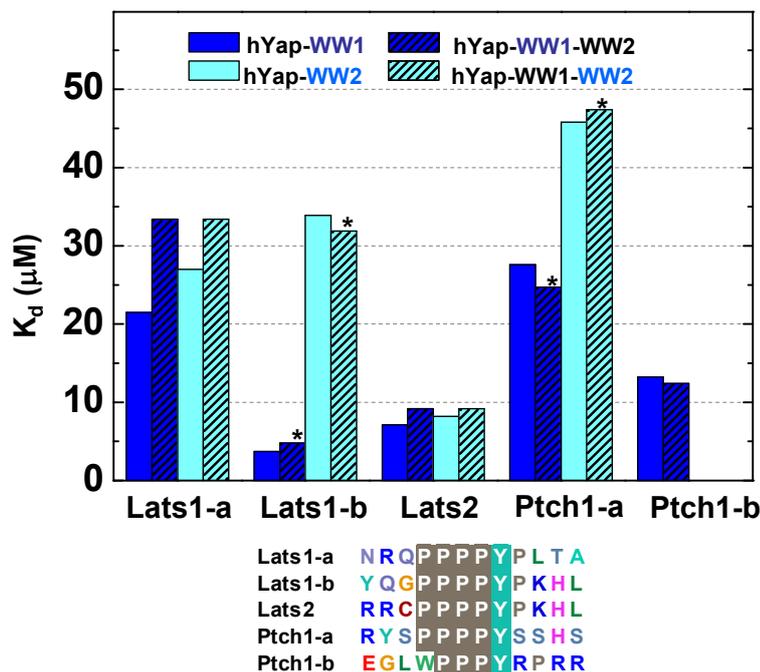
En los análisis de las titulaciones con los ligandos **Lats1-a**, **Lats2** y **Ptch1-b** (realizados con el modelo de unión SII), el valor de  $n$  (el número de sitios) se ha dejado flotar durante el ajuste, obteniéndose en todos los casos valores en torno a dos. La única excepción se observa para la interacción del tándem con el ligando **Ptch1-b**, cuyo valor de  $n$  es igual a 1.23. En este caso, al estequiometría de unión confirma que, al igual que ocurría con los dominios individuales, este ligando sólo es reconocido por el primer dominio WW de hYap, hYap-WW1. Posiblemente exista una interacción residual con el segundo dominio con una afinidad muy débil, varios órdenes de magnitud inferior, lo que se refleja en que el número de sitios sea algo superior a 1. En los análisis realizados para las titulaciones de **Lats2** y **Ptch1-a** (según el modelo SDI), el valor  $n$  se ha fijado en 1 para cada dominio WW para minimizar el número de variables ajustables en estos análisis.

En la Figura 4.12, junto a las constantes de disociación obtenidas a partir de las titulaciones de los dominios individuales con cada uno de los ligandos, se han representado las constantes de disociación correspondientes a las titulaciones del tándem. En esta figura se observa que, en general, las constantes de disociación de los complejos con el tándem hYap-WW1-WW2 prácticamente no varían con respecto a las obtenidas para los dominios individuales hYap-WW1 y hYap-WW2. Los ligandos **Lats1-a** y **Lats2** no muestran preferencia de unión por uno u otro dominio en el tándem, los ligandos **Lats1-b** y **Ptch1-a** tienen mayor afinidad por uno de los dos dominios, y el ligando **Ptch1-b** sólo une a uno de los dominios del tándem. Estos resultados sugieren que, como cabía esperar a partir de los datos de DSC, no hay ninguna clase de cooperatividad entre los dominios WW en tándem durante el reconocimiento de estos ligandos naturales de hYap.

**Tabla 4.4. Parámetros termodinámicos para la unión de ligandos naturales a los dominios WW de hYap.** Todos los datos se han obtenido a partir del análisis mediante el modelo de unión de un ligando a una macromolécula con “n” sitios de unión idénticos e independientes<sup>1</sup> (SII) (ecuación X.III del Apéndice III) o mediante el modelo de unión de un ligando a una macromolécula con dos clases de sitios de unión diferentes e independientes<sup>2</sup> (SDI) (ecuación IV.XI del Apéndice IV) de distintos experimentos de ITC con dominios WW individuales y en tándem realizados a 25 °C en tampón fosfato sódico 20 mM pH 7.0 (más β-mercaptoetanol 10 mM en los experimentos con el ligando **Lats2**).

	Ligandos	n	<sup>a</sup> Dominio hYap-WW1				<sup>a</sup> Dominio hYap-WW2				
			K <sub>d</sub> (μM)	ΔG <sub>ap</sub> (kJ·mol <sup>-1</sup> )	ΔH <sub>ap</sub> (kJ·mol <sup>-1</sup> )	T·ΔS <sub>ap</sub> (kJ·mol <sup>-1</sup> )	n	K <sub>d</sub> (μM)	ΔG <sub>ap</sub> (kJ·mol <sup>-1</sup> )	ΔH <sub>ap</sub> (kJ·mol <sup>-1</sup> )	T·ΔS <sub>ap</sub> (kJ·mol <sup>-1</sup> )
Dominios WW individuales	<sup>1</sup> LATS1-a NRQPPPPYPLTA	0.96	21.5	-26.7	-67.0	40.3	1.09	27.0	-26.1	-76.2	50.1
	<sup>1</sup> LATS1-b YQGPPPPYPKHL	0.91	3.7	-31.0	-69.0	38.0	0.94	33.9	-25.6	-70.2	44.6
	<sup>1</sup> LATS2 RRCPPPPYPKHL	1.05	7.1	-28.8	-78.5	49.7	1.08	8.2	-29.1	-57.8	28.7
	<sup>1</sup> PTCH1-a RYSPPPPYSSHS	0.95	27.6	-26.1	-65.0	38.9	0.91	45.8	-24.8	-65.9	41.1
	<sup>1</sup> PTCH1-b EGLWPPPYRPRR	1.09	13.2	-27.9	-75.2	47.3	n.i.	--	--	--	--
	<sup>1</sup> LATS1-a NRQPPPPYPLTA	2.04	33.4	-25.6	-72.0	46.4					
	<sup>2</sup> LATS1-b YQGPPPPYPKHL	<sup>b</sup> 1	4.8	-30.4	-70.8	40.4	<sup>b</sup> 1	31.9	-25.7	-51.8	26.1
Tándem dominios WW	<sup>1</sup> LATS2 RRCPPPPYPKHL	2.02	9.2	-28.8	-68.8	40.0					
	<sup>2</sup> PTCH1-a RYSPPPPYSSHS	<sup>b</sup> 1	24.7	-26.4	-52.9	26.5	<sup>b</sup> 1	47.4	-24.7	-75.2	50.5
	<sup>1</sup> PTCH1-b EGLWPPPYRPRR	1.23	16.9	-27.3	-80.2	52.9					

<sup>1</sup>Datos analizados mediante el modelo de unión SII. <sup>2</sup>Datos analizados mediante el modelo de unión SDI. n.i.: No hay interacción. <sup>a</sup>Los errores en los ajustes por regresión no lineal de los datos experimentales deben estimarse en torno al 5% para la entalpía de unión y al 10% para la constante de disociación. <sup>b</sup>El valor del número de sitios, n, se ha mantenido fijo en 1 durante el análisis de estos experimentos.



**Figura 4.12. Afinidad de unión de los ligandos naturales a los dominios individuales y en tándem de hYap.** Representación en forma de barras de las constantes de disociación para la unión de los ligandos **Lats1-a**, **Lats1-b**, **Lats2**, **Ptch1-a** y **Ptch1-b** al tándem hYap-WW1-WW2, junto con las constantes de los dominios individuales, hYap-WW1 y hYap-WW2, de la Figura 4.4. Se ha empleado el mismo código de colores que en dicha figura, correspondiendo las barras de colores lisos a los dominios individuales y las de colores rayados a los dominios en tándem. Todas las constantes se han obtenido mediante distintos experimentos de ITC en tampón fosfato sódico 20 mM pH 7.0 a 25 °C (en presencia de β-mercaptoetanol 10 mM en aquellos experimentos con el ligando **Lats2**). Los errores asociados a las constantes de disociación se estiman en torno al 10 %. El asterisco indica que las constantes de disociación del tándem se han obtenido mediante el análisis con el modelo de SDI (Apéndice IV), mientras que el resto corresponden a análisis con el modelo SII (Apéndice III). En la parte inferior, se ha representado un alineamiento de las secuencias de los ligandos empleados en el estudio.

En la Tabla 4.4 se recogen los cambios de energía libre para la unión del tándem hYap-WW1-WW2 con los distintos ligandos y las correspondientes disecciones termodinámicas de estas interacciones. A partir de estos datos podemos comparar las contribuciones entálpicas y entrópicas a la afinidad de unión del tándem con las obtenidas para las interacciones de los dominios WW individuales. En el caso de los ligandos **Lats1-a** y **Lats2**, el modelo SII promedia el cambio de entalpía si se compara con los valores obtenidos para las titulaciones de los dominios individuales. En cambio, este mismo modelo SII, obtiene valores muy similares para las magnitudes termodinámicas determinadas para las interacciones de **Ptch1-b** con el tándem (cuyo valor de n es 1.23) y de este ligando

con el dominio hYap-WW1 ( $n=1.09$ ). Por otro lado, el modelo SDI reconoce como sitios de unión diferentes los dominios WW del tándem en sus interacciones con los ligandos **Lats1-b** y **Ptch1-a**. Aunque este modelo proporciona una medida realista de la afinidad de cada sitio de unión, al ser las constantes bastante parecidas, la correlación entre los parámetros termodinámicos ( $\Delta H_{ap}$  y  $n$ ) de ambos sitios es elevada y no permite distinguir la energética de cada sitio. Así, en la interacción de **Lats1-b** con el tándem, los incrementos de entalpía y entropía correspondientes al segundo dominio WW son menores que los determinados para el dominio individual hYap-WW2 (Tabla 4.4), mientras que, para la unión del ligando **Ptch1-a** al tándem, los términos energéticos están subestimados para el primer dominio y sobrestimados para el segundo dominio WW, en comparación con los parámetros obtenidos para la interacción con los dominios individuales.

En resumen, los resultados para el reconocimiento de ligandos, recogidos en la Figura 4.11 y en la Tabla 4.4, están en la línea de lo descrito para el equilibrio conformacional del tándem de dominios WW. Estos datos indican que no existe cooperatividad entre los dominios WW en el tándem de la isoforma hYap2 durante la unión a los ligandos naturales que hemos seleccionado.

#### **4.5. Implicaciones funcionales del segundo dominio WW en la isoforma hYap2**

A partir de la información obtenida, y teniendo en cuenta las diferentes afinidades de los dominios WW de hYap por los ligandos estudiados, es posible plantear ciertas hipótesis sobre el mecanismo mediante el cual la presencia o ausencia del dominio hYap-WW2 pudiera contribuir a la regulación funcional de las distintas isoformas de hYap. Así, por ejemplo, la unión de los dominios WW de la forma hYap2 a los dos motivos PPxY encontrados en la quinasa de serina Last1, sugiere que la orientación durante la fosforilación de algunas de las distintas dianas de serina está doblemente modulada por los dominios WW que presenta esta isoforma y, por tanto, su especificidad será diferente.

La quinasa Lats2 sólo puede ser reconocida mediante un único motivo PPxY (el ligando **Lats2**) y éste presenta una afinidad similar por los dos dominios WW de hYap. Estos resultados sugieren que esta proteína se une indistintamente

al primer dominio WW de cualquiera de las dos isoformas sin que se puedan pronosticar diferencias funcionales entre ambas isoformas de hYap. Por otro lado, la unión al segundo dominio WW en la isoforma hYap2 podría inducir fosforilaciones en posiciones diferentes, dotando a esta isoforma de otras funcionalidades (Komuro et al. 2003).

Es posible que el reclutamiento citosólico de la forma hYap2 mediante la unión a Ptch1, presenta una complejidad mayor que el de la isoforma hYap1. Esta última reconoce dos sitios de unión en Ptch1 con afinidades bastante similares, pero la otra isoforma, hYap2, presenta dos dominios WW que reconocen la secuencia **Ptch1-a** con diferente afinidad y un sólo dominio (hYap-WW1) que reconoce a **Ptch1-b**. Además, en el supuesto de que hYap2 reconozca a la vez las dos secuencias de Ptch1, esta interacción proteína-proteína tendría una orientación clara debida a las preferencias de unión de los dominios WW, lo cual sugiere de nuevo, posibles efectos de modulación en la funcionalidad de hYap2.

En resumen, nuestro análisis indica que la presencia de dos dominios WW con capacidad para reconocer distintos ligandos permite modular la especificidad y el modo de interacción de las dos isoformas de hYap con sus distintos ligandos y, por tanto, añadir un nivel adicional de regulación funcional. En el contexto de la construcción en tándem de los dos dominios WW aislados del resto de la proteína, hemos descartado la existencia de cooperatividad entre ambos dominios, como se ha descrito para otros dominios WW (Fedoroff et al. 2004; Jennings et al. 2007; Ramirez-Espain et al. 2007; Huang et al. 2009; Chong et al. 2010). No obstante, es importante tener en cuenta que efectos cooperativos adicionales podrían surgir asociados a factores contextuales, bien originados por interacciones con ligandos con motivos de unión consecutivos en su secuencia o bien, en el contexto de proteínas completas, originados por interacciones adicionales establecidas entre los dominios WW y otros elementos en la estructura de hYap. En este sentido, el análisis del equilibrio conformacional y la unión de ligandos en construcciones más complejas de hYap suscitan un gran interés.

## CAPÍTULO 5



## CAPÍTULO 5. Conclusiones

### Capítulo 2

De la caracterización del equilibrio conformacional de los distintos dominios WW, se concluya que:

1. Estos dominios pueden clasificarse como proteínas que contienen una lámina tipo  $\beta$ -II.
2. El equilibrio conformacional se encuentra en el límite de lo que puede considerarse un comportamiento cooperativo, lo que indica que la barrera que separa los macroestados nativo y desnaturalizado es de pequeña magnitud.

### Capítulo 3.

Del análisis termodinámico-estructural de la unión de ligandos L víricos al tercer dominio WW de la ligasa de ubiquitina Nedd4 se extraen las siguientes conclusiones:

3. El análisis termodinámico de la unión de varios dominios WW a un conjunto de ligandos indica que el proceso está gobernado por una fuerte componente entálpica y penalizado por efectos de naturaleza entrópica
4. Las secuencias de dominios L aisladas del resto de la proteína vírica reconocen de forma específica al tercer dominio WW de la ligasa Nedd4.
5. El análisis de la termodinámica de unión y la caracterización estructural de los complejos formados entre el dominio hNedd4-WW3 y los ligandos **HTLV1-b** y **Ébola-a** revelan dos modos diferentes de reconocimiento de estos ligandos relacionados con la elevada flexibilidad conformacional del dominio hNedd4-WW3.

#### Capítulo 4.

Del estudio de reconocimiento de ligandos naturales por el tándem de dominios WW de hYap se obtienen las siguientes conclusiones:

6. Los cinco ligandos naturales, **Lats1-a**, **Lats1-b**, **Lats2**, **Ptch1-a** y **Ptch1-b** son reconocidos por al menos unos de los dominios WW de hYap.
7. La presencia el dominio hYap-WW2 modula la especificidad de hYap, dotando a la isoforma hYap2 de la posibilidad de establecer interacciones adicionales de diferente naturaleza a las que puede establecer el dominio hYap-WW1.
8. Los dominios WW de hYap2 se comportan como unidades independientes, tanto en lo que respecta a su equilibrio conformacional como al reconocimiento de ligandos.

## CAPÍTULO 6



## **CAPÍTULO 6. Materiales y Métodos**

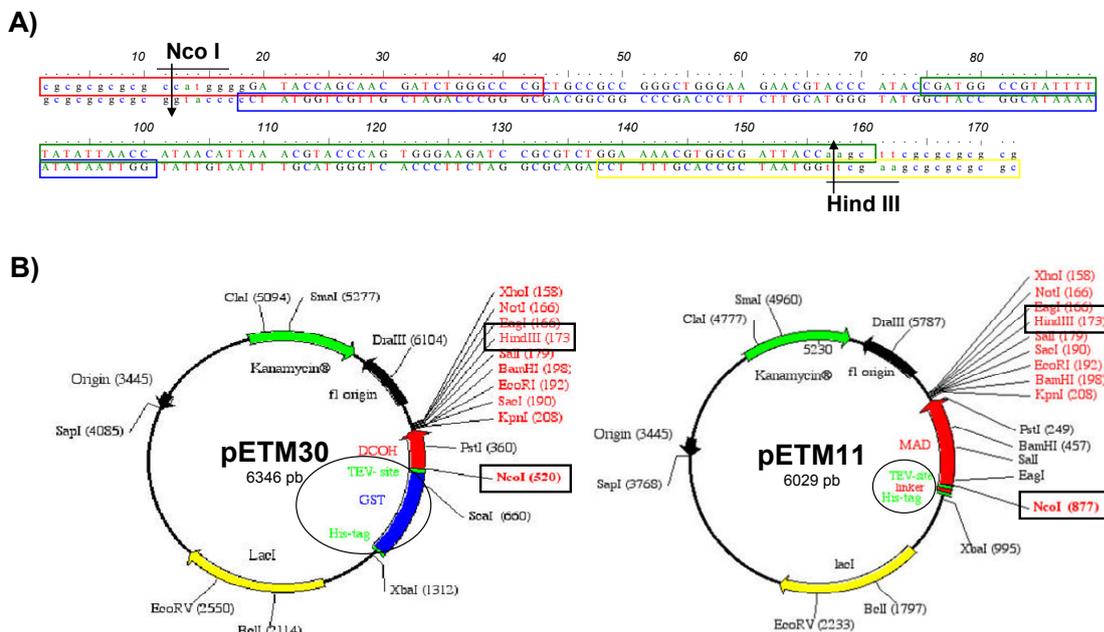
### **6.1. Clonado, expresión y purificación de los dominios WW**

#### **6.1.1. Dominios WW**

Los genes que codifican los dominios WW3 y WW4 de la ligasa de ubiquitina humana Nedd4 (código P46934 de la base de datos SwissProt/TrEMBL), los dominios WW1 y WW2 de la proteína humana Fbp11 (código O75400) y del mutante L30K del dominio WW1 de Yap (Macias et al. 1996) fueron construidos en nuestro laboratorio mediante el método de la reacción en cadena de la polimerasa (PCR: Polymerase Chain Reaction) (Mullis & Faloona 1987; Ausubel 1992; Creighton 1993) a partir de cuatro nucleótidos solapantes. La secuencia resultante se diseñó para que cada secuencia de nucleótidos que codifica los dominios WW esté situada entre dos dianas de las enzimas de restricción NcoI y HindIII como se muestra en la Figura 6.1 para el dominio hNedd4-WW4. El producto de la PCR fue, a continuación, digerido por dichas enzimas y clonado en el plásmido pETM30 (G. Stier, EMBL, Heidelberg, Alemania) que se muestra en la Figura 6.1B. La identidad e integridad de los genes clonados fueron confirmadas mediante secuenciación de nucleótidos por el Servicio de Secuenciación del Instituto de Parasitología y Biomedicina “López Neyra” del Consejo Superior de investigaciones Científicas (C.S.I.C), en Granada. Los genes que codifican los dominios WW1 y WW2 de Yap o el tándem de ambos dominios WW (código P46937-2) fueron adquiridos a la empresa GENEART AG (Regensburg, Alemania) y reclonados en el vector pETM30 siguiendo el mismo proceso.

Los genes de todos los dominios WW fueron expresados en células *E. Coli* BL21-DE3 (NOVAGEN) como proteínas de fusión a Glutación-S-Transferasa (GST) y a una cola de histidinas (6xHis tag) en el extremo N-terminal, con un sitio de reconocimiento para la proteasa del mosaico del tabaco (TEV). Para su expresión, se empleó como medio de cultivo el Luria-Bertani (LB) en presencia de kanamicina 1mM y se añadió IPTG como agente inductor hasta alcanzar una concentración final de 1mM, cuando la absorbancia del cultivo celular a 600 nm

alcanzó valores entre 0.6 y 0.8. La incubación se llevó a cabo durante toda la noche a 37 °C.



**Figura 6.1. Clonación y expresión de dominios WW.** A) Secuencia de nucleótidos completa de hNedd4-WW4 y en cuadros de diferentes colores, los oligonucleótidos seleccionados para la producción del dominio mediante dos RCP: oligonucleótido 1 en rojo, oligonucleótido 2 en azul, oligonucleótido 3 en verde y oligonucleótido 4 en amarillo. En negro se han resaltado las dianas de corte de las enzimas de restricción Nco I y Hind III. B) Representación esquemática o mapa de los elementos presentes en los vectores plasmídicos utilizados para el clonaje de los genes de los dominios WW. En la izquierda, se representado el mapa del plásmido pETM30, y en la derecha, el mapa del plásmido pETM11

Las células se recogieron por centrifugación a 7000 rpm y 4 °C durante 10 minutos (centrífuga HETTICH modelo Roto Super 40, rotor A6.9) y posteriormente se resuspendieron en tampón fosfato sódico 50 mM, cloruro sódico 0.3 M, pH 8.0 (tampón A); después se lisaron mediante prensa French (AMINCO) a una presión de 1200 psi y se recogieron por ultracentrifugación (BECKMAN Optima LE-80, rotor 45Ti) a 30.000 rpm y 4 °C durante 30 minutos. Los dominios WW se purificaron mediante cromatografía de afinidad a Níquel, para lo cual el sobrenadante resultante de la centrifugación se pasó por una columna con 5 mL de resina Ni-NTA (QUIAGEN) previamente equilibrada con 50 mL de tampón A y se llevaron a cabo sucesivos lavados de 50 mL con tampón A, tampón A + 20 mM imidazol y tampón A + 50 mM imidazol. Por último, la proteína se eluyó con 30 mL de tampón A + 500 mM imidazol y aquellas fracciones que contenían la proteína se

dializaron frente a tampón A para eliminar el imidazol. A continuación, se eliminó la cola de histidinas en posición N-terminal mediante una etapa de hidrólisis controlada para lo cual la proteína purificada se incubó con la proteasa TEV a temperatura ambiente durante 4 horas en una relación de peso aproximada de 100:1. Los productos de la hidrólisis se sometieron a una segunda etapa de cromatografía de afinidad en Ni-NTA, de modo que la cola de histidinas, la proteína no hidrolizada y la propia TEV, que también posee una cola de histidinas, quedaron retenidas en la columna mientras que la proteína libre de la cola de histidinas se recogió en las primeras fracciones de elución en ausencia de imidazol.

**Tabla 6.1. Coeficientes de extinción molar a 280 nm y pesos moleculares determinados para la los distintos dominios WW.**

Dominios WW	Coeficiente de extinción molar ( $M^{-1} cm^{-1}$ )	Peso molecular ( $g mol^{-1}$ )
hNedd4-WW3	11380	5568
hNedd4-WW4	11660	5720
hYap-WW1	12550	5524
hYap-WW1(L30K)	12600	5539
hYap-WW2	13960	5410
hYap-WW1-WW2	26400	12529
hFbp11-WW1	15440	5323
hFbp11-WW2	16990	5697

Las diferentes etapas de la purificación así como la pureza final fueron controladas mediante electroforesis en gel de acrilamida (SDS-page) al 16 %. Aquellas fracciones que contenían alguna impureza, se sometieron a una etapa adicional de cromatografía de exclusión molecular utilizando una columna Hi-Load Superdex-75 26/60 (AMERSHAM BIOSCIENCIAS) equilibrada en tampón A. Posteriormente, mediante espectrometría de masas, se comprobó la pureza de la proteína (superior al 99 %) y se determinó el peso molecular correcto, cuyos valores son recogidos en la Tabla 6.1. Finalmente, las fracciones de proteína purificada se concentraron hasta  $2-3 mg \cdot mL^{-1}$ , se congelaron en nitrógeno líquido y se almacenaron a  $-80\text{ }^{\circ}C$  en el mismo tampón A. En estas condiciones los dominios WW son estables durante varios meses. El rendimiento final aproximado

para cada purificación fue de unos 5-10 mg de proteína por cada litro de medio de cultivo.

### 6.1.2. Dominio hNedd4-WW3 enriquecido en $^{13}\text{C}$ y/o $^{15}\text{N}$

El gen que codifica el dominio WW3 de la ligasa de ubiquitina humana Nedd4, clonado en el plásmido pETM30, se cortó mediante una digestión con las enzimas de restricción NcoI y HindIII y se clonó de nuevo en el plásmido pETM11 (G. Stier, EMBL, Heidelberg, Alemania) (Figura 6.1B). La expresión y purificación de este dominio enriquecido en  $^{13}\text{C}$  y/o  $^{15}\text{N}$  (en dos procedimientos independientes para obtener una muestra sólo enriquecida en  $^{15}\text{N}$  y otra enriquecida en los dos isótopos,  $^{13}\text{C}$  y  $^{15}\text{N}$ ) se realizó en colaboración con el Dr. Francisco J. Blanco en el Centro Nacional de Investigaciones Oncológicas (CNIO, Madrid).

El dominio hNedd4-WW3 se expresó en células *E. Coli* BL21-DE3 con una cola de histidinas (6xHis tag) en el extremo N-terminal y con un sitio de reconocimiento para la proteasa TEV. Varias colonias se incubaron en LB como preinóculos de 50 mL por cada litro de medio LB final a 37 °C durante toda la noche para aumentar la masa celular de partida. Las células se recogieron por centrifugación (5000 g, 10 minutos y 4 °C), se resuspendieron y se incubaron a 37 °C en 5 litros de medio de cultivo LB en presencia de kanamicina 1mM hasta que la absorbancia del cultivo celular a 600 nm alcanzó valores entre 0.6 y 0.8. Entonces, se hicieron dos lavados sucesivos con 800 mL y 500 mL de medio mínimo salino M9 preparado según se detalla en la Tabla 6.2 con el fin de eliminar subproductos del metabolismo de las bacterias (como lactatos, piruvatos, etc) que pudieran afectar a su crecimiento, y se recogieron de nuevo las células por centrifugación. A continuación, las células se resuspendieron, en 2 L de medio mínimo de cultivo en las proporciones que se detallan en la Tabla 6.2, compuesto por una solución salina M9, un complemento de vitaminas, algunos oligoelementos,  $^{15}\text{NH}_4\text{Cl}$  como fuente de nitrógeno y  $^{12}\text{C}$ - ó  $^{13}\text{C}$ -glucosa como fuente de carbono (dependiendo si se expresa una muestra de proteína sencilla o doblemente etiquetada). Entonces, las células se incubaron en el medio de cultivo nuevo durante 1 h a 37 °C, se indujo su expresión con 1mM IPTG y se incubaron

durante 4-5 horas a 20 °C. Después, las células fueron recogidas por centrifugación y resuspendidas en tampón A.

**Tabla 6.2. Composición en 1 L de medio mínimo M9.**

Medio salino M9	Vitaminas y oligoelementos (concentración final)	Fuentes de nitrógeno y carbono
3 g KH <sub>2</sub> PO <sub>4</sub>	0.001 g mL <sup>-1</sup> Tiamina	1g <sup>15</sup> NH <sub>4</sub> Cl
6.8 g Na <sub>2</sub> HPO <sub>4</sub>	2 mM MgSO <sub>4</sub>	4g <sup>13</sup> C-glucosa
0.5 g NaCl	0.1 mM CaCl <sub>2</sub>	ó
	0.4 g mL <sup>-1</sup> FeCl <sub>3</sub>	4g <sup>12</sup> C-glucosa

En esta ocasión, las células fueron lisadas mediante el sistema de ultrafiltración Emulsiflex-C50 (Avestin, Canadá) que presenta múltiples ventajas frente a otros métodos, como son: una fácil limpieza, con lo que se evitan posibles contaminaciones; un preciso control tanto de presión (en este caso, entre 15000 y 20000 psi) como de temperatura (4 °C), con lo que se evitan posibles efectos de desnaturalización de la proteína sobre-expresada; y además, baja fragmentación del ADN, lo que facilita posteriores pasos de la purificación. Posteriormente, se procedió a la purificación siguiendo el mismo protocolo descrito en el Apartado 6.1.1, con las siguientes excepciones: en lugar de tampón A, en todos los pasos se utilizó el tampón 0.5 M NaCl, 50 mM Tris, pH 8.0 (Tampón B); para la cromatografía de afinidad a Níquel se utilizó una columna HisTrap FF de 5 mL en un sistema de FPLC (Fast Protein Liquid Chromatography) (ambos de AMERSHAM BIOSCIENCES) equilibrada en tampón B + 20 mM imidazol; y la elución de la proteína se realizó mediante un gradiente de concentración de imidazol desde 20 mM hasta 500 mM en tampón B. El rendimiento final aproximado para las muestras de hNedd-WW3 enriquecidas en isótopos fue de unos 4 mg de proteína por cada litro de medio de cultivo LB inicial, algo menor que el obtenido para las proteínas no marcadas isotópicamente. Finalmente, las fracciones de proteína purificada se concentraron según lo requerido en cada experimento de RMN y se almacenaron a 4 °C mientras no fue necesaria su utilización.

## 6.2. Preparación de muestras

### 6.2.1. Preparación de disoluciones de proteína

Las soluciones tampón se prepararon a partir de disoluciones de las formas ácida y básica del sistema tamponante en el 80 % del volumen final. Tras medir el valor de pH, se enrasó al volumen final y se volvió a comprobar nuevamente el pH de la disolución. Alternativamente, el tampón se preparó a partir de su forma ácida o básica, ajustando el pH al valor deseado añadiendo la cantidad adecuada de una disolución de hidróxido sódico o ácido clorhídrico. Tal procedimiento se siguió, por ejemplo, para preparar los tampones glicina, MES, Tris-HCl e Imidazol.

Las alícuotas de proteína descongeladas se dializaron frente a un volumen de tampón unas 500 o 600 veces mayor que el de la muestra, para lo que se introdujeron en una membrana de diálisis de SPECTRA-POR con un tamaño de poro de 1.0 KDa durante al menos el tiempo mínimo necesario para alcanzar el equilibrio de la diálisis (unas 8 horas) a 4° C con agitación suave. Para asegurar que la muestra se encuentra en las condiciones experimentales deseadas, se realizaron sucesivas etapas de diálisis (al menos 2) con tampón fresco. La disolución dializada de proteína se extrajo de la bolsa de diálisis y se filtró con un filtro de nitrocelulosa de tamaño de poro de 0.45 µm (SARSTEDT FiltroPur).

Con el fin de minimizar posibles diferencias en la composición de las disoluciones de proteína y ligando (ver Apartado 6.2.2), para los experimentos de ITC, las disoluciones de proteína se prepararon a una concentración de proteína superior al doble de la necesaria en el experimento; de forma análoga, el tampón de diálisis también contenía el doble de concentración sales del sistema tamponante (2x) deseadas en el experimento. Las disoluciones finales a la concentración deseada para cada experimento se prepararon añadiendo las cantidades necesarias proteína dializada, de agua Milli-Q y de tampón (2x) del último cambio de diálisis (prácticamente sin necesidad de corregir posteriormente el pH).

La determinación de la concentración de las diferentes proteínas se llevó a cabo midiendo la absorbancia a 280 nm y empleando coeficientes de extinción

recogidos en la Tabla 6.1, que fueron calculados empleando el método de Gill y von Hippel (Gill & von Hippel 1989).

### **6.2.2. Preparación de las disoluciones de ligando**

Los ligandos peptídicos fueron adquiridos en las casas comerciales PEPTIDE 2.0 (Chantilly, EEUU) y SYNBIOSCI (Livermore, EEUU). En todos los casos, los péptidos estaban acetilados y amidados en sus extremos carboxilo y amino terminal, respectivamente; y su pureza era superior al 98 %.

Las disoluciones de ligando para los experimentos de fluorescencia se prepararon en el momento de la realización del experimento pesando la cantidad necesaria de péptido liofilizado y disolviéndola en un volumen adecuado de tampón de diálisis, corrigiendo posteriormente el pH. Dado que los experimentos de ITC requieren disoluciones de ligando bastante concentradas (unos  $5 \text{ mg} \cdot \text{mL}^{-1}$ ), éstas se prepararon a partir de una disolución madre en agua Milli-Q de alta concentración, que generalmente se encuentra a un pH ácido debido las sales que acompañan a los péptidos comerciales. Las disoluciones finales a la concentración deseada para cada experimento se prepararon añadiendo las cantidades necesarias de la disolución madre, de agua y de tampón (2x) del último cambio de diálisis, como se ha mencionado en el apartado anterior. En el caso de algunos ligandos fue necesario corregir posteriormente el pH. Por último, en todos los casos, se comprobó espectrofotométricamente la concentración final de la disolución de péptido.

Todos los ligandos utilizados contenían al menos un resto de tirosina, por lo que sus concentraciones se determinaron, bien mediante medidas de absorbancia a 276 nm considerando un coeficiente de extinción molar de  $1450 \text{ M}^{-1} \text{ cm}^{-1}$  para cada aminoácido de tirosina en su secuencia, o bien mediante medidas de absorbancia a 280 nm para los péptidos **Rabia** y **Ptch1b**, que contienen un resto triptófano en su secuencia. Estos coeficientes de absorción y pesos moleculares de los ligandos se resumen en la Tabla 6.3.

**Tabla 6.3. Coeficientes de extinción molar a 276 nm y pesos moleculares determinados para todos los distintos ligandos.**

Ligandos	Coeficiente de extinción molar (M <sup>-1</sup> cm <sup>-1</sup> )	Peso molecular (g mol <sup>-1</sup> )
p53bp2 (EYPPYPPPYPSG)	4350	1501.2
PPPY (PPPY)	1450	510.0
HTLV1-a (PPYVEPTAP)	1450	1107.1
HTLV1-b (SDPQIPPPYVEP)	1450	1379.5
Ébola-a (ILPTAPPEYME)	1450	1302.1
Ébola-b (ILPTAPPEY)	1450	1041.1
Marburgo (MQYLNPPPYADH)	2900	1465.6
Rabia (DLWLPPPEYVPL)	6990*	1478.6
Lats1a (NRQPPPYPLTA)	1450	1392.0
Lats1b (YQGPPPYPKHL)	2900	1436.0
Lats2 (RRCPPPYPKHL)	1450	1503.0
Ptch1a (RYSPPPYSSHS)	2900	1415.5
Ptch1b (EGLWPPPYRPRR)	6990*	1563.7

\*Son coeficientes de extinción molar a 280 nm determinados teóricamente en base a la secuencia por el herramienta informática ProtParam (<http://expasy.org/tools/protparam.html>)

### 6.2.3. Preparación de disoluciones de agente desnaturizante (urea)

Las disoluciones de urea empleadas en los experimentos de desplegamiento térmico en presencia de agente desnaturizante se prepararon a partir de una disolución madre concentrada de urea 9.8 M en tampón fosfato sódico 20 mM, pH 7.0. Para ello, se pesó la cantidad necesaria de urea que se disolvió, debido al aumento de volumen que tiene lugar durante la solubilización, mediante sucesivas adiciones sobre un pequeño volumen de agua Milli-Q hasta alcanzar aproximadamente el 80 % del volumen final deseado y calentando ligeramente para acelerar la solubilización, dado que estamos muy cerca de la saturación (10.49 M a 25 °C) (Pace 1986). Una vez comprobada la concentración, procedimos a purificar la disolución de urea de los iones cianato y amonio, fruto de

su descomposición, añadiendo una pequeña cantidad de una resina de intercambio iónico, AG501-X8(D) (Biorad), que se separó por filtración después de 30 minutos. A continuación se añadieron las sales correspondientes para la preparación del tampón y se confirmó el pH. Entonces se enrasó al volumen final, se filtró la disolución resultante con un filtro de 0.45  $\mu\text{m}$  y se comprobaron de nuevo los valores de pH y concentración de urea finales.

La concentración exacta de las disoluciones de urea se determinó a partir de la medida del índice de refracción a temperatura ambiente (Pace 1986), mediante un refractómetro de mano ATAGO modelo R5000. La diferencia entre el índice de refracción de la disolución de urea y el tampón,  $\Delta n = n_{\text{urea}} - n_{\text{tampón}}$ , se relaciona con la concentración de urea a través de la ecuación (Warren & Gordon 1966):

$$[\text{urea}](M) = 117.66(\Delta n) + 29.753(\Delta n)^2 + 185.56(\Delta n)^3 \quad (6.2.1)$$

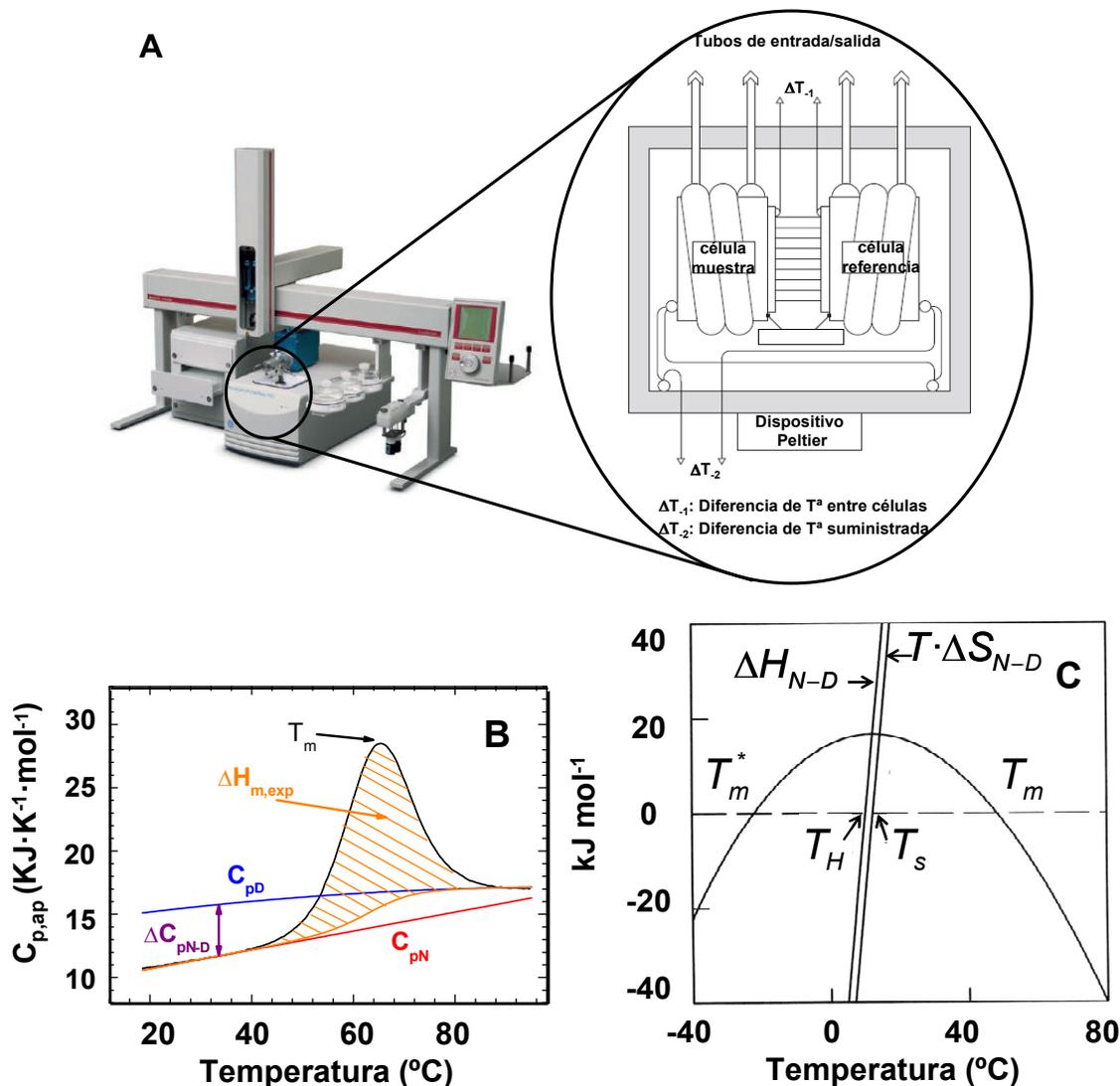
Las disoluciones finales a la concentración deseada, tanto de proteína como de urea, para cada experimento se prepararon añadiendo las cantidades necesarias de disolución de proteína dializada, disolución madre de urea y de tampón del último cambio de diálisis. Posteriormente se determinaron las concentraciones de urea y de proteína tal como se ha mencionado anteriormente.

### **6.3. Calorimetría diferencial de barrido (DSC)**

#### **6.3.1. Introducción**

La calorimetría diferencial de barrido (en inglés "Differential Scanning Calorimetry", DSC) es una técnica ampliamente utilizada para estudiar las transiciones de fase inducidas térmicamente y los cambios conformacionales en los sistemas biológicos, como las proteínas, ácidos nucleicos y las membranas lipídicas (Sanchez-Ruiz 1995; Cobos et al. 2001; Dragan & Privalov 2002; Garcia-Mira et al. 2002; Cobos et al. 2004; Ruiz-Sanz et al. 2004; Privalov & Dragan 2007; Streicher & Makhatadze 2007; Candel et al. 2008; Varela et al. 2009; Farber et al. 2010; Morel et al. 2010). Esta técnica mide la diferencia en la capacidad calorífica entre la muestra y las soluciones de referencia en función de la temperatura. A

modo de ejemplo, en el Panel A de la Figura 6.2, se ilustra un esquema de un bloque calorimétrico con los principales componentes de estos calorímetros.



**Figura 6.2. Fundamentos de la Calorimetría Diferencial de Barrido (DSC).** **A)** Sistema automatizado de un calorímetro capilar VP-DSC de Microcal LLC y un diagrama esquemático de su bloque calorimétrico. **B)** En trazo continuo negro se representa, a modo de ejemplo, la curva de capacidad calorífica en función de la temperatura del dominio N-terminal del repesor R69 del fago 434 en acetato sódico 20 mM a pH 4.5 (Ruiz-Sanz et al. 1999).  $C_{p,N}$  y  $C_{p,D}$  representan las capacidades caloríficas de los estados nativo y desnaturalizado en rojo y azul, respectivamente. La  $T_m$  corresponde a la temperatura de desnaturalización. La curva sigmoidal en color naranja representa la función de la capacidad calorífica interna (ecuación 6.4.13). El área bajo la transición, con rayas naranjas, corresponde al cambio total de entalpía para la temperatura de desnaturalización,  $\Delta H_{m,exp}$  o entalpía calorimétrica. **C)** Dependencia de las variaciones de entalpía, entropía y energía de Gibbs con la temperatura para una proteína hipotética. En él se muestran las temperaturas de desnaturalización a alta y baja temperatura,  $T_m$  y  $T_m^*$  respectivamente, así como  $T_S$  y la temperatura de inversión de entalpía  $T_H$ .

La capacidad calorífica se obtiene mediante la medida de la potencia necesaria para mantener la solución de la muestra a la misma temperatura que la de la solución de referencia ( $\Delta T_{-1}$ ), mientras la temperatura de ambos se aumenta de forma lineal. Así, a medida que se produce el cambio de temperatura durante una calorimetría de barrido, la proteína en la solución de la muestra sufre una transición de una manera cooperativa. Esta transición se relaciona con la destrucción de numerosas pequeñas fuerzas que estabilizan la estructura nativa. Dado que este proceso suele ser endotérmico, los cambios observados en la entalpía del sistema, como resultado de dichas perturbaciones, dan lugar a un descenso en la temperatura. Por lo tanto, el calorímetro proporcionará energía a la muestra para mantener su temperatura en el mismo valor que el de la solución de referencia. Esta energía se mide como potencia suministrada y produce la salida de un conjunto de datos brutos de potencia en función de la temperatura. En los calorímetros modernos la diferencia de potencial se convierte, mediante calibraciones internas, en los valores molares parciales capacidad calorífica. El posterior análisis del termograma obtenido utilizando modelos basados en la termodinámica del equilibrio proporciona la caracterización termodinámica completa del proceso de desplegamiento térmico de la macromolécula.

El modelo de equilibrio más simple que existe para el análisis del desplegamiento térmico de proteínas es el modelo de dos estados, el cual supone que las proteínas sólo pueden encontrarse en estado nativo o en estado desplegado (ver Apéndice I). Otros modelos pueden suponer la existencia de intermedios de plegamiento significativamente poblados y/o estados con diferentes grados de asociación (Luque et al. 2002; Martínez et al. 2011). En el Panel B de la Figura 6.2, se ha representado un ejemplo de una endoterma del desplegamiento térmico de una proteína con los principales parámetros que pueden determinarse a partir de un análisis calorimétrico. La temperatura del máximo del pico de la transición,  $T_m$ , corresponde a la temperatura donde el 50% de las moléculas de proteína se encuentran desnaturalizadas en el caso de una transición de dos estados sin ningún tipo de proceso acoplado de asociación o de disociación. El calor de la transición o entalpía calorimétrica correspondiente a la temperatura  $T_m$ ,  $\Delta H_{m,exp}$ , es el área de la curva de capacidad calorífica. Las líneas de base el pre- y

post- transición son las funciones de las capacidades caloríficas dependientes de la temperatura para el estado desplegado,  $C_{p,D}$ , y el estado nativo,  $C_{p,N}$ , de la proteína. La diferencia entre estas funciones,  $\Delta C_{p,N-D}$ , es el cambio de capacidad calorífica asociado a la desnaturalización de la proteína. A partir de estos parámetros obtenidos del análisis se completa la caracterización termodinámica del desplegamiento mediante distintas relaciones matemáticas, de acuerdo con el modelo en equilibrio utilizado para el análisis, como por ejemplo, las representadas en el Panel C de la Figura 6.2: el cambio de entalpía,  $\Delta H_{N-D}(T)$ , de entropía,  $\Delta S_{N-D}(T)$ , la energía de Gibbs,  $\Delta G_{N-D}(T)$ , la temperatura de máxima estabilidad,  $T_S$ , la temperatura de inversión de la entalpía,  $T_H$ , etc.

Una descripción más detallada de los fundamentos y características de esta técnica así como de los modelos termodinámicos más comunes se puede encontrar en algunas tesis doctorales desarrolladas en nuestro grupo de investigación (Cobos 2002; Casares 2003; Candel 2008b), y en numerosas referencias bibliográficas (Sturtevant 1977; Privalov 1979; Privalov 1982; Mateo et al. 1984; Cooper & Johnson 1994; Freire 1995; Sanchez-Ruiz 1995; Martinez et al. 2011).

### 6.3.2. Realización del experimento calorimétrico

Para la realización de todos los experimentos de DSC se han utilizado dos calorímetros VP-DSC (Microcal LLC., Northampton, Massachusetts) (Plotnikov et al. 1997) con diferente arquitectura de células calorimétricas: uno con células de moneda de 0.524 mL y otro con células capilares en espiral de 137  $\mu$ L (Figura 6.2A). El experimento calorimétrico se inicia llenando ambas células con el tampón filtrado y desgasificado procedente del último cambio de la diálisis. Entonces, se registran varios barridos de temperatura sucesivos que deben ser reproducibles y el último de ellos se guarda como línea base instrumental. Después, se llena la célula de muestra con la disolución de proteína desgasificada (en la célula de referencia se mantiene la disolución de tampón), y se realiza un primer barrido de temperatura de la muestra manteniendo las mismas condiciones experimentales. Posteriormente, se realiza un segundo barrido de temperatura de la muestra sin extraer la proteína del calorímetro para comprobar la reversibilidad del proceso de

desplegamiento. La reversibilidad depende de la temperatura final de calentamiento. La transición se considera que es reversible cuando, después de calentar la muestra pasando del máximo de la transición en dos barridos sucesivos, se recupera al menos un 70% de la endoterma inicial.

### 6.3.3. Cálculo de la capacidad calorífica molar parcial de la proteína

Como se ha descrito anteriormente, los calorímetros modernos generan archivos de datos experimentales como curvas de capacidad calorífica aparente de exceso en función de la temperatura. Sin embargo, con el fin de obtener los parámetros correctos de la evolución del proceso termodinámico ( $T_m$ ,  $\Delta H_{m,exp}$ ,  $\Delta C_{p,N-D}$ , etc) es necesario realizar un tratamiento previo y las correcciones de estos archivos de datos. El primer paso es la sustracción de archivos de datos de la línea base instrumental al primer y segundo barridos de temperatura de la muestra, por lo que se obtiene la dependencia de la capacidad calorífica de la proteína con la temperatura, y las posibles contribuciones instrumentales a la capacidad calorífica se eliminan. Además, la medida se ve afectada por un tiempo de respuesta conocido del bloque térmico del instrumento ( $\tau$ ), al no ser instantánea la transferencia de calor que se produce en las células del calorimétrico. La brecha temporal entre el calor producido y la señal registrada afecta y distorsiona los termogramas obtenidos durante el experimento. Esta distorsión es mayor a medida que el valor de la velocidad de barrido se incrementa. Por tanto, es importante corregir el efecto del tiempo de respuesta del calorímetro (Lopez Mayorga & Freire 1987):

$$C_{p,app,corr} = C_{p,app} + \tau \cdot \nu \cdot \frac{dC_{p,ap}}{dT} \quad (6.3.1)$$

donde  $\nu$  es la velocidad de barrido de temperatura.

A continuación, se expresan los datos brutos (potencia eléctrica frente a temperatura) en términos de capacidad calorífica molar mediante la siguiente expresión:

$$C_{p,ap} \left( \frac{kJ}{K \cdot mol} \right) = C_{p,ap} \left( \frac{mcal}{min} \right) \frac{4.186 \times 10^{-6} \left( \frac{kJ}{mcal} \right)}{\nu (K/min) \cdot V(L) \cdot c(M)} \quad (6.3.2)$$

donde  $\nu$  es la velocidad de barrido de temperatura,  $c$  la concentración de proteínas de la solución de la muestra del experimento y  $V$  el volumen del calorímetro de la célula. Hay una diferencia de capacidad calorífica aparente entre la línea de base instrumental y el barrido de temperatura de la muestra, debido a las moléculas del disolvente reemplazadas por moléculas de proteína en la célula de la muestra del calorímetro al realizar el experimento. La capacidad calorífica de estas moléculas de disolvente desplazadas presentan un valor mucho mayor que el de las moléculas de la muestra y por tanto esta diferencia generalmente es negativa. Para corregir este efecto en los datos obtenidos de  $C_{p,ap}$  debidos a la sustracción de referencia, se obtiene la capacidad calorífica molar parcial,  $C_{p,p}$ :

$$C_{p,p} = C_{p,ap} + \Delta C_{p,ap} \quad (6.3.3)$$

La diferencia de capacidad calorífica aparente,  $\Delta C_{p,ap}$ , se puede expresar como:

$$\Delta C_{p,ap} = \Delta m_s \cdot C_{p,s} = \left( m_p \cdot \frac{\nu_p}{\nu_s} \right) \cdot C_{p,s} \quad (6.3.4)$$

donde  $\Delta m_s$  son las moléculas del disolvente sustituidas,  $C_{p,s}$  la capacidad calorífica del disolvente,  $m_p$  la masa de la proteína en la celda calorimétrica y,  $\nu_p$  y  $\nu_s$  los volúmenes parciales específicos del disolvente y la proteína, respectivamente. Mientras que el disolvente sea un tampón acuoso diluido, se pueden utilizar los valores del agua pura  $1 \text{ mL} \cdot \text{g}^{-1}$  y  $4.186 \text{ J} \cdot \text{K}^{-1} \cdot \text{g}^{-1}$ , para el  $\nu_s$  y la  $C_{p,s}$ , respectivamente (Privalov & Potekhin 1986), y un valor de  $0,73 \text{ g} \cdot \text{mL}^{-1}$  se utiliza para  $\nu_p$  como promedio supone para las proteínas globulares (Privalov & Potekhin 1986; Makhatadze & Privalov 1990; Privalov & Makhatadze 1990). Por lo tanto, para convertir el  $C_{p,ap}$  obtenido de la resta de referencia en  $C_{p,p}$  la transformación que se utiliza es:

$$C_{p,p} (\text{kJ}/\text{K} \cdot \text{mol}) = C_{p,ap} (\text{kJ}/\text{K} \cdot \text{mol}) + 0.73 \cdot P_m (\text{g} \cdot \text{mol}^{-1}) \cdot 4.184 \cdot 10^{-3} (\text{kJ} \cdot \text{K}^{-1} \cdot \text{g}^{-1}) \quad (6.3.5)$$

donde  $P_m$  es el peso molecular de la proteína.

Todas las operaciones de normalización y tratamiento previo de los datos brutos del DSC hasta obtenerlos en unidades de capacidad calorífica molar parcial han sido realizadas mediante el programa *Origin DSC Data Analysis de OriginLab* (Microcal LLC., Northampton, Massachusetts).

#### 6.3.4. Análisis de las trazas calorimétricas

Como se ha mencionado anteriormente, del análisis de la capacidad calorífica molar parcial de la proteína en función de la temperatura de acuerdo con un modelo de equilibrio adecuado, se obtienen los parámetros termodinámicos del proceso de desnaturalización térmica. El modelo de equilibrio de dos estados es el modelo más sencillo que se puede utilizar, según el cual, durante el proceso de desplegamiento de la proteína sólo existen dos estados poblados en el equilibrio, que son el estado nativo (N) y el estado desnaturalizado o desplegado (D). En el Apéndice I se describe el desarrollo matemático completo de todas las ecuaciones utilizadas en los análisis realizados con este modelo. Así, las curvas experimentales de  $C_p$  se ajustaron mediante el método de mínimos cuadrados no lineales a la ecuación I.X del Apéndice I, cuya rutina ha sido implementada en el programa *SigmaPlot (SPSS Inc)*:

$$C_p = C_{p,N} + \frac{(\Delta H_D)^2}{RT^2} \frac{K_D}{(1+K_D)^2} + x_D \Delta C_p \quad (\text{I.X ó 6.3.6})$$

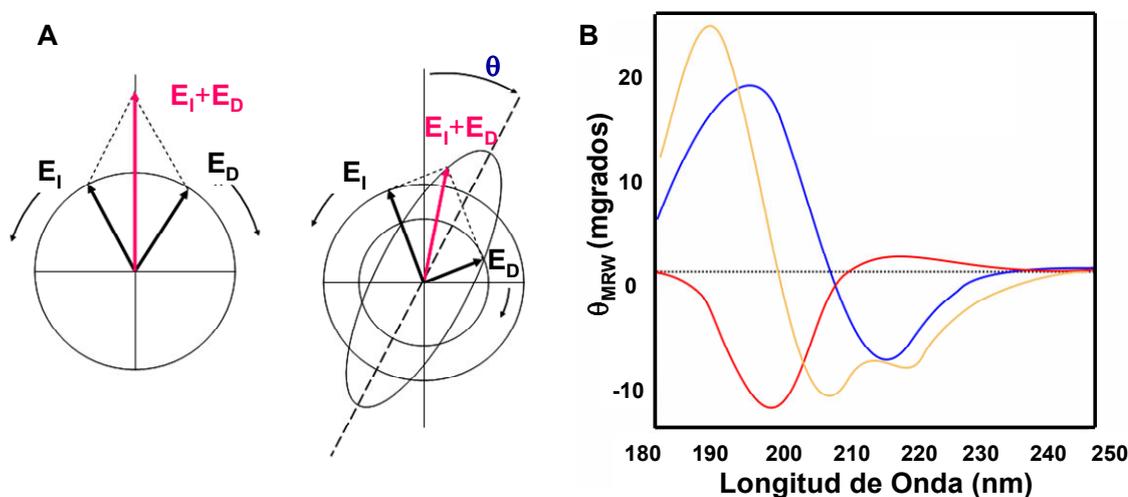
### 6.4. Dicroísmo circular

#### 6.4.1. Introducción

La espectroscopía 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 (Padmanabhan et al. 1999; Cobos et al. 2002; Martín-Sierra et al. 2003; Oliva & Muñoz 2004). Un rayo de luz polarizado en un plano puede considerarse formado por dos componentes polarizados circularmente, uno a la derecha y el otro a la izquierda. Estos componentes están en fase y son de la misma amplitud. Al pasar por un medio ópticamente activo,

cada componente interactúa de manera diferente con los centros quirales de las moléculas presentes. La interacción de la radiación con la muestra induce un desfase y un cambio de magnitud diferenciales en ambas componentes circularmente polarizadas de la luz, y estos fenómenos provocan una rotación del plano de polarización en un ángulo  $\theta$  y la distorsión de este plano genera una elipse como se representa en el Panel A la Figura 6.3. La rotación del plano y la diferente absorción de los componentes circularmente polarizados varían de acuerdo con la longitud de onda, pudiéndose obtener espectros de estos fenómenos, esto es, gráficas de la rotación o elipticidad,  $\theta$ , frente a la longitud de onda.

Los espectros de dicroísmo circular se obtienen generalmente en las regiones del ultravioleta cercano (250 a 350 nm) y lejano (180 a 250 nm) de la radiación electromagnética. En la región del ultravioleta cercano, los cromóforos más importantes son los grupos aromáticos de las cadenas laterales de triptófano, tirosina, y fenilalanina. Ya que la asimetría de estos grupos químicos se debe exclusivamente a su entorno y como los residuos aromáticos generalmente se encuentran distribuidos en toda la macromolécula, los espectros en esta región son extremadamente sensibles a los cambios en la conformación global de la proteína. Los espectros de dicroísmo en la región del ultravioleta lejano, se deben principalmente a los enlaces amida que unen los residuos de los aminoácidos entre sí. La asimetría de estos cromóforos se debe a la distribución espacial de la cadena principal de la proteína (ángulos diedros  $\phi$  y  $\psi$ ), por lo tanto, las señales de dicroísmo circular se pueden interpretar en términos del contenido de estructura secundaria presente, es decir, se pueden relacionar con el porcentaje de residuos que se encuentran en las diferentes conformaciones estructurales: hélices  $\alpha$ , láminas  $\beta$ , giros  $\beta$  o estructuras desordenadas (Compton & Johnson 1986; Johnson 1990). En el Panel B Figura 6.3, se representan los espectros característicos de cada una de estas conformaciones, y se pueden observar importantes diferencias entre ellos.



**Figura 6.3. Fundamentos de Dicroísmo Circular (CD).** **A)** La luz linealmente polarizada se representa como la superposición ( $E_L + E_R$ ) de luz polarizada circularmente a izquierdas ( $E_L$ ) y derechas ( $E_R$ ) de igual amplitud y fase. Cuando se produce una absorción diferencial de sus componentes polarizadas se produce la rotación óptica o elipticidad que se define en base al ángulo,  $\theta$ , resultante. **B)** Representación conjunta de los espectros correspondientes a cuatro situaciones de estructura secundaria, en el intervalo de longitud de onda comprendido entre 180 y 250 nm: Hélice  $\alpha$  en verde, lámina  $\beta$  en azul y estructura de ovillo-desordenado o “random-coil” en rojo.

#### 6.4.2. Obtención de espectros de dicroísmo circular

Los experimentos de CD se llevaron a cabo en un espectropolarímetro JASCO modelo J-715 equipado con un control de temperatura del compartimento de células (PTC-348WI).

Los espectros de CD en la región del ultravioleta lejano se registraron entre 190 nm y 260 nm a una velocidad de barrido de  $50 \text{ nm min}^{-1}$ , usando una cubeta de 1 mm de paso de luz y una concentración de proteína en torno a  $0.3 \text{ mg mL}^{-1}$  ( $50 \text{ }\mu\text{M}$ ). Los espectros se realizaron a diferentes temperaturas, en un intervalo desde  $2 \text{ }^\circ\text{C}$  hasta  $98 \text{ }^\circ\text{C}$ , generalmente obteniendo un espectro cada  $10 \text{ }^\circ\text{C}$  y otro después del calentamiento. Antes de registrar cada espectro se dejó equilibrar la muestra a la temperatura deseada al menos 5 minutos. Los espectros en la región del UV-cercano se registraron entre 350 nm y 250 nm a una velocidad de barrido de  $50 \text{ nm}\cdot\text{min}^{-1}$ , usando una cubeta de 5 mm de paso de luz y una concentración de proteína en torno a  $1 \text{ mg mL}^{-1}$  ( $180 \text{ }\mu\text{M}$ ).

Para eliminar tanto las contribuciones de tampón como las contribuciones instrumentales se recogió también un espectro con el tampón del último cambio de diálisis para cada condición experimental que se restó al espectro de proteína. En todos los casos se acumularon un total de 10 espectros consecutivos, número suficiente, en nuestro caso, para obtener una aceptable relación señal/ruido.

En todos los casos, las medidas de elipticidad se han normalizado a unidades de elipticidad molar media por residuo ( $\theta_{MRW}$ ) ( $\text{grados} \cdot \text{dmol}^{-1} \cdot \text{cm}^2$ ), para lo cual se ha utilizado la siguiente ecuación:

$$\theta_{MRW} = \frac{\theta}{\frac{C}{P_m} \cdot l \cdot 10 \cdot n} \quad (6.4.1)$$

donde  $\theta$  es la elipticidad medida experimentalmente en milésimas de grado,  $C$  es la concentración de la proteína en  $\text{mg} \cdot \text{mL}^{-1}$ ,  $l$  es el paso de luz de la célula en  $\text{cm}$ ,  $P_m$  es la masa molecular de la proteína en  $\text{g} \cdot \text{mol}^{-1}$  y  $n$  es el número de residuos de la proteína. El factor de 10 es el resultado de convertir las unidades de concentración de  $\text{mg} \cdot \text{mL}^{-1}$  a  $\text{dmol} \cdot \text{cm}^{-3}$  y la elipticidad a grados.

### 6.4.3. Estudio del desplegamiento térmico de dominios WW

El estudio del desplegamiento térmico mediante CD de los dominios hNedd4-WW4, hYap-WW1(L30K), hFbp11-WW1 y hFbp11-WW2 se ha llevado a cabo, por una parte a diferentes condiciones de pH y, por otra, en presencia de urea, registrando los cambios de elipticidad observados en la región del ultravioleta lejano, a una longitud de onda de 230 nm, variando linealmente la temperatura de la muestra en un intervalo de 2 a 98 °C. El desplegamiento del dominio hNedd4-WW4, además, se ha monitorizado en esta región a 203 nm, y en el ultravioleta cercano a 260 nm. En todos los casos la velocidad de barrido de temperatura empleada ha sido de  $1.5 \text{ }^\circ\text{C} \cdot \text{min}^{-1}$ .

El procedimiento experimental fue el mismo para todo el estudio (tanto muestras de proteína como blancos con el tampón correspondiente). En primer lugar, se disminuyó la temperatura hasta 2 °C en el instrumento registrando un espectro previo antes de comenzar el barrido de temperatura hasta 98 °C, temperatura a la que se registró otro espectro. A continuación, se enfrió

rápidamente hasta 2 °C para reducir la irreversibilidad causada por los fenómenos de oligomerización que tienen lugar a alta temperatura una vez se encuentra la proteína en el estado desnaturalizado. En estas condiciones, se registró entonces un espectro para comprobar la reversibilidad del proceso de desplegamiento, mediante la comparación de éste con el espectro registrado antes de comenzar el calentamiento. La señal del desplegamiento térmico seguido mediante CD se expresó como la diferencia entre los barridos temperatura de la muestra de proteína y del blanco correspondiente. En todos los casos, la temperatura se determinó mediante una sonda de temperatura en el mismo bloque térmico donde se situó la cubeta de cuarzo, muy próxima a ella, debido a que el estrecho paso de luz requerido para estos experimentos no permitía introducir la sonda directamente en la cubeta.

#### 6.4.4. Análisis de las curvas de desplegamiento térmico

En el apéndice II se ha realizado el desarrollo matemático necesario para seguir por CD la desnaturalización térmica de acuerdo con el modelo de equilibrio de dos estados. El análisis de las curvas de CD se realizó mediante el ajuste por mínimos cuadrados no lineales de acuerdo con este modelo. Para ello, se ha utilizado la ecuación II.IX, implementada en el programa SigmaPlot versión 6.10 (SPSS Inc).

$$\theta = \frac{[\theta_{N,0} + m_N(T - T_r)] + [\theta_{D,0} + m_D(T - T_r)] \exp\left\{-\left(\Delta H_m \left(1 - \frac{1}{T_m}\right) - \Delta C_{p,D} \left[T_m - T + T \ln\left(\frac{T}{T_m}\right)\right]\right) / R \cdot T\right\}}{1 + \exp\left\{-\left(\Delta H_m \left(1 - \frac{1}{T_m}\right) - \Delta C_{p,D} \left[T_m - T + T \ln\left(\frac{T}{T_m}\right)\right]\right) / R \cdot T\right\}} \quad (\text{II.IX ó 6.4.2})$$

### 6.5. Espectroscopía de fluorescencia

#### 6.5.1. Introducción

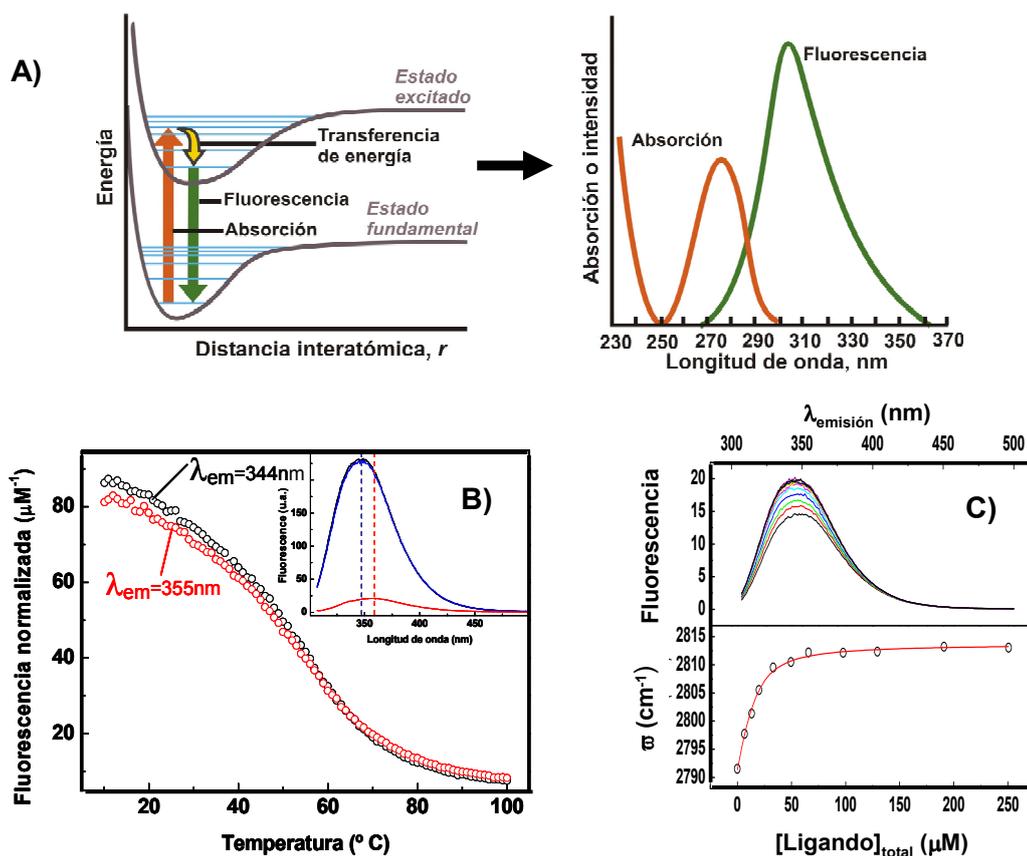
La fluorescencia molecular es un proceso en el cual las moléculas son excitadas por la absorción de radiación electromagnética. Como se muestra en la Figura 6.4A, las especies excitadas se relajan al estado fundamental, liberando su exceso de energía en forma de fotones, en lo que se denomina emisión de

fluorescencia. Una vez excitada, la molécula pasa a un modo de menos energía dentro del estado excitado (el modo vibracional) debido a un proceso de transferencia de energía. Desde este nivel se produce la emisión de un fotón que devuelve a la molécula al estado fundamental. Este fotón emitido por fluorescencia tiene menos energía que el absorbido, y por tanto, posee una frecuencia menor, es decir se emite a mayor longitud de onda (Figura 6.4A). En muestras biológicas la molécula fluorescente (fluoróforo) está normalmente en disolución y puede interactuar con otras moléculas y, como consecuencia, se pueden producir variaciones en la emisión fluorescente. En el caso de las proteínas, en general, los aminoácidos aromáticos son capaces de emitir fluorescencia cuando son excitados con luz ultravioleta y, en particular, el triptófano es el que tiene un rendimiento cuántico mayor.

La fluorescencia de proteínas puede emplearse para el estudio de procesos de unión de ligandos a las proteínas gracias a la alta sensibilidad del triptófano a los cambios en su entorno próximo, de tal manera que se pueden observar cambios en los espectros de emisión del triptófano. En algunas ocasiones, dichos ligandos poseen fluorescencia propia, que puede cambiar cuando se unen a una proteína (Daniel & Weber 1966; Anderson 1969; Anderson & Weber 1969; Condie & Quay 1983). Sin embargo, el proceso de unión de ligandos se puede seguir aún cuando éstos carezcan de fluorescencia propia, ya que la unión de ligandos a proteínas normalmente induce cambios en la estructura tridimensional de las mismas que, si suponen una alteración del entorno de algún fluoróforo, provocarán cambios en los espectros de fluorescencia. Todos estos cambios en la emisión de fluorescencia a una longitud de onda determinada, pueden emplearse para obtener información sobre la constante de unión.

En nuestro caso, las medidas de fluorescencia realizadas sirvieron para obtener una estimación inicial de la constante de disociación ( $K_d$ ) de los diferentes ligandos peptídicos a los diferentes dominios WW, lo que nos fue de gran utilidad para el diseño de los experimentos de titulación calorimétrica. En estos experimentos de fluorescencia, se siguieron las variaciones en la intensidad de fluorescencia y el desplazamiento del máximo de longitud de onda del espectro de fluorescencia. Por otro lado, durante la caracterización termodinámica del dominio

hNedd4-WW4 también se monitorizó el cambio en la fluorescencia durante la desnaturalización térmica de este dominio, como se ilustra en el Panel B de la Figura 6.4. En ambos casos, las medidas de fluorescencia se realizaron en un espectrofluorímetro Cary Eclipse (*Varian Inc*) equipado con un sistema *peltier* de control de temperatura utilizando una cubeta cuadrangular de 1 cm de paso de luz.



**Figura 6.4. Fundamentos de la espectroscopía de fluorescencia. A)** Fenómeno de la emisión de fluorescencia en proteínas. A la izquierda, diagrama de energía que representa la absorción de un fotón por una molécula en el estado fundamental y su posterior emisión de fluorescencia desde el estado excitado. A la derecha, los correspondientes espectros de absorción y emisión. **B)** Experimentos de desnaturalización térmica seguida por emisión fluorescencia a dos longitudes de onda diferentes: 344 nm (en círculos negros) y 355 nm (en círculos rojo). En el recuadro, los espectros de fluorescencia a diferentes temperaturas del dominio: a 10  $^{\circ}\text{C}$  antes de calentar (en línea negra), a 100  $^{\circ}\text{C}$  (en línea roja) y a 10  $^{\circ}\text{C}$  después de calentar (en línea azul). **C)** Experimento de titulación de una proteína con un ligando seguido por emisión fluorescencia. Arriba, se representan los espectros de emisión recogidos durante la titulación, y abajo, la curva de titulación resultante de la representación del centro de masas ( $\bar{m}$ ) de cada espectro frente a concentración total de ligando añadido. En rojo se muestra el ajuste de la curva de titulación a un modelo de unión.

### 6.5.2. Experimento de desplegamiento térmico seguido por fluorescencia

El estudio del desplegamiento térmico mediante fluorescencia del dominio hNedd4-WW4 se ha realizado únicamente a pH 7.0, de forma análoga a como se ha descrito para la desnaturalización térmica seguida por CD: se recogen espectros a 10 °C, a 100 °C y de nuevo a 10 °C después de enfriar una vez terminado el barrido de temperatura para comprobar la reversibilidad. El rango de temperatura escogido fue entre 10 y 100 °C, la velocidad de barrido fue de 1.5 °C min<sup>-1</sup> y los tiempos de espera antes de registrar cada espectro fueron de 5 minutos. Los espectros y el barrido de temperatura se recogieron tanto con la muestra a una concentración de 20 μM como con el tampón para luego restarlo (blanco). La excitación de la muestra se realizó a una longitud de onda de 298 nm y 5 nm de ancho de rendija. Los espectros se registraron con 5 nm de anchura de rendija a longitudes de onda comprendidas entre 305 y 500 nm y la desnaturalización térmica se siguió simultáneamente a dos longitudes diferentes, a 344 nm donde se encuentra el máximo de fluorescencia de la proteína nativa y a 355 nm donde se encuentra el máximo teórico para el grupo indol del triptófano en disolución. La intensidad de fluorescencia se normalizó por la concentración de proteína total en la cubeta ( $M_T$ ) y se representaron las curvas de desnaturalización seguidas a distintas longitudes de onda frente a la temperatura como se puede observar en la Figura 6.4 B. Ambas curvas se analizaron mediante un ajuste por mínimos cuadrados de acuerdo con el modelo de dos estados de forma análoga a los experimentos de desnaturalización térmica seguidos por CD, haciendo uso de la ecuación II.IX ó 6.4.2 del Apéndice II.

### 6.5.3. Experimento de titulación proteína-ligando mediante espectroscopía de fluorescencia

Para determinar las diferentes constantes de disociación ( $K_d$ ) se realizó una serie de experimentos de titulación por adición de volúmenes variables de disoluciones de los diferentes péptidos sobre disoluciones de los dominios WW.

Como se ha descrito en el Apartado 6.2, estas titulaciones implicaron la preparación de dos disoluciones, una disolución de proteína 20 μM y otra de ligando con una concentración final entre 1.2 mM y 4 mM, ambas en tampón

fosfato sódico 20 mM a pH 7.0. Inicialmente se introdujeron 120  $\mu\text{L}$  de la disolución de proteína en una cubeta de cuarzo para fluorescencia, de 3 mm de paso de luz. Después se atemperó la muestra a 25 °C durante 5 minutos en el instrumento, y se recogió un espectro de la proteína libre. Sobre esta disolución inicial se adicionaron progresivamente, mediante el uso de distintas jeringas Hamilton de alta precisión, volúmenes crecientes (entre 0.5 y 25  $\mu\text{L}$ ) de la disolución de ligando. Tras cada adición de ligando, la disolución de muestra se agitó para conseguir su homogeneización y posteriormente se registraron los espectros de emisión de fluorescencia entre 305 y 500 nm. Para el análisis de las curvas de fluorescencia es necesario tener en cuenta la fluorescencia intrínseca del ligando debida a la presencia de restos de tirosina en su secuencia. Para ello se registraron espectros correspondientes a las mismas adiciones de ligando sobre el mismo volumen inicial de disolución de tampón que se restaron a cada espectro equivalente de la titulación proteína-ligando. Todos los espectros se registraron a 25 °C, utilizando una longitud de onda de excitación de 298 nm, y anchos de rendijas de excitación y emisión de 5 nm.

#### 6.5.4. Análisis de un experimento de titulación mediante espectroscopía de fluorescencia

Una vez corregida la contribución del ligando libre, cada espectro se normalizó por la concentración de proteína total en la cubeta ( $M_T$ ), teniendo en cuenta la dilución de ésta tras cada adición de ligando. De este modo, la señal de intensidad de fluorescencia en cada medida del experimento de titulación viene dada por la ecuación:

$$F_{i,n} = \frac{F_i}{M_T} = F_M + (F_{ML} - F_M) \cdot \theta_i = F_M + \Delta F \cdot \Theta_i \quad (6.5.1)$$

donde  $F_{i,n}$  se refiere a la intensidad de fluorescencia corregida por la fluorescencia debida al ligando y normalizada por la concentración total de proteína;  $F_M$  a la intensidad de fluorescencia de la proteína libre;  $F_{ML}$  a la intensidad de fluorescencia del complejo ML, y  $\Theta_i$  a la fracción de saturación en cada punto de la titulación, que se define como:

$$\Theta_i = \frac{(K_d + M_T + L_T) \pm \sqrt{(K_d + M_T + L_T)^2 - 4 \cdot M_T \cdot L_T}}{2 \cdot M_T} \quad (6.5.2)$$

donde  $M_T$  y  $L_T$  son las concentraciones totales de proteína y ligando respectivamente.

Sustituyendo 6.5.2 en 6.5.1 se obtiene la ecuación correspondiente al modelo de unión a un único sitio, al cual se ajustaron los datos experimentales:

$$F_{i,n} = F_M + (F_{ML} - F_M) \cdot \frac{(M_T + L_T + K_d) - \sqrt{(K_d + M_T + L_T)^2 - 4 \cdot M_T \cdot L_T}}{2 \cdot M_T} \quad (6.5.3)$$

Además de producirse un cambio en la intensidad de fluorescencia se observó que para algunos ligandos tiene lugar un desplazamiento en la longitud de onda del máximo de emisión. En estos casos, para poder analizar adecuadamente los datos experimentales se utilizó el centro de masas espectral (Mohana-Borges et al. 1999), el cual se define como:

$$\varpi = \frac{\sum \varpi_i \cdot F_i}{\sum F_i} \quad (6.5.4)$$

donde  $\varpi$  es el centro de masas en números de onda ( $\text{cm}^{-1}$ ) y  $F_i$  es la fluorescencia emitida para cada número de onda  $\varpi_i$ . Empleando este método de análisis se consideran simultáneamente tanto los cambios en la intensidad como los desplazamientos en longitud de onda del máximo de emisión tras cada adición de ligando. La representación del centro de masas  $\varpi$  frente a la concentración de ligando total en la cubeta ( $L_T$ ) proporciona la curva de unión como la mostrada en el Panel C de la Figura 6.4, cuyo ajuste por regresión no lineal de mínimos cuadrados empleando la ecuación 6.5.3 (siendo  $\varpi \equiv F_{i,n}$ ), permite determinar el valor de la constante de disociación ( $K_d$ ).

## **6.6. Calorimetría isotérmica de titulación (ITC)**

### **6.6.1. Introducción**

La calorimetría isotérmica de titulación (Isothermal Titration Calorimetry, ITC) es una técnica que mide directamente el calor liberado o absorbido en cualquier proceso de interacción intermolecular, como es el caso de asociaciones proteína-ligando, proteína-proteína, etc (Ladbury & Chowdhry 1996). Un experimento de ITC es un experimento de valoración calorimétrica de un cierto volumen de disolución de uno de los reactivos (normalmente macromolécula) con cantidades controladas de otro reactivo (normalmente ligando) a presión y temperatura constantes. Por tanto, el calor medido directamente durante la valoración equivale a la entalpía de dicho proceso (Wiseman et al. 1989). La calorimetría isotérmica de titulación permite obtener, de forma relativamente sencilla, una caracterización termodinámica completa y precisa del proceso de unión. Así, si el efecto térmico es lo suficientemente grande y la magnitud de la constante de unión adecuada, un único experimento de ITC permite determinar la constante de equilibrio de unión,  $K$ , el cambio de entalpía aparente de unión,  $\Delta H_{app}$ , así como la estequiometría de la reacción,  $n$ . Además, si se realizan los experimentos a diferentes temperaturas es posible determinar el cambio en la capacidad calorífica del proceso,  $\Delta C_p$ .

Los calorímetros de titulación más comunes son adiabáticos y funcionan por compensación del efecto térmico que tiene lugar durante la adición de reactivo en la célula de muestra localizada en un entorno adiabático (Wiseman et al. 1989). En el Panel A de la Figura 6.5, se ilustra un diagrama general de uno de estos instrumentos, en donde se puede observar los dispositivos termoeléctricos que miden la diferencia de temperatura entre las células de muestra y de referencia ( $\Delta T_{-1}$ ) y entre cada célula y la camisa adiabática ( $\Delta T_{-2}$ ). A medida que se produce la reacción en la célula de muestra se disminuye el valor de  $\Delta T_{-1}$  hasta 0 mediante el calentamiento de la célula muestra (si la reacción es endotérmica) o de la célula de referencia (si es exotérmica). De esta manera, se obtiene un pico característico sobre la línea base de potencia estacionaria y la integral de este potencial

requerido para llevar  $\Delta T_{-1}$  a 0 en el tiempo necesario para recobrar el equilibrio es el calor correspondiente a cada inyección.

Normalmente, en un experimento de ITC, se realizan sucesivas adiciones hasta superar la saturación del equilibrio entre los reactantes de modo que las últimas adiciones no produzcan un calor significativo como se muestra en el Panel B (arriba) de la Figura 6.5. El termograma neto se obtiene de la integración de cada pico individualmente, estableciendo los límites de integración en la línea base que precede y sigue a dicho pico (Figura 6.5B), mediante la ecuación:

$$Q = \int_{t_i}^{t_f} W(t) \cdot dt \quad (6.6.1)$$

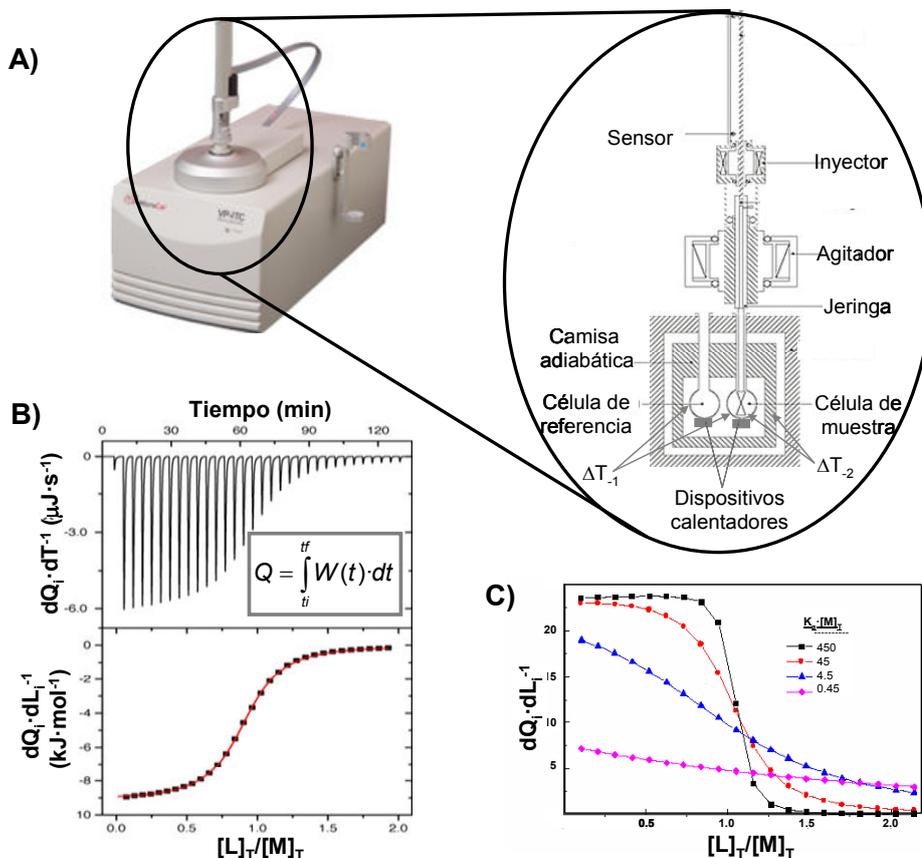
Los límites prácticos para una correcta determinación de los parámetros termodinámicos de unión utilizando esta técnica vienen determinados por el producto entre la constante de asociación,  $K_a$ , y la concentración total de macromolécula,  $[M]_T$ :  $1 < K_a [M]_T > 1000$  (Wiseman et al. 1989). En el Panel C de la Figura 6.5 se muestran las simulaciones de un experimento convencional de ITC con adiciones de idéntico volumen para las que se han considerado cuatro posibles valores de la constante de asociación,  $2.5 \times 10^6$ ,  $2.5 \times 10^5$ ,  $2.5 \times 10^4$  y  $2.5 \times 10^3 \text{ M}^{-1}$ , que abarcan un intervalo desde una afinidad relativamente alta hasta una afinidad baja (cuyo producto va  $K_a \cdot [M]_T$  desde 450 hasta 0.45). Como puede observarse cuando el producto  $K_a \cdot [M]_T$  está dentro del intervalo adecuado, es fácil obtener una isoterma de unión de forma sigmoideal, adecuada para realizar un análisis con unos errores estándar aceptables. En los casos en que el producto  $K_a \cdot [M]_T$  se aproxima a los límites se puede tratar de optimizar la isoterma haciendo algunas variaciones en el diseño del experimento en las concentraciones de macromolécula total o del ligando en la jeringa, así como en el perfil de volúmenes de inyección del ligando. Cuando el producto  $K \cdot [M]_T$  está fuera del intervalo adecuado, la solución puede ser realizar experimentos de titulación por desplazamiento, en los que el ligando problema compite por el mismo sitio de unión con otro cuya interacción ha sido caracterizada previamente (Zhang & Zhang 1998; Sigurskjold 2000; Velazquez-Campoy & Freire 2006).

### 6.6.2. Realización de los experimentos de ITC

Los experimentos se realizaron utilizando dos calorímetros isotérmicos de titulación MCS-ITC y VP-ITC (*MicroCal LLC.*, Northampton, MA). Las disoluciones de proteína y ligando preparadas según los Apartados 6.2.1 y 6.2.2 fueron filtradas, desgasificadas para evitar la formación de burbujas y equilibradas a la temperatura experimental. La mayoría de los experimentos se realizaron a  $25 \pm 0.1$  °C, aunque para el cálculo de la capacidad calorífica de unión también se registraron experimentos a 20 y 30 °C. El tiempo esperado para que la potencia suministrada por el calorímetro recupere la línea base entre cada adición fue de 450 s.

Con el fin de obtener una isoterma adecuada, previamente, diversas simulaciones (como las mostradas en la Figura 6.5B) a partir de las constantes de disociación obtenidas por titulaciones seguidas mediante fluorescencia (Apartado 6.5.4) han servido para optimizar las concentraciones de macromolécula total, de ligando en la jeringa o los perfiles de volúmenes de inyección. Las disoluciones de proteína en la célula de muestra con concentraciones entre 20  $\mu$ M y 60  $\mu$ M fueron tituladas con ligandos con concentraciones entre 0.6 mM y 4 mM, y utilizando generalmente un perfil de volúmenes de 25 a 28 inyecciones desde 4  $\mu$ L hasta 22  $\mu$ L para poder definir mejor la curva de titulación. En aquellos casos que el producto  $K_a \cdot [M]_T$  fue menor que 1, se realizaron experimentos de ITC por desplazamiento. Para ello, se introdujo una disolución de proteína saturada con el ligando problema (ambas concentraciones conocidas) en la célula de muestra y se realizó la titulación con un ligando de afinidad moderada-alta (el ligando p53bp2) que compitió por el sitio de unión de la macromolécula hasta que desplazó por completo el equilibrio del primer ligando y se alcanzó nuevamente la saturación.

Los termogramas correspondientes a los experimentos de dilución se obtuvieron titulando el ligando sobre la disolución tampón en ausencia de proteína utilizando el mismo perfil de volúmenes.



**Figura 6.5. Fundamentos de la Calorimetría Isoterma de Titulación (ITC).** **A)** Calorímetro Isotérmico de Titulación VP-ITC de Microcal LLC y un diagrama esquemático de su bloque calorimétrico con los principales componentes. **B)** Experimento de titulación de una proteína con un ligando seguido por ITC. Arriba, se muestra el termograma de titulación como calores por unidad de tiempo liberado tras cada inyección de ligando sobre la proteína. En el recuadro se muestra la ecuación utilizada para obtener el calor liberado tras cada inyección de ligando. Abajo, la isoterma de unión o dependencia del calor liberado en cada inyección frente al cociente concentración de ligando total/concentración de proteína total. Los cuadrados representan los datos experimentales y la línea el ajuste al modelo de "n" sitios de unión idénticos e independientes (ecuación III.X). **C)** Simulación del calor por mol de ligando añadido asociado a cada inyección para un experimento de ITC con los siguientes parámetros experimentales:  $V_C=1.347$  mL;  $[M]$  inicial en célula de  $1.8 \times 10^{-4}$  M; 20 inyecciones de 5  $\mu$ L;  $[L]$  en la jeringa de inyección de 5 mM; un solo sitio de unión para el ligando, n de 1 y cuatro posibles valores de la constante de asociación,  $2.5 \times 10^6$ ,  $2.5 \times 10^5$ ,  $2.5 \times 10^4$  y  $2.5 \times 10^3$   $M^{-1}$

### 6.6.3. Análisis de los experimentos de ITC

Una vez corregida la línea base, el termograma neto se integra para obtener los calores correspondientes de cada inyección de la titulación, los cuales se corrigen por los calores de dilución del ligando y se normalizan por la concentración de ligando total en la célula tras cada inyección. Estos valores del

calor transferido por mol de ligando añadido ( $dQ_i/dL_{T,i}$ ) representados frente a la fracción molar ( $[L]_T/[M]_T$ ) permiten obtener la isoterma de unión (Figura 6.5B).

El análisis de las isotermas se realizó mediante el ajuste no lineal por mínimos cuadrados de acuerdo con distintas ecuaciones dependiendo del tipo de unión del ligando a la macromolécula. Todos los ajustes de las curvas a los diferentes modelos se realizaron con el programa Origin 7.0. En los Apéndices III, IV y V, se han recogido los desarrollos matemáticos de donde derivan las ecuaciones para el ajuste de los siguientes modelos de unión:

**Apéndice III.** El modelo de unión de un ligando a una macromolécula con “n” Sitios Idénticos e Independientes (SII).

$$\frac{1}{V_c} \cdot \frac{dQ}{d[L]_T} \approx \frac{1}{V_c} \cdot \frac{\Delta Q}{\Delta[L]_T} = \frac{\Delta H_{ap}}{2} \left[ 1 - \frac{1 + [M]_T \cdot \theta - n \cdot K \cdot [M]_T}{\sqrt{(1 + K \cdot [M]_T \cdot \theta + n \cdot K \cdot [M]_T)^2 - 4 \cdot n \cdot K^2 \cdot [M]_T^2 \cdot \theta}} \right] \quad (\text{III.X ó 6.6.2})$$

siendo  $V_c$  el volumen de la célula y  $[M]_T$  la concentración de macromolécula total.

A partir del ajuste a esta ecuación, se obtuvieron la estequiometría,  $n$ ; la constante de unión en el equilibrio,  $K$ ; y la entalpía aparente de asociación,  $\Delta H_{app}$ , directamente del ajuste de los datos experimentales.

**Apéndice IV.** El modelo de unión de un ligando a una macromolécula con “m” clases de Sitios Diferentes e Independientes (SDI).

$$Q(N) = V_c \cdot [M]_T \left[ \Delta H_{ap,1} \frac{n_1 \cdot K_1 \cdot [L]}{1 + K_1 \cdot [L]} + \Delta H_{ap,2} \frac{n_2 \cdot K_2 \cdot [L]}{1 + K_2 \cdot [L]} \right] \quad (\text{IV.VII ó 6.6.3})$$

$$\Delta Q(j) = Q_T(j) - Q_T(j-1) + \frac{V_{in}}{V_c} \left( \frac{Q_T(j) + Q_T(j-1)}{2} \right) \quad (\text{IV.XI ó 6.6.4})$$

Las ecuaciones para este modelo se han desarrollado considerando dos clases de sitios diferentes e independientes. De modo que a partir del ajuste no lineal de las isotermas haciendo uso de estas ecuaciones, se obtuvo la estequiometría correspondiente para cada sitio,  $n_1$  y  $n_2$ ; las constantes de asociación en el equilibrio,  $K_1$  y  $K_2$ ; y las entalpías aparentes de unión para cada sitio,  $\Delta H_{app,1}$  y  $\Delta H_{app,2}$ .

**Apéndice V.** El modelo de unión a una macromolécula (con un solo sitio de unión) de un ligando A con afinidad moderada por desplazamiento de otro B de afinidad baja.

$$\Delta Q_i = V_c [M]_T \cdot \{ \Delta H_A \cdot (\Delta x_{MA,i} - f_i \cdot x_{MA,i-1}) + \Delta H_B \cdot (\Delta x_{MB,i} - f_i \cdot x_{MB,i-1}) \}$$

(V.XIV ó 6.6.5)

Esta ecuación utilizada en el ajuste por regresión no lineal de mínimos cuadrados, requiere definir previamente las fracciones molares de las especies formadas por la macromolécula,  $x_{MA}$ ,  $x_{MB}$  y  $x_M$  (ecuaciones V.V, V.VI y V.VII); y el factor de dilución tras cada inyección,  $f_i$  (ecuación V.XIII). Para el análisis con este modelo se utilizó una plantilla para el programa Origin 7.0 que generosamente fue ofrecida por el Dr. Sigurskjold (Universidad de Copenhague, Dinamarca) (Sigurskjold 2000). El término  $\Delta x_{MA,i} - f_i \cdot x_{MA,i-1}$  será siempre positivo, de modo que la contribución de la unión del ligando A al calor tendrá siempre el mismo signo que  $\Delta H_A$ . Mientras que el término  $\Delta x_{MB,i} - f_i \cdot x_{MB,i-1}$  será siempre negativo, de modo que la disociación del ligando B del complejo MB contribuirá al calor total con el signo opuesto a  $\Delta H_B$ . El modelo de unión por desplazamiento permite obtener directamente del ajuste los parámetros de unión del ligando B a la macromolécula: la constante de asociación en el equilibrio,  $K_B$ ; y la entalpía aparente de asociación,  $\Delta H_B$ .

Para todos los modelos utilizados, el resto de parámetros termodinámicos que caracterizan la interacción se obtienen haciendo uso de las ecuaciones fundamentales  $\Delta G = -RT \cdot \ln(K)$  y  $\Delta G = \Delta H - T \cdot \Delta S$  (Jelesarov & Bosshard 1999; Velazquez Campoy & Freire 2005).

## 6.7. Resolución estructural de complejos proteína-ligando mediante Resonancia Magnética Nuclear (RMN)

### 6.7.1. Introducción

La Resonancia Magnética Nuclear es un fenómeno físico basado en las propiedades magnéticas que poseen algunos núcleos atómicos. La mayoría de técnicas basadas en este fenómeno utilizan núcleos con número másico impar

como  $H^1$ ,  $C^{13}$ ,  $N^{15}$ ,  $O^{17}$ ,  $F^{19}$  y  $P^{31}$ . Cuando estos núcleos están sometidos a un campo magnético externo ( $B_0$ ) y se les aplica una onda electromagnética dentro de las frecuencias de radio, algunos de estos núcleos generan una pequeña corriente oscilante de radiofrecuencia que puede ser detectada cuando cesa la radiación por una bovina. Posteriormente, esta señal es amplificada y un ordenador genera una gráfica Voltaje/Tiempo que mediante la Transformada de Fourier es convertida en una gráfica Frecuencia/Intensidad que es lo que se conoce como el espectro de RMN como se ha ilustrado en la Figura 6.6. En un mismo campo magnético, la radiación con ondas de diferente frecuencia afectan a núcleos diferentes ( $H^1$ ,  $C^{13}$ ,  $N^{15}$ ). Por ejemplo, en un campo magnético de 14,1 Tesla, una onda de 600 MHz hará que los protones de la muestra generen una corriente oscilante de la misma frecuencia. Mientras una radiación de frecuencia 150,86 MHz en el mismo campo magnético afectará a los núcleos de  $^{13}C$ . Siguiendo con el ejemplo del protón, como estos núcleos se comportan como pequeños imanes, los núcleos vecinos alteran el campo magnético al que cada protón está sometido, por lo que en función del entorno químico, el campo magnético real ( $B_{\text{efectivo}}$ ) percibido por cada protón es diferente. Como la frecuencia de resonancia varía con el campo magnético, las diferencias en el  $B_{\text{efectivo}}$  hacen que cada protón resuene a una frecuencia distinta y, por tanto, que cada protón genere una señal diferente en el espectro. El desplazamiento de la frecuencia de resonancia de cada protón,  $\nu_{\text{núcleo}}$ , respecto a la frecuencia teórica,  $\nu_{\text{referencia}}$ , es lo que se conoce como desplazamiento químico,  $\delta_{\text{núcleo}}$ . Para poder comparar señales adquiridas en espectrómetros con  $B_0$  distintos, es necesario dividir los desplazamientos químicos por la frecuencia de operación del espectrofotómetro correspondiente,  $\nu_0$ . Los valores obtenidos son muy pequeños, por lo que se suelen expresar en parte por millón (ppm):

$$\delta_{\text{núcleo}}(\text{ppm}) = \frac{(\nu_{\text{núcleo}} - \nu_{\text{referencia}})}{\nu_0} \cdot 10^6 \quad (6.7.1)$$

La espectroscopía de RMN es una herramienta útil para identificar regiones de proteínas que interactúan con ligandos y para resolver estructuras moleculares, entre otras aplicaciones, ya que las variaciones del entorno químico

de los núcleos quedan reflejadas en los espectros de RMN por los cambios de sus frecuencias de resonancia. Existe una gran diversidad de experimentos de RMN basados en la aplicación de diferentes secuencias de pulsos que permiten seleccionar los núcleos observados ( $^1\text{H}$ s conectados a  $^1\text{H}$ ,  $^{15}\text{N}$  y  $^{13}\text{C}$  en nuestro caso). La utilización de la resonancia magnética nuclear en el estudio de proteínas tiene una importante limitación, el denominado problema del solapamiento. El gran número de átomos similares en una proteína hace que exista un elevado número de señales por espectro. La probabilidad de que estas estén superpuestas es muy elevada, lo cual dificulta su identificación. Sin embargo, mediante la utilización de la espectroscopía de bidimensional (2D) y tridimensional (3D) tanto homonuclear como heteronuclear, es posible la interpretación de los complejos espectros de RMN de proteínas y la obtención de los parámetros necesarios para la determinación de sus estructuras. Los espectros 2D se obtienen a partir de una serie de espectros monodimensionales (1D) con una variación sistemática de un tiempo de la secuencia de pulsos llamado tiempo de evolución,  $t_1$ . Esta variación hace que los picos aparezcan en diferentes 1D y al apilar los 1D se obtiene una segunda dimensión como se ilustran en la Figura 6.6. Para realizar un espectro 3D se acumula una serie de 2D a los que durante la adquisición del experimento se añade un segundo tiempo de evolución,  $t_2$  (Figura 6.6)

A continuación se presenta una relación de todos los espectros de RMN utilizados en la resolución estructural de dos complejos formados entre el dominio hNedd4-WW3 y dos ligandos peptídicos: el péptido **HTLV1-b** y el péptido **Ébola-a**.

Espectros monodimensionales (1D):

$^1\text{H}$ -monodimensional:

Los espectros monodimensionales de protón son los espectros más sencillos. Cuando la proteína está plegada, su espectro presenta dispersión en las señales de los protones del enlace peptídico (resuenan entre 6 y 10 ppm) y aparecen señales por debajo de 0 ppm debido a la interacción terciaria entre cadenas laterales. Cuando la proteína está desnaturalizada, la dispersión de la señal de los protones amida disminuye drásticamente y desaparecen las señales por debajo de 0 ppm debido a la pérdida de las interacciones terciarias. Además

de la dispersión de la señal de los  $^1\text{H}$  característica de todas las proteínas plegadas, los dominios WW presentan dos señales alrededor de 10 ppm generadas por la amida del indol de los 2 triptófanos característicos de estos dominios y también se han utilizado de forma estándar para comprobar que los dominios WW están plegados (Macias et al. 2000; Socolich et al. 2005; Jager et al. 2006). En la resolución estructural de los complejos de hNedd4-WW3, su principal aplicación fue la de comprobar si la proteína estaba plegada en las diferentes preparaciones de muestra realizadas.

Espectros bidimensionales (2D):

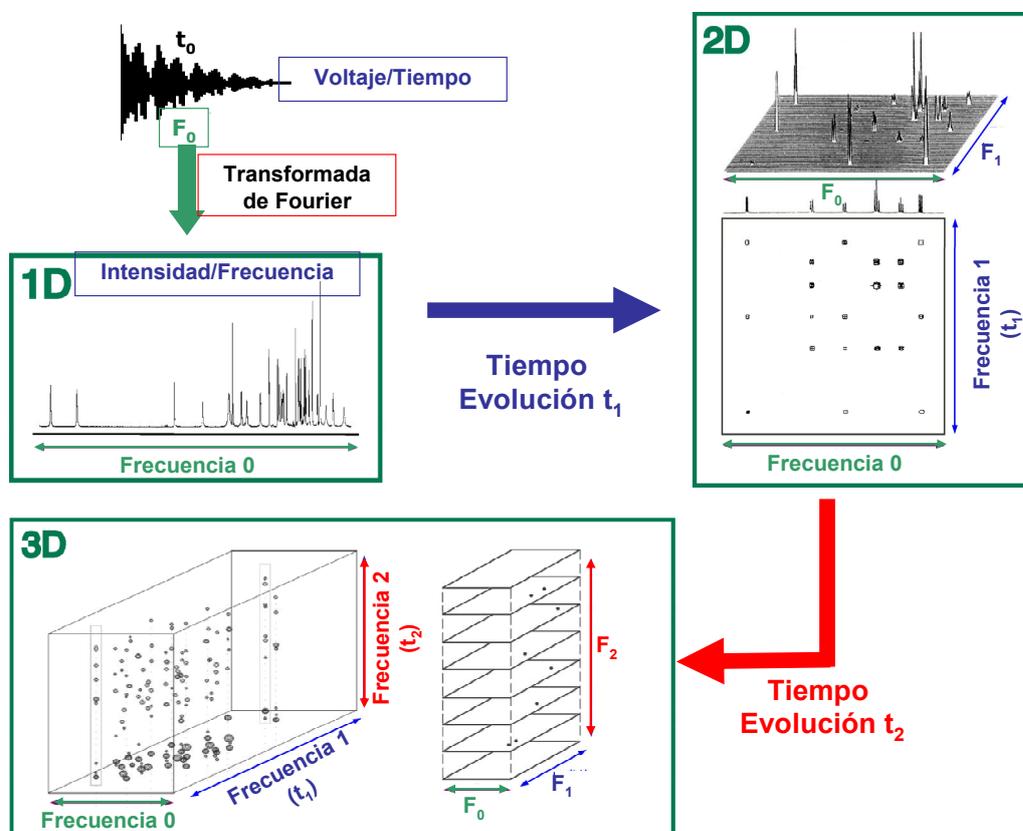
$^{-15}\text{N}$ -HSQC:

El experimento heteronuclear de 2D  $^{15}\text{N}$  HSQC (*"Heteronuclear Single Quantum Coherence"*) necesita que la proteína haya sido enriquecida isotópicamente en  $^{15}\text{N}$ , pues sirve para correlacionar el protón de la amida con el nitrógeno al que está unido. Por tanto, en estos experimentos se genera un espectro con tantas señales como residuos hay en la secuencia (exceptuando las prolinas). Además se observan pares de picos para las cadenas laterales con amidas como las de glutaminas (Q), asparaginas (N), argininas (R) y triptófano (W). Mediante el registro de sucesivos espectros HSQC se puede seguir el cambio en el entorno químico de varios grupos amida a la vez. Este espectro resulta de gran utilidad para monitorizar la interacción con un ligando, para determinar el efecto de la variación de la temperatura o del pH, o para detectar simplemente si una proteína esta plegada. En nuestro caso, estos espectros se han utilizado para monitorizar qué residuos se ven implicados en la interacción del dominio hNedd4-WW3 con cada uno de los dos ligandos.

$^{-1}\text{H}$ -TOCSY:

El TOCSY (*"Total Correlation Spectroscopy"*) es un espectro de 2D de  $^1\text{H}$  en el que los protones conectados por enlaces covalentes generan una señal como resultado de la transferencia de la magnetización a través de los enlaces. Se dice que estos protones están acoplados escalarmente. Este espectro se ha utilizado

durante la asignación para localizar y relacionar todos los protones pertenecientes a un mismo aminoácido.



**Figura 6.6. Fundamentos de Resonancia Magnética Nuclear (RMN).** Esquema de la adquisición de espectros de RMN de proteínas. Después de una secuencia de pulsos adecuada para un núcleo, se hace la Transformada de Fourier a la representación de la diferencia de potencial frente al tiempo ( $t_0$ ) y se obtiene un **espectro 1D** en el que los núcleos afectados por la radiación presentan diferentes frecuencias e intensidades debido a las diferencias de su entorno químico. Adquiriendo una serie de espectros 1D en los que el  $t_1$  se incrementa en cada espectro se obtiene un **espectro 2D**. Se muestra una visión de la serie de 1Ds apilados y una visión frontal del 2D en la que los picos se ven como contornos. Adquiriendo una serie de 2D en los que el  $t_2$  incrementa en cada espectro se obtiene un **espectro 3D**. Hay que realizar la Transformada de Fourier en las 3 dimensiones para que todas ellas se muestren como frecuencias ( $F_0, F_1, F_2$ ). En verde se muestra la dimensión directa (frecuencias adquiridas) y en azul y en rojo, las dimensiones indirectas.

$^1\text{H}$ -NOESY:

Este experimento se basa en el efecto NOE (*"Nuclear Overhauser Effect"*). El NOE se produce debido a una interacción a través del espacio entre 2 núcleos cercanos, concretamente a una distancia de hasta 5 Å. En contraste con el acoplamiento escalar (a través de enlaces), el NOE es el resultado de una interacción o acoplamiento dipolar (interacción entre dipolos magnéticos a través

del espacio). Si nos centramos en los espectros de proteínas, en el  $^1\text{H}$ -NOESY se pueden observar todos los picos del  $^1\text{H}$ -TOCSY, ya que los protones a tres enlaces de distancia no están separados espacialmente por una distancia superior a 5 Å. La característica más interesante del  $^1\text{H}$ -NOESY es que se observan otras señales entre protones que en la secuencia de aminoácidos son lejanos pero que debido a la estructura secundaria y terciaria de la proteína están situados a menos de 5 Å. Estas señales proporcionan la información principal que ha sido utilizada para resolver los dos complejos.

Espectros tridimensionales (3D):

-CBCANH, CACB(CO)NH y CC(CO)NH:

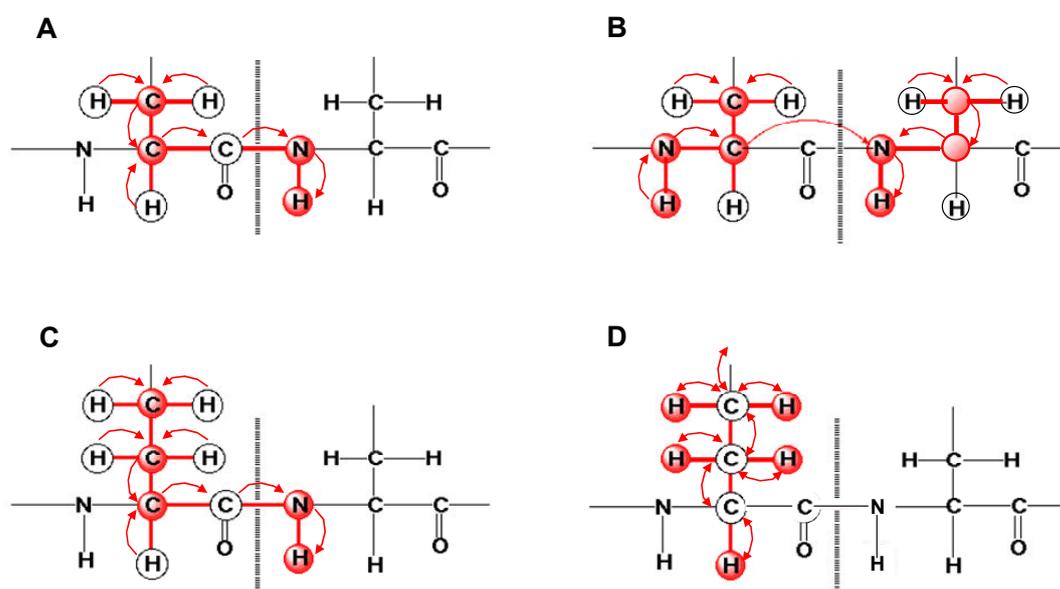
Los experimentos 3D CBCANH, CACB(CO)NH y CC(CO)NH transfieren la magnetización a través de los aminoácidos y del enlace peptídico permitiendo observar señales o picos de conexión entre aminoácidos consecutivos, es decir el esqueleto polipeptídico. A modo de ejemplo, en la Figura 6.7 se muestra la transferencia de la magnetización y que señales aparecen en varios de estos espectros. Debido a que transfieren la magnetización entre protones, nitrógeno y carbono las muestras de proteína para estos experimentos deben estar enriquecidas tanto en  $^{15}\text{N}$  como en  $^{13}\text{C}$ . Con estos espectros se asignan las frecuencias de todos los carbonos de una proteína y de las de los protones enlazados a éstos.

- $^{15}\text{N}$ -HSQC-TOCSY,  $^{15}\text{N}$ -HSQC-NOESY y HCCH-TOCSY:

En el experimento  $^{15}\text{N}$ -TOCSY-HSQC la magnetización se transfiere a través de protones vecinos y mediante el heteronúcleo  $^{15}\text{N}$  se incrementa la resolución del experimento aumentando en una dimensión el mismo. El resultado de este experimento es un cubo con dos dimensiones de  $^1\text{H}$  y una de  $^{15}\text{N}$ . Para cada plano de  $^{15}\text{N}$ - $^1\text{H}$  se observan los protones del mismo aminoácido.

En el experimento  $^{15}\text{N}$ -NOESY-HSQC, la magnetización se transfiere desde el protón amida mediante correlaciones de tipo NOE a los protones cercanos y para cada plano de  $^{15}\text{N}$ - $^1\text{H}$  se observan los NOEs intrarresiduo e interresiduo.

En el experimento HCCH-TOCSY la magnetización se transfiere por toda la cadena lateral a través de acoplamientos directos entre  $^1\text{H}$ - $^{13}\text{C}$  ó  $^{13}\text{C}$ - $^{13}\text{C}$  unidos por un solo enlace y obtenemos los planos  $^{13}\text{C}$ - $^1\text{H}$  para cada  $^{13}\text{C}$ , donde se pueden asignar los picos de cruce con los protones (Panel D de la Figura 6.7). Estos espectros se asignaron completamente con el fin de evitar ambigüedades debidas al solapamiento de señales y fueron de gran utilidad para completar la asignación del experimento bidimensional  $^1\text{H}$ -NOESY.



**Figura 6.7. Experimentos tridimensionales de RMN de proteínas.** Ejemplos de algunos experimentos de triple resonancia usados habitualmente para la asignación de proteínas: **A)** experimento CBCA(CO)NH, **B)** experimento CBCANH, **C)** experimento CC(CO)NH y **D)** experimento HCCH-TOCSY. Los distintos dipéptidos mostrados en la figura muestran las conectividades detectadas en cada experimento (como círculos rojos) y los átomos que no se detectan (como círculos vacíos), pero que se usan para la transferencia de la magnetización (línea roja).

### 6.7.2. Adquisición, procesado y asignación de los espectros de RMN

Todos los experimentos de RMN fueron adquiridos a 15 °C en 2 espectrómetros de RMN Bruker de 700MHz y 600MHz en la Unidad de Espectroscopía y RMN del CNIO de Madrid en colaboración con el Dr. Francisco J. Blanco (actualmente, Unidad de Biología Estructural del Centro de Investigación Científica BioGune, Vizcaya). Las muestras se prepararon en tampón fosfato sódico 20 mM pH 7.0 de acuerdo con el procedimiento descrito en el Apartado 6.2, usando concentraciones 1 mM del dominio hNedd4-WW3 y 3 mM de ligando

**HTLV1-b** ó 10 mM de ligando **Ébola-a** en cada caso. Todas las medidas se realizaron en tubos de RMN Shigemi Inc. que contenían 300  $\mu$ L de muestra en presencia de 25 mM DSS (ácido 4,4-dimetil-4-silapentano-1-sulfónico) y 1% acida sódica.

Para optimizar el tiempo de medida disponible en los dos espectrómetros, sólo se obtuvo un juego completo de espectros 2D y 3D para el complejo de hNedd4-WW3 con el ligando **HTLV1-b**. Una vez se asignó completamente el dominio WW en este primer complejo, se utilizó dicha asignación del dominio para resolver un segundo complejo de hNedd4-WW3 con el ligando **Ébola-a**, del que sólo se obtuvo el conjunto de espectros 2D. Para una mayor claridad, en la Tabla 6.4 se resumen los espectros registrados con cada una de las muestras (enriquecidas en  $^{13}\text{C}$  y/o  $^{15}\text{N}$ ), así como la información que se puede obtener de cada uno de ellos.

**Tabla 6.4. Resumen de los experimentos de RMN adquiridos.** Estos experimentos se ha utilizado para asignación y cálculo de la estructura tridimensional de los complejos formados por el dominio hNedd-WW3.

*Experimento	Muestra	Correlaciones	Información obtenida
<b>2D</b>			
$^{15}\text{N}$ -HSQC	$^{15}\text{N}$ $^1\text{H}$	H-N	Asignación esqueleto
$^1\text{H}$ -TOCSY	$^1\text{H}$	H-H	Asignación cadenas laterales
$^1\text{H}$ -NOESY	$^1\text{H}$	H-H	Restricciones distancia
<b>3D</b>			
CBCA(CO)NH	$^{15}\text{N}$ $^{13}\text{C}$	H, N, $\text{C}_\alpha$ , $\text{C}_\beta$	Asignación esqueleto
CACBNH	$^{15}\text{N}$ $^{13}\text{C}$	H, N, $\text{C}_\alpha$ , $\text{C}_\beta$	Asignación esqueleto
CC(CO)NH	$^{15}\text{N}$ $^{13}\text{C}$	H, N, $\text{C}_\alpha$ , $\text{C}_\beta$	Asignación esqueleto
HCCH-TOCSY	$^{15}\text{N}$ $^{13}\text{C}$	H, C, H	Asignación cadenas laterales
HCCH-NOESY	$^{15}\text{N}$ $^{13}\text{C}$	H, C, H	Restricciones distancia cadenas laterales
$^{15}\text{N}$ -TOCSY-HSQC	$^{15}\text{N}$ $^1\text{H}$	H, N, H	Asignación cadenas laterales
$^{15}\text{N}$ -NOESY-HSQC	$^{15}\text{N}$ $^1\text{H}$	H, N, H	Restricciones distancia

\*El juego completo de experimentos 2D y 3D, sólo se recogieron para el complejo con el ligando **HTLV1-b**. Para la estructura del complejo con el ligando del **Ébola-a** sólo se adquirieron experimentos 2D.

Una vez adquiridos, los espectros se procesaron con el programa TOPSPIN (Bruker Biospin, Karlsruhe, Alemania). Para ello se realizó la transformada de

Fourier de los datos brutos y se aplicó una corrección de línea base en todas las dimensiones. Entonces se asignaron de manera secuencial los átomos del esqueleto polipeptídico mediante los experimentos 3D CBCA(CO)NH, CBCANH y CC(CO)NH. A continuación, se identificaron las cadenas laterales alifáticas mediante la combinación de información procedente de los espectros 3D HCCH-TOCSY,  $^{15}\text{N}$ -TOCSY-HSQC y  $^{15}\text{N}$ -NOESY-HSQC. Luego, se obtuvo la asignación los protones aromáticos a partir de la información procedente de espectros 2D  $^1\text{H}$ -TOCSY y  $^1\text{H}$ -NOESY realizados tanto en agua (95%  $\text{H}_2\text{O}$  y 5%  $\text{D}_2\text{O}$ ) como en agua deuterada (100%  $\text{D}_2\text{O}$ ), concluyendo así la asignación de la proteína. Por último, combinando la información de los espectros 2D,  $^1\text{H}$ -TOCSY y  $^1\text{H}$ -NOESY (Tabla 6.4), se identificaron los átomos de cada péptido así como los NOEs intermoleculares de cada complejo. Para la visualización y asignación de los distintos espectros se utilizó el programa XEASY (Bartels et al. 1995).

La asignación de las señales de RMN de ambos complejos fue depositada en la base de datos del *BioMagResBank*, BMRB (<http://www.bmrb.wisc.edu/>), con los números de acceso 16574 para el complejo con el ligando **HTLV1-b** y 16575 para el complejo con el ligando del **Ébola-a**.

### 6.7.3. Cálculo de las estructuras

La determinación de la estructura tridimensional de una proteína en disolución consiste en la obtención de las coordenadas atómicas compatibles con las restricciones experimentales derivadas de los espectros de RMN. Normalmente las restricciones utilizadas para el cálculo de coordenadas son distancias interatómicas y una serie de ángulos diedros admisibles para las distintas regiones de la proteína de acuerdo con las diferentes estructuras secundarias presentes. Los programas informáticos para el cálculo de la estructura tridimensional de una proteína tratan de satisfacer tantas restricciones como sea posible, así como determinadas propiedades generales de las proteínas como son la longitud de los enlaces y los ángulos. Haciendo uso de distintos algoritmos, estos programas convierten las restricciones impuestas y las propiedades generales de las proteínas en una función de energía y, realizando diversas iteraciones, tratan de que el valor de esta función sea mínimo. En el proceso se genera un conjunto de

estructuras que convergen si los datos son suficientes y tienen la calidad adecuada.

Las restricciones de distancia entre protones intermoleculares e intramoleculares de ambos complejos formados por hNedd4-WW3 se obtuvieron a partir de la integración del volumen (o intensidad) de las señales de correlación por el efecto NOE en los espectros  $^1\text{H}$ -NOESY mediante el programa XEASY (Bartels et al. 1995).

Las restricciones para los ángulos diedros “*phi*” ( $\phi$ ) y “*psi*” ( $\psi$ ) de la cadena principal de la proteína utilizados en el cálculo se estimaron a partir de la predicción de la estructura secundaria de la proteína a través del índice de desplazamiento químico (*Chemical-Shift Index*; CSI) (Wishart & Sykes 1994). Este método consiste en comparar los desplazamientos químicos experimentales de los núcleos  $^1\text{H}_\alpha$ ,  $^{13}\text{C}_\alpha$ ,  $^{13}\text{C}_\beta$  y  $^{13}\text{CO}$  con los tabulados para estos en una conformación desordenada, de modo que se asigna un índice de +1, 0 ó -1 según esas diferencias sean positivas o negativas y superen en magnitud ciertos umbrales previamente establecidos. Teniendo en cuenta los patrones de estos índices que normalmente presentan las estructuras secundarias para cada uno de estos núcleos, se puede identificar cuáles de los residuos se encuentran formando dichas estructuras.

Los cálculos estructurales se realizaron en colaboración con la Dra. Maria J. Macias del Laboratorio de RMN de Proteínas del Instituto de Investigación Biomédicas de Barcelona. Para ello se utilizó el programa “Structcalc” (<http://maciasnmr.net/>), basado en CNS (“Cristallography and NMR System”) (Brunger et al. 1998), que ha sido desarrollado en este laboratorio por Pau Martín. En este proceso se calcularon 120 modelos estructurales de acuerdo con las restricciones correspondientes para cada complejo. De todos estos modelos, se seleccionaron los 20 modelos de menor energía.

#### **6.7.4. Evaluación de las estructuras**

Para analizar la calidad de las estructuras se suelen evaluar tanto la precisión de las coordenadas como la desviación de sus ángulos en términos estadísticos.

La precisión de las coordenadas de las estructuras se analizó mediante una inspección visual de los 20 modelos de menor energía superpuestos con el programa PYMOL (<http://www.pymol.org/>) y se calculó la desviación cuadrática media (RMSD) de las coordenadas de estos 20 modelos mediante el programa MOLMOL (Koradi et al. 1996).

Por otro lado, también se analizaron los ángulos entre los distintos elementos de estructura secundaria de los 20 modelos de menor energía con el programa PROCHECK-NMR (Laskowski et al. 1996). Este programa representa los ángulos diedros sobre el mapa de Ramachandran, un gráfico que representa los valores estándar de estos ángulos para proteínas. PROCHECK-NMR muestra así las desviaciones de los ángulos para los modelos analizados.

Las coordenadas de los 20 modelos derivados del cálculo de la estructura de cada complejo han sido depositadas en el “*Protein Data Bank*”, PDB (<http://www.rcsb.org/pdb/>), con los números de acceso 2KPZ y 2KP0 para los complejos con el péptido **HTLV1-b** y con el péptido del **Ébola-a**, respectivamente.

### **6.8. Cálculo del potencial electrostático de una proteína mediante el programa DelPhi**

Todos los cálculos de potencial electrostático tanto de los dominios WW como de los péptidos que los reconocen, en distintas estructuras de complejos, se han realizado mediante el programa DelPhi implementado en el conjunto de aplicaciones informáticas Discovery Studio (*Accelrys Inc.*). El programa DelPhi calcula la energía electrostática basada en un potencial continuo y resolviendo numéricamente una aproximación de la ecuación de Poisson-Boltzmann. Estos cálculos los realiza en dos pasos: en el primero, se realiza la estimación de la energía coulombiana para la proteína o el péptido. En nuestro caso se asignaron las cargas propias de todos los átomos en el sistema de acuerdo con los parámetros del campo de fuerza CHARMM (Brooks et al. 2009). En el segundo, se estima la energía libre de solvatación. La energía coulombiana es la energía libre correspondiente a la disposición las cargas atómicas desde el infinito a un medio dieléctrico de constante igual a la del interior de la proteína. La energía libre de solvatación es la energía resultante de la transferencia de la proteína de un medio

con una constante dieléctrica dada a un medio con una constante dieléctrica del agua. La evaluación de estas cantidades para la proteína y para el solvente permite calcular la energía electrostática.

### **6.8.1. Generación de modelos estructurales**

Para aquellos complejos para los que no se dispone de una estructura obtenida experimentalmente, se decidió modelarla a partir de otra estructura existente. A ser posible, se utilizó la estructura de un complejo formado con el mismo dominio u otro muy similar. Para producir los archivos de coordenadas de estos dominios WW se separaron de las coordenadas del ligando y se realizaron las siguientes operaciones mediante las herramientas de diseño que contiene el programa Discovery Studio. Con el objetivo de alterar lo menos posible la estructura del dominio WW de estructura conocida, tan sólo se modificaron aquellas cadenas laterales de los residuos del archivo que no coinciden en un alineamiento de secuencia con el dominio WW problema. De forma análoga se sustituyeron las cadenas laterales del péptido por las del que nos interesa. Por tanto, se realizaron los cálculos del potencial electrostático de dos moléculas independientes: la estructura del dominio WW libre y de un péptido libre; pero ambos con la conformación que tuvieran en el complejo.



## BIBLIOGRAFÍA



## BIBLIOGRAFÍA

1. Alderton GK (2010) Signalling: All roads lead to YAP1. *Nat Rev Cancer* 10:6-6.
2. Anderson SR (1969) Fluorescence polarization studies of conjugates of beef heart lactic dehydrogenase with 1-dimethylaminonaphthalene-5-sulfonyl chloride. *Biochemistry* 8:1394-1396.
3. Anderson SR, Weber G (1969) Fluorescence polarization of the complexes of 1-anilino-8-naphthalenesulfonate with bovine serum albumin. Evidence for preferential orientation of the ligand. *Biochemistry* 8:371-377.
4. Arbely E, Rutherford TJ, Sharpe TD, Ferguson N, Fersht AR (2009) Downhill versus barrier-limited folding of BBL 1: energetic and structural perturbation effects upon protonation of a histidine of unusually low pKa. *J Mol Biol* 387:986-992.
5. Arold S, O'Brien R, Franken P, Strub MP, Hoh F, Dumas C, Ladbury JE (1998) RT loop flexibility enhances the specificity of Src family SH3 domains for HIV-1 Nef. *Biochemistry* 37:14683-14691.
6. Arold ST, Ulmer TS, Mulhern TD, Werner JM, Ladbury JE, Campbell ID, Noble ME (2001) The role of the Src homology 3-Src homology 2 interface in the regulation of Src kinases. *J Biol Chem* 276:17199-17205.
7. Ausubel FM, Brent, R. Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K. 1992. *Short protocols in molecular biology*, Wiley & Sons, New York (USA).
8. Azuaga AI, Galisteo ML, Mayorga OL, Cortijo M, Mateo PL (1992) Heat and cold denaturation of beta-lactoglobulin B. *FEBS Lett* 309:258-260.
9. Ball LJ, Kuhne R, Schneider-Mergener J, Oschkinat H (2005) Recognition of proline-rich motifs by protein-protein-interaction domains. *Angew Chem Int Ed Engl* 44:2852-2869.
10. Barila D, Superti-Furga G (1998) An intramolecular SH3-domain interaction regulates c-Abl activity. *Nat Genet* 18:280-282.
11. Bartels C, Xia TH, Billeter M, Guntert P, Wuthrich K (1995) The Program Xeasy for Computer-Supported Nmr Spectral-Analysis of Biological Macromolecules. *Journal of Biomolecular Nmr* 6:1-10.

12. Basu S, Totty NF, Irwin MS, Sudol M, Downward J (2003) Akt phosphorylates the Yes-associated protein, YAP, to induce interaction with 14-3-3 and attenuation of p73-mediated apoptosis. *Mol Cell* 11:11-23.
13. Bedford MT, Chan DC, Leder P (1997) FBP WW domains and the Abl SH3 domain bind to a specific class of proline-rich ligands. *EMBO J* 16:2376-2383.
14. Bedford MT, Reed R, Leder P (1998) WW domain-mediated interactions reveal a spliceosome-associated protein that binds a third class of proline-rich motif: the proline glycine and methionine-rich motif. *Proc Natl Acad Sci U S A* 95:10602-10607.
15. Bedford MT, Sarbassova D, Xu J, Leder P, Yaffe MB (2000) A novel pro-Arg motif recognized by WW domains. *J Biol Chem* 275:10359-10369.
16. Belgareh-Touze N, Leon S, Erpapazoglou Z, Stawiecka-Mirota M, Urban-Grimal D, Haguenaer-Tsapis R (2008) Versatile role of the yeast ubiquitin ligase Rsp5p in intracellular trafficking. *Biochem Soc Trans* 36:791-796.
17. Bellomaria A, Barbato G, Melino G, Paci M, Melino S (2010) Recognition of p63 by the E3 ligase ITCH: Effect of an ectodermal dysplasia mutant. *Cell Cycle* 9:3730-3739.
18. Bertini E, Oka T, Sudol M, Strano S, Blandino G (2009) YAP: at the crossroad between transformation and tumor suppression. *Cell Cycle* 8:49-57.
19. Bhakuni V, Xie D, Freire E (1991) Thermodynamic identification of stable folding intermediates in the B-subunit of cholera toxin. *Biochemistry* 30:5055-5060.
20. Bieniasz PD (2006) Late budding domains and host proteins in enveloped virus release. *Virology* 344:55-63.
21. Blot V, Perugi F, Gay B, Prevost MC, Briant L, Tangy F, Abriel H, Staub O, Dokhelar MC, Pique C (2004) Nedd4.1-mediated ubiquitination and subsequent recruitment of Tsg101 ensure HTLV-1 Gag trafficking towards the multivesicular body pathway prior to virus budding. *J Cell Sci* 117:2357-2367.
22. Bork P, Sudol M (1994) The WW domain: a signalling site in dystrophin? *Trends Biochem Sci* 19:531-533.
23. Bouamr F, Melillo JA, Wang MQ, Nagashima K, de Los Santos M, Rein A, Goff SP (2003) PPPYVEPTAP motif is the late domain of human T-cell leukemia

- virus type 1 Gag and mediates its functional interaction with cellular proteins Nedd4 and Tsg101 [corrected]. *J Virol* 77:11882-11895.
24. Brabek J, Mojzita D, Novotny M, Puta F, Folk P (2002) The SH3 domain of Src can downregulate its kinase activity in the absence of the SH2 domain-pY527 interaction. *Biochem Biophys Res Commun* 296:664-670.
  25. Brannetti B, Helmer-Citterich M (2003) iSPOT: A web tool to infer the interaction specificity of families of protein modules. *Nucleic Acids Res* 31:3709-3711.
  26. Brasher BB, Roumiantsev S, Van Etten RA (2001) Mutational analysis of the regulatory function of the c-Abl Src homology 3 domain. *Oncogene* 20:7744-7752.
  27. Brooks BR, Brooks CL, Mackerell AD, Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S, Caflisch A, Caves L, Cui Q, Dinner AR, Feig M, Fischer S, Gao J, Hodoscek M, Im W, Kuczera K, Lazaridis T, Ma J, Ovchinnikov V, Paci E, Pastor RW, Post CB, Pu JZ, Schaefer M, Tidor B, Venable RM, Woodcock HL, Wu X, Yang W, York DM, Karplus M (2009) CHARMM: The Biomolecular Simulation Program. *J Comput Chem* 30:1545-1614.
  28. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D* 54:905-921.
  29. Bruscolini P, Pelizzola A, Zamparo M (2007) Downhill versus two-state protein folding in a statistical mechanical model. *J Chem Phys* 126:215103.
  30. Bruylants G, Wouters J, Michaux C (2005) Differential scanning calorimetry in life science: thermodynamics, stability, molecular recognition and application in drug design. *Curr Med Chem* 12:2011-2020.
  31. Candel AM. Diseño y caracterización de proteínas quiméricas para el estudio del reconocimiento molecular entre ligandos péptidicos y dominios SH3 (2008a) Química-Física. Universidad de Granada, Granada.
  32. Candel AM. Tesis Doctoral: Diseño y caracterización de proteínas quiméricas para el estudio del reconocimiento molecular entre ligandos peptídicos y

- dominios SH3. (2008b) Departamento de Química-Física Universidad de Granada, Granada.
33. Candel AM, Cobos ES, Conejero-Lara F, Martinez JC (2009) Evaluation of folding co-operativity of a chimeric protein based on the molecular recognition between polyproline ligands and SH3 domains. *Protein Eng Des Sel* 22:597-606.
  34. Candel AM, van Nuland NA, Martin-Sierra FM, Martinez JC, Conejero-Lara F (2008) Analysis of the thermodynamics of binding of an SH3 domain to proline-rich peptides using a chimeric fusion protein. *J Mol Biol* 377:117-135.
  35. Casares S. Tesis Doctoral: Análisis de las rutas cooperativas de propagación de estabilidad en el dominio SH3. Aplicación al estudio de la interacciones proteína ligando (2003) Departamento de Química-Física Universidad de Granada, Granada.
  36. Cobos ES. Tesis Doctoral: Estabilidad, plegamiento y estructura de la proteína circular AS-48 y de mutantes del dominio SH3. (2002) Departamento de Química-Física Universidad de Granada, Granada.
  37. Cobos ES, Filimonov VV, Galvez A, Maqueda M, Valdivia E, Martinez JC, Mateo PL (2001) AS-48: a circular protein with an extremely stable globular structure. *FEBS Lett* 505:379-382.
  38. Cobos ES, Filimonov VV, Galvez A, Valdivia E, Maqueda M, Martinez JC, Mateo PL (2002) The denaturation of circular enterocin AS-48 by urea and guanidinium hydrochloride. *Biochim Biophys Acta* 1598:98-107.
  39. Cobos ES, Pisabarro MT, Vega MC, Lacroix E, Serrano L, Ruiz-Sanz J, Martinez JC (2004) A miniprotein scaffold used to assemble the polyproline II binding epitope recognized by SH3 domains. *J Mol Biol* 342:355-365.
  40. Compton LA, Johnson WC, Jr. (1986) Analysis of protein circular dichroism spectra for secondary structure using a simple matrix multiplication. *Anal Biochem* 155:155-167.
  41. Condie CC, Quay SC (1983) Conformational studies of aqueous melittin. Characteristics of a fluorescent probe binding site. *J Biol Chem* 258:8231-8234.

42. Cooper A (2005) Heat capacity effects in protein folding and ligand binding: a re-evaluation of the role of water in biomolecular thermodynamics. *Biophys Chem* 115:89-97.
43. Cooper A, Johnson CM (1994) Differential scanning calorimetry. *Methods Mol Biol* 22:125-136.
44. Crane JC, Koepf EK, Kelly JW, Gruebele M (2000) Mapping the transition state of the WW domain beta-sheet. *J Mol Biol* 298:283-292.
45. Creighton TE. 1993. *Proteins: Structures and Molecular Properties*, WH Freeman and Co., New York (USA).
46. Chen C, Li F, Montelaro RC (2001) Functional roles of equine infectious anemia virus Gag p9 in viral budding and infection. *J Virol* 75:9762-9770.
47. Chong PA, Lin H, Wrana JL, Forman-Kay JD (2010) Coupling of tandem Smad ubiquitination regulatory factor (Smurf) WW domains modulates target specificity. *Proc Natl Acad Sci U S A* 107:18404-18409.
48. Dalgarno DC, Botfield MC, Rickles RJ (1997) SH3 domains and drug design: ligands, structure, and biological function. *Biopolymers* 43:383-400.
49. Daniel E, Weber G (1966) Cooperative effects in binding by bovine serum albumin. I. The binding of 1-anilino-8-naphthalenesulfonate. Fluorimetric titrations. *Biochemistry* 5:1893-1900.
50. Deechongkit S, Nguyen H, Powers ET, Dawson PE, Gruebele M, Kelly JW (2004) Context-dependent contributions of backbone hydrogen bonding to beta-sheet folding energetics. *Nature* 430:101-105.
51. Demirov DG, Freed EO (2004) Retrovirus budding. *Virus Res* 106:87-102.
52. Dragan AI, Privalov PL (2002) Unfolding of a leucine zipper is not a simple two-state transition. *J Mol Biol* 321:891-908.
53. Eaton WA, Munoz V, Hagen SJ, Jas GS, Lapidus LJ, Henry ER, Hofrichter J (2000) Fast kinetics and mechanisms in protein folding. *Annu Rev Biophys Biomol Struct* 29:327-359.
54. Espanel X, Sudol M (2001) Yes-associated protein and p53-binding protein-2 interact through their WW and SH3 domains. *J Biol Chem* 276:14514-14523.

55. Faraldo-Gomez JD, Roux B (2007a) Characterization of conformational equilibria through Hamiltonian and temperature replica-exchange simulations: assessing entropic and environmental effects. *J Comput Chem* 28:1634-1647.
56. Faraldo-Gomez JD, Roux B (2007b) On the importance of a funneled energy landscape for the assembly and regulation of multidomain Src tyrosine kinases. *Proc Natl Acad Sci U S A* 104:13643-13648.
57. Farber P, Darmawan H, Sprules T, Mittermaier A (2010) Analyzing protein folding cooperativity by differential scanning calorimetry and NMR spectroscopy. *J Am Chem Soc* 132:6214-6222.
58. Fedoroff OY, Townson SA, Golovanov AP, Baron M, Avis JM (2004) The structure and dynamics of tandem WW domains in a negative regulator of notch signaling, Suppressor of deltex. *J Biol Chem* 279:34991-35000.
59. Felitsky DJ, Record MT, Jr. (2003) Thermal and urea-induced unfolding of the marginally stable lac repressor DNA-binding domain: a model system for analysis of solute effects on protein processes. *Biochemistry* 42:2202-2217.
60. Ferguson N, Johnson CM, Macias M, Oschkinat H, Fersht A (2001a) Ultrafast folding of WW domains without structured aromatic clusters in the denatured state. *Proc Natl Acad Sci U S A* 98:13002-13007.
61. Ferguson N, Pires JR, Toepert F, Johnson CM, Pan YP, Volkmer-Engert R, Schneider-Mergener J, Daggett V, Oschkinat H, Fersht A (2001b) Using flexible loop mimetics to extend phi-value analysis to secondary structure interactions. *Proc Natl Acad Sci U S A* 98:13008-13013.
62. Fernandez LA, Northcott PA, Dalton J, Fraga C, Ellison D, Angers S, Taylor MD, Kenney AM (2009) YAP1 is amplified and up-regulated in hedgehog-associated medulloblastomas and mediates Sonic hedgehog-driven neural precursor proliferation. *Genes Dev* 23:2729-2741.
63. Ferreon JC, Hilser VJ (2004) Thermodynamics of binding to SH3 domains: the energetic impact of polyproline II (PII) helix formation. *Biochemistry* 43:7787-7797.
64. Freed EO (2002) Viral late domains. *J Virol* 76:4679-4687.

65. Freire E. Differential scanning calorimetry. In: Shirley BA, Ed. (1995) Protein stability and folding: Theory and practice. Humana Press Inc., Totowa, NJ, pp. 191-218.
66. Freire E (2008) Do enthalpy and entropy distinguish first in class from best in class? *Drug Discov Today* 13:869-874.
67. Freire E, Hayne DT, Xie D (1993) Molecular basis of cooperativity in protein folding. 4. Core - a general cooperative folding model. *Proteins: Struct Funct Genet* 17:111-123.
68. 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.
69. Fung A, Li P, Godoy-Ruiz R, Sanchez-Ruiz JM, Munoz V (2008) Expanding the realm of ultrafast protein folding: gpW, a midsize natural single-domain with alpha+beta topology that folds downhill. *J Am Chem Soc* 130:7489-7495.
70. Garcia-Mira MM, Sadqi M, Fischer N, Sanchez-Ruiz JM, Munoz V (2002) Experimental identification of downhill protein folding. *Science* 298:2191-2195.
71. Garnier L, Wills JW, Verderame MF, Sudol M (1996) WW domains and retrovirus budding. *Nature* 381:744-745.
72. Garrus JE, von Schwedler UK, Pornillos OW, Morham SG, Zavitz KH, Wang HE, Wettstein DA, Stray KM, Cote M, Rich RL, Myszka DG, Sundquist WI (2001) Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. *Cell* 107:55-65.
73. Gill SC, von Hippel PH (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem* 182:319-326.
74. Ginsburg A, Carroll WR (1965) Some specific ion effects on the conformation and thermal stability of ribonuclease. *Biochemistry* 4:2159-2174.
75. Gomez J, Hilser VJ, Xie D, Freire E (1995) The heat capacity of proteins. *Proteins: Struct Funct Genet* 22:404-412.
76. Gottlinger HG, Dorfman T, Sodroski JG, Haseltine WA (1991) Effect of mutations affecting the p6 gag protein on human immunodeficiency virus particle release. *Proc Natl Acad Sci U S A* 88:3195-3199.
77. Hao Y, Chun A, Cheung K, Rashidi B, Yang X (2008) Tumor suppressor LATS1 is a negative regulator of oncogene YAP. *J Biol Chem* 283:5496-5509.

78. Harty RN, Brown ME, McGettigan JP, Wang G, Jayakar HR, Huibregtse JM, Whitt MA, Schnell MJ (2001) Rhabdoviruses and the cellular ubiquitin-proteasome system: a budding interaction. *J Virol* 75:10623-10629.
79. Harty RN, Brown ME, Wang G, Huibregtse J, Hayes FP (2000) A PPxY motif within the VP40 protein of Ebola virus interacts physically and functionally with a ubiquitin ligase: implications for filovirus budding. *Proc Natl Acad Sci U S A* 97:13871-13876.
80. Harty RN, Paragas J, Sudol M, Palese P (1999) A proline-rich motif within the matrix protein of vesicular stomatitis virus and rabies virus interacts with WW domains of cellular proteins: implications for viral budding. *J Virol* 73:2921-2929.
81. Harvey KF, Kumar S (1999) Nedd4-like proteins: an emerging family of ubiquitin-protein ligases implicated in diverse cellular functions. *Trends Cell Biol* 9:166-169.
82. Hendsch ZS, Jonsson T, Sauer RT, Tidor B (1996) Protein stabilization by removal of unsatisfied polar groups: computational approaches and experimental tests. *Biochemistry* 35:7621-7625.
83. Hesselberth JR, Miller JP, Golob A, Stajich JE, Michaud GA, Fields S (2006) Comparative analysis of *Saccharomyces cerevisiae* WW domains and their interacting proteins. *Genome Biol* 7:R30.
84. Hilser VJ, Gomez J, Freire E (1996) The enthalpy change in protein folding and binding: refinement of parameters for structure-based calculations. *Proteins* 26:123-133.
85. Hu H, Columbus J, Zhang Y, Wu D, Lian L, Yang S, Goodwin J, Luczak C, Carter M, Chen L, James M, Davis R, Sudol M, Rodwell J, Herrero JJ (2004) A map of WW domain family interactions. *Proteomics* 4:643-655.
86. Huang M, Orenstein JM, Martin MA, Freed EO (1995) p6Gag is required for particle production from full-length human immunodeficiency virus type 1 molecular clones expressing protease. *J Virol* 69:6810-6818.
87. Huang X, Beullens M, Zhang J, Zhou Y, Nicolaescu E, Lesage B, Hu Q, Wu J, Bollen M, Shi Y (2009) Structure and function of the two tandem WW domains

- of the pre-mRNA splicing factor FBP21 (formin-binding protein 21). *J Biol Chem* 284:25375-25387.
88. Huang X, Poy F, Zhang R, Joachimiak A, Sudol M, Eck MJ (2000) Structure of a WW domain containing fragment of dystrophin in complex with beta-dystroglycan. *Nat Struct Biol* 7:634-638.
89. Ibragimova GT, Wade RC (1999) Stability of the beta-sheet of the WW domain: A molecular dynamics simulation study. *Biophys J* 77:2191-2198.
90. Im YJ, Kuo L, Ren X, Burgos PV, Zhao XZ, Liu F, Burke TR, Jr., Bonifacino JS, Freed EO, Hurley JH (2010) Crystallographic and functional analysis of the ESCRT-I /HIV-1 Gag PTAP interaction. *Structure* 18:1536-1547.
91. Ingham RJ, Gish G, Pawson T (2004) The Nedd4 family of E3 ubiquitin ligases: functional diversity within a common modular architecture. *Oncogene* 23:1972-1984.
92. Jackson SE, Fersht AR (1991) Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition. *Biochemistry* 30:10428-10435.
93. Jaenicke R (2000) Do ultrastable proteins from hyperthermophiles have high or low conformational rigidity? *Proc Natl Acad Sci U S A* 97:2962-2964.
94. Jager M, Deechongkit S, Koepf EK, Nguyen H, Gao J, Powers ET, Gruebele M, Kelly JW (2008) Understanding the mechanism of beta-sheet folding from a chemical and biological perspective. *Biopolymers* 90:751-758.
95. Jager M, Fischer J, Schultheis A, Lensing-Hohn S, Krauspe R (2006) Extensive H(+) release by bone substitutes affects biocompatibility in vitro testing. *J Biomed Mater Res A* 76:310-322.
96. Jager M, Nguyen H, Crane JC, Kelly JW, Gruebele M (2001) The folding mechanism of a beta-sheet: the WW domain. *J Mol Biol* 311:373-393.
97. Jager M, Nguyen H, Dendle M, Gruebele M, Kelly JW (2007) Influence of hPin1 WW N-terminal domain boundaries on function, protein stability, and folding. *Protein Sci* 16:1495-1501.
98. Jelen F, Oleksy A, Smietana K, Otlewski J (2003) PDZ domains - common players in the cell signaling. *Acta Biochim Pol* 50:985-1017.

99. Jelesarov I, Bosshard HR (1999) Isothermal titration calorimetry and differential scanning calorimetry as complementary tools to investigate the energetics of biomolecular recognition. *J Mol Recognit* 12:3-18.
100. Jennings MD, Blankley RT, Baron M, Golovanov AP, Avis JM (2007) Specificity and autoregulation of Notch binding by tandem WW domains in suppressor of Deltex. *J Biol Chem* 282:29032-29042.
101. Johnson WC, Jr. (1990) Protein secondary structure and circular dichroism: a practical guide. *Proteins* 7:205-214.
102. Kanai F, Marignani PA, Sarbassova D, Yagi R, Hall RA, Donowitz M, Hisaminato A, Fujiwara T, Ito Y, Cantley LC, Yaffe MB (2000) TAZ: a novel transcriptional co-activator regulated by interactions with 14-3-3 and PDZ domain proteins. *EMBO J* 19:6778-6791.
103. Kanelis V, Bruce MC, Skrynnikov NR, Rotin D, Forman-Kay JD (2006) Structural determinants for high-affinity binding in a Nedd4 WW3\* domain-Comm PY motif complex. *Structure* 14:543-553.
104. Kanelis V, Rotin D, Forman-Kay JD (2001) Solution structure of a Nedd4 WW domain-ENaC peptide complex. *Nat Struct Biol* 8:407-412.
105. Karanicolas J, Brooks CL, 3rd (2003) The structural basis for biphasic kinetics in the folding of the WW domain from a formin-binding protein: lessons for protein design? *Proc Natl Acad Sci U S A* 100:3954-3959.
106. Karanicolas J, Brooks CL, 3rd (2004) Integrating folding kinetics and protein function: biphasic kinetics and dual binding specificity in a WW domain. *Proc Natl Acad Sci U S A* 101:3432-3437.
107. Kardinal C, Konkol B, Schulz A, Posern G, Lin H, Adermann K, Eulitz M, Estrov Z, Talpaz M, Arlinghaus RB, Feller SM (2000) Cell-penetrating SH3 domain blocker peptides inhibit proliferation of primary blast cells from CML patients. *FASEB J* 14:1529-1538.
108. Kasanov J, Pirozzi G, Uveges AJ, Kay BK (2001) Characterizing Class I WW domains defines key specificity determinants and generates mutant domains with novel specificities. *Chem Biol* 8:231-241.

109. Kay BK, Williamson MP, Sudol M (2000) The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *FASEB J* 14:231-241.
110. Knapp S, Mattson PT, Christova P, Berndt KD, Karshikoff A, Vihinen M, Smith CI, Ladenstein R (1998) Thermal unfolding of small proteins with SH3 domain folding pattern. *Proteins* 31:309-319.
111. Koepf EK, Petrassi HM, Ratnaswamy G, Huff ME, Sudol M, Kelly JW (1999a) Characterization of the structure and function of W --> F WW domain variants: identification of a natively unfolded protein that folds upon ligand binding. *Biochemistry* 38:14338-14351.
112. Koepf EK, Petrassi HM, Sudol M, Kelly JW (1999b) WW: An isolated three-stranded antiparallel beta-sheet domain that unfolds and refolds reversibly; evidence for a structured hydrophobic cluster in urea and GdnHCl and a disordered thermal unfolded state. *Protein Sci* 8:841-853.
113. Komuro A, Nagai M, Navin NE, Sudol M (2003) WW domain-containing protein YAP associates with ErbB-4 and acts as a co-transcriptional activator for the carboxyl-terminal fragment of ErbB-4 that translocates to the nucleus. *J Biol Chem* 278:33334-33341.
114. Koradi R, Billeter M, Wuthrich K (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J Mol Graph* 14:51-55, 29-32.
115. Korzhnev DM, Salvatella X, Vendruscolo M, Di Nardo AA, Davidson AR, Dobson CM, Kay LE (2004) Low-populated folding intermediates of Fyn SH3 characterized by relaxation dispersion NMR. *Nature* 430:586-590.
116. Kuriyan J, Cowburn D (1997) Modular peptide recognition domains in eukaryotic signaling. *Annu Rev Biophys Biomol Struct* 26:259-288.
117. Ladbury JE, Arold S (2000) Searching for specificity in SH domains. *Chem Biol* 7:R3-8.
118. Ladbury JE, Chowdhry BZ (1996) Sensing the heat: the application of isothermal titration calorimetry to thermodynamic studies of biomolecular interactions. *Chem Biol* 3:791-801.

119. Lafont V, Armstrong AA, Ohtaka H, Kiso Y, Mario Amzel L, Freire E (2007) Compensating enthalpic and entropic changes hinder binding affinity optimization. *Chem Biol Drug Des* 69:413-422.
120. Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* 8:477-486.
121. Lee KY, Hyeok Yoon JH, Kim M, Roh S, Lee YS, Seong BL, Kim K (2002) A dipalmitoyl peptide that binds SH3 domain, disturbs intracellular signal transduction, and inhibits tumor growth in vivo. *Biochem Biophys Res Commun* 296:434-442.
122. Li P, Oliva FY, Naganathan AN, Munoz V (2009) Dynamics of one-state downhill protein folding. *Proc Natl Acad Sci U S A* 106:103-108.
123. Li Z, Lazaridis T (2005) The effect of water displacement on binding thermodynamics: concanavalin A. *J Phys Chem B* 109:662-670.
124. Liu F, Du D, Fuller AA, Davoren JE, Wipf P, Kelly JW, Gruebele M (2008) An experimental survey of the transition between two-state and downhill protein folding scenarios. *Proc Natl Acad Sci U S A* 105:2369-2374.
125. Liu F, Dumont C, Zhu Y, DeGrado WF, Gai F, Gruebele M (2009) A one-dimensional free energy surface does not account for two-probe folding kinetics of protein alpha(3)D. *J Chem Phys* 130:061101.
126. Lopez Mayorga O, Freire E (1987) Dynamic analysis of differential scanning calorimetry data. *Biophys Chem* 27:87-96.
127. Lovell SC, Davis IW, Arendall WB, 3rd, de Bakker PI, Word JM, Prisant MG, Richardson JS, Richardson DC (2003) Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins* 50:437-450.
128. 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.
129. Macias MJ, Gervais V, Civera C, Oschkinat H (2000) Structural analysis of WW domains and design of a WW prototype. *Nat Struct Biol* 7:375-379.

130. Macias MJ, Hyvonen M, Baraldi E, Schultz J, Sudol M, Saraste M, Oschkinat H (1996) Structure of the WW domain of a kinase-associated protein complexed with a proline-rich peptide. *Nature* 382:646-649.
131. Macias MJ, Wiesner S, Sudol M (2002) WW and SH3 domains, two different scaffolds to recognize proline-rich ligands. *FEBS Lett* 513:30-37.
132. Mah AS, Elia AE, Devgan G, Ptacek J, Schutkowski M, Snyder M, Yaffe MB, Deshaies RJ (2005) Substrate specificity analysis of protein kinase complex Dbf2-Mob1 by peptide library and proteome array screening. *BMC Biochem* 6:22.
133. Mahoney WM, Jr., Hong JH, Yaffe MB, Farrance IK (2005) The transcriptional co-activator TAZ interacts differentially with transcriptional enhancer factor-1 (TEF-1) family members. *Biochem J* 388:217-225.
134. Makhatadze GI (1998) Heat capacities of amino acids, peptides and proteins. *Biophys Chem* 71:133-156.
135. 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.
136. Makhatadze GI, Privalov PL (1995) Energetics of protein structure. *Adv Protein Chem* 47:307-425.
137. Manavalan P, Johnson WC, Jr. (1987) Variable selection method improves the prediction of protein secondary structure from circular dichroism spectra. *Anal Biochem* 167:76-85.
138. Martin-Serrano J, Perez-Caballero D, Bieniasz PD (2004) Context-dependent effects of L domains and ubiquitination on viral budding. *J Virol* 78:5554-5563.
139. Martin-Serrano J, Zang T, Bieniasz PD (2001) HIV-1 and Ebola virus encode small peptide motifs that recruit Tsg101 to sites of particle assembly to facilitate egress. *Nat Med* 7:1313-1319.
140. Martin-Sierra FM, Candel AM, Casares S, Filimonov VV, Martinez JC, Conejero-Lara F (2003) A binding event converted into a folding event. *FEBS Lett* 553:328-332.
141. Martinez JC, Cobos ES, Luque I, Ruiz-Sanz J. Differential scanning calorimetry: thermodynamic analysis of the unfolding transitions of proteins, domains and

- peptidic fragments by using equilibrium models. In: Walters EC, Ed. (2011) Protein folding. Nova Science Publishers Hauppauge, New York.
142. Martinez JC, el Harrou 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.
143. Martinez JC, Pisabarro MT, Serrano L (1998) Obligatory steps in protein folding and the conformational diversity of the transition state. *Nat Struct Biol* 5:721-729.
144. Martinez JC, Serrano L (1999) The folding transition state between SH3 domains is conformationally restricted and evolutionarily conserved. *Nat Struct Biol* 6:1010-1016.
145. Martinez JC, Viguera AR, Berisio R, Wilmanns M, Mateo PL, Filimonov VV, Serrano L (1999) Thermodynamic analysis of alpha-spectrin SH3 and two of its circular permutants with different loop lengths: discerning the reasons for rapid folding in proteins. *Biochemistry* 38:549-559.
146. Matallanas D, Romano D, Yee K, Meissl K, Kucerova L, Piazzolla D, Baccarini M, Vass JK, Kolch W, O'Neill E (2007) RASSF1A elicits apoptosis through an MST2 pathway directing proapoptotic transcription by the p73 tumor suppressor protein. *Mol Cell* 27:962-975.
147. Mateo PL, Baron C, Lopez-Mayorga O, Jimenez JS, Cortijo M (1984) AMP and IMP binding to glycogen phosphorylase b. A calorimetric and equilibrium dialysis study. *J Biol Chem* 259:9384-9389.
148. Mayer BJ (2001) SH3 domains: complexity in moderation. *J Cell Sci* 114:1253-1263.
149. Mayer BJ, Eck MJ (1995) SH3 domains. Minding your p's and q's. *Curr Biol* 5:364-367.
150. McPherson PS (1999) Regulatory role of SH3 domain-mediated protein-protein interactions in synaptic vesicle endocytosis. *Cell Signal* 11:229-238.
151. Meiyappan M, Birrane G, Ladas JA (2007) Structural basis for polyproline recognition by the FE65 WW domain. *J Mol Biol* 372:970-980.
152. Mohana-Borges R, Silva JL, Ruiz-Sanz J, de Prat-Gay G (1999) Folding of a pressure-denatured model protein. *Proc Natl Acad Sci U S A* 96:7888-7893.

153. Mohler J, Seecoomar M, Agarwal S, Bier E, Hsai J (2000) Activation of knot (kn) specifies the 3-4 intervein region in the *Drosophila* wing. *Development* 127:55-63.
154. Morales B, Ramirez-Espain X, Shaw AZ, Martin-Malpartida P, Yraola F, Sanchez-Tillo E, Farrera C, Celada A, Royo M, Macias MJ (2007) NMR structural studies of the ItchWW3 domain reveal that phosphorylation at T30 inhibits the interaction with PPxY-containing ligands. *Structure* 15:473-483.
155. Morel B, Varela L, Conejero-Lara F (2010) The thermodynamic stability of amyloid fibrils studied by differential scanning calorimetry. *J Phys Chem B* 114:4010-4019.
156. Mullis KB, Faloona FA (1987) Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction. *Methods Enzymol* 155:335-350.
157. Munoz V, Sanchez-Ruiz JM (2004) Exploring protein-folding ensembles: a variable-barrier model for the analysis of equilibrium unfolding experiments. *Proc Natl Acad Sci U S A* 101:17646-17651.
158. Naganathan AN, Li P, Perez-Jimenez R, Sanchez-Ruiz JM, Munoz V (2010) Navigating the downhill protein folding regime via structural homologues. *J Am Chem Soc* 132:11183-11190.
159. Naganathan AN, Perez-Jimenez R, Sanchez-Ruiz JM, Munoz V (2005) Robustness of downhill folding: guidelines for the analysis of equilibrium folding experiments on small proteins. *Biochemistry* 44:7435-7449.
160. Neduva V, Linding R, Su-Angrand I, Stark A, de Masi F, Gibson TJ, Lewis J, Serrano L, Russell RB (2005) Systematic discovery of new recognition peptides mediating protein interaction networks. *PLoS Biol* 3:e405.
161. Nguyen H, Jager M, Moretto A, Gruebele M, Kelly JW (2003) Tuning the free-energy landscape of a WW domain by temperature, mutation, and truncation. *Proc Natl Acad Sci U S A* 100:3948-3953.
162. Nicholson EM, Scholtz JM (1996) Conformational stability of the *Escherichia coli* HPr protein: test of the linear extrapolation method and a thermodynamic characterization of cold denaturation. *Biochemistry* 35:11369-11378.

163. Nojima H, Ikai A, Oshima T, Noda H (1977) Reversible thermal unfolding of thermostable phosphoglycerate kinase. Thermostability associated with mean zero enthalpy change. *J Mol Biol* 116:429-442.
164. Ohno H, Stewart J, Fournier MC, Bosshart H, Rhee I, Miyatake S, Saito T, Gallusser A, Kirchhausen T, Bonifacino JS (1995) Interaction of tyrosine-based sorting signals with clathrin-associated proteins. *Science* 269:1872-1875.
165. Oka T, Mazack V, Sudol M (2008) Mst2 and Lats kinases regulate apoptotic function of Yes kinase-associated protein (YAP). *J Biol Chem* 283:27534-27546.
166. Oka T, Remue E, Meerschaert K, Vanloo B, Boucherie C, Gfeller D, Bader GD, Sidhu SS, Vandekerckhove J, Gettemans J, Sudol M (2010) Functional complexes between YAP2 and ZO-2 are PDZ domain-dependent, and regulate YAP2 nuclear localization and signalling. *Biochem J* 432:461-472.
167. Okumura A, Pitha PM, Harty RN (2008) ISG15 inhibits Ebola VP40 VLP budding in an L-domain-dependent manner by blocking Nedd4 ligase activity. *Proc Natl Acad Sci U S A* 105:3974-3979.
168. Oliva FY, Munoz V (2004) A simple thermodynamic test to discriminate between two-state and downhill folding. *J Am Chem Soc* 126:8596-8597.
169. Omerovic J, Puggioni EM, Napoletano S, Visco V, Fraioli R, Frati L, Gulino A, Alimandi M (2004) Ligand-regulated association of ErbB-4 to the transcriptional co-activator YAP65 controls transcription at the nuclear level. *Exp Cell Res* 294:469-479.
170. Oneyama C, Nakano H, Sharma SV (2002) UCS15A, a novel small molecule, SH3 domain-mediated protein-protein interaction blocking drug. *Oncogene* 21:2037-2050.
171. Otte L, Wiedemann U, Schlegel B, Pires JR, Beyermann M, Schmieder P, Krause G, Volkmer-Engert R, Schneider-Mergener J, Oschkinat H (2003) WW domain sequence activity relationships identified using ligand recognition propensities of 42 WW domains. *Protein Sci* 12:491-500.
172. Pace CN (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol* 131:266-280.

173. Padmanabhan S, Laurents DV, Fernandez AM, Elias-Arnanz M, Ruiz-Sanz J, Mateo PL, Rico M, Filimonov VV (1999) Thermodynamic analysis of the structural stability of phage 434 Cro protein. *Biochemistry* 38:15536-15547.
174. Palencia A. Reconocimiento de secuencias ricas en prolina por dominios modulares de interacción proteína-proteína. (2008) Departamento de Química-Física Universidad de Granada, Granada.
175. Palencia A, Camara-Artigas A, Pisabarro MT, Martinez JC, Luque I (2010) Role of interfacial water molecules in proline-rich ligand recognition by the Src homology 3 domain of Abl. *J Biol Chem* 285:2823-2833.
176. Palencia A, Cobos ES, Mateo PL, Martinez JC, Luque I (2004) Thermodynamic dissection of the binding energetics of proline-rich peptides to the Abl-SH3 domain: implications for rational ligand design. *J Mol Biol* 336:527-537.
177. Parent LJ, Bennett RP, Craven RC, Nelle TD, Krishna NK, Bowzard JB, Wilson CB, Puffer BA, Montelaro RC, Wills JW (1995) Positionally independent and exchangeable late budding functions of the Rous sarcoma virus and human immunodeficiency virus Gag proteins. *J Virol* 69:5455-5460.
178. Park SJ, Seo MD, Lee SK, Ikeda M, Longnecker R, Lee BJ (2005) Expression and characterization of N-terminal domain of Epstein-Barr virus latent membrane protein 2A in *Escherichia coli*. *Protein Expr Purif* 41:9-17.
179. Pawson T (1995) Protein modules and signalling networks. *Nature* 373:573-580.
180. Pawson T, Raina M, Nash P (2002) Interaction domains: from simple binding events to complex cellular behavior. *FEBS Lett* 513:2-10.
181. Pawson T, Scott JD (1997) Signaling through scaffold, anchoring, and adaptor proteins. *Science* 278:2075-2080.
182. Pereboev AV, Ahmed N, thi Man N, Morris GE (2001) Epitopes in the interacting regions of beta-dystroglycan (PPxY motif) and dystrophin (WW domain). *Biochim Biophys Acta* 1527:54-60.
183. Petrovich M, Jonsson AL, Ferguson N, Daggett V, Fersht AR (2006) Phi-analysis at the experimental limits: mechanism of beta-hairpin formation. *J Mol Biol* 360:865-881.

184. Pires JR, Parthier C, Aido-Machado R, Wiedemann U, Otte L, Bohm G, Rudolph R, Oschkinat H (2005) Structural basis for APPTPPPLPP peptide recognition by the FBP11WW1 domain. *J Mol Biol* 348:399-408.
185. Pires JR, Taha-Nejad F, Toepert F, Ast T, Hoffmuller U, Schneider-Mergener J, Kuhne R, Macias MJ, Oschkinat H (2001) Solution structures of the YAP65 WW domain and the variant L30 K in complex with the peptides GTPPPPYTVG, N-(n-octyl)-GPPPY and PLPPY and the application of peptide libraries reveal a minimal binding epitope. *J Mol Biol* 314:1147-1156.
186. Plotnikov VV, Brandts JM, Lin LN, Brandts JF (1997) A new ultrasensitive scanning calorimeter. *Anal Biochem* 250:237-244.
187. Polo JM, Dubensky TW, Jr. (2002) Virus-based vectors for human vaccine applications. *Drug Discov Today* 7:719-727.
188. Porter CJ, Matthews JM, Mackay JP, Pursglove SE, Schmidberger JW, Leedman PJ, Pero SC, Krag DN, Wilce MC, Wilce JA (2007) Grb7 SH2 domain structure and interactions with a cyclic peptide inhibitor of cancer cell migration and proliferation. *BMC Struct Biol* 7:58.
189. Privalov PL (1979) Stability of proteins. Small globular proteins. *Adv Protein Chem* 33:167-241.
190. Privalov PL (1982) Stability of proteins. Proteins which do not present a single cooperative system. *Adv Protein Chem* 35:1-104.
191. Privalov PL (1990) Cold denaturation of proteins. *Crit Rev Biochem Mol Biol* 25:281-305.
192. Privalov PL, Dragan AI (2007) Microcalorimetry of biological macromolecules. *Biophys Chem* 126:16-24.
193. Privalov PL, Khechinashvili NN (1974) A thermodynamic approach to the problem of stabilization of globular protein structure: a calorimetric study. *J Mol Biol* 86:665-684.
194. 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.
195. Privalov PL, Potekhin SA (1986) Scanning microcalorimetry in studying temperature-induced changes in proteins. *Methods Enzymol* 131:4-51.

196. Puffer BA, Watkins SC, Montelaro RC (1998) Equine infectious anemia virus Gag polyprotein late domain specifically recruits cellular AP-2 adapter protein complexes during virion assembly. *J Virol* 72:10218-10221.
197. Ramirez-Espain X, Ruiz L, Martin-Malpartida P, Oschkinat H, Macias MJ (2007) Structural characterization of a new binding motif and a novel binding mode in group 2 WW domains. *J Mol Biol* 373:1255-1268.
198. Razvi A, Scholtz JM (2006) Lessons in stability from thermophilic proteins. *Protein Sci* 15:1569-1578.
199. Reid KL, Rodriguez HM, Hillier BJ, Gregoret LM (1998) Stability and folding properties of a model beta-sheet protein, *Escherichia coli* CspA. *Protein Sci* 7:470-479.
200. Reiss DJ, Schwikowski B (2004) Predicting protein-peptide interactions via a network-based motif sampler. *Bioinformatics* 20 Suppl 1:i274-282.
201. Renzoni DA, Pugh DJ, Siligardi G, Das P, Morton CJ, Rossi C, Waterfield MD, Campbell ID, Ladbury JE (1996) Structural and thermodynamic characterization of the interaction of the SH3 domain from Fyn with the proline-rich binding site on the p85 subunit of PI3-kinase. *Biochemistry* 35:15646-15653.
202. Robertson AD, Murphy KP (1997) Protein Structure and the Energetics of Protein Stability. *Chem Rev* 97:1251-1268.
203. Rotin D, Schild L (2008) ENaC and its regulatory proteins as drug targets for blood pressure control. *Curr Drug Targets* 9:709-716.
204. 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. *Eur J Biochem* 271:1497-1507.
205. Sadqi M, Casares S, Lopez-Mayorga O, Conejero-Lara F (2002) The temperature dependence of the hydrogen exchange in the SH3 domain of alpha-spectrin. *FEBS Lett* 527:86-90.
206. Sakurai A, Yasuda J, Takano H, Tanaka Y, Hatakeyama M, Shida H (2004) Regulation of human T-cell leukemia virus type 1 (HTLV-1) budding by ubiquitin ligase Nedd4. *Microbes Infect* 6:150-156.

207. Sanchez-Ruiz JM (1992) Theoretical analysis of Lumry-Eyring models in differential scanning calorimetry. *Biophys J* 61:921-935.
208. Sanchez-Ruiz JM (1995) Differential scanning calorimetry of proteins. *Subcell Biochem* 24:133-176.
209. Schleinkofer K, Wiedemann U, Otte L, Wang T, Krause G, Oschkinat H, Wade RC (2004) Comparative structural and energetic analysis of WW domain-peptide interactions. *J Mol Biol* 344:865-881.
210. Schlundt A, Sticht J, Piotukh K, Kosslick D, Jahnke N, Keller S, Schuemann M, Krause E, Freund C (2009) Proline-rich sequence recognition: II. Proteomics analysis of Tsg101 ubiquitin-E2-like variant (UEV) interactions. *Mol Cell Proteomics* 8:2474-2486.
211. Schuler B, Kremer W, Kalbitzer HR, Jaenicke R (2002) Role of entropy in protein thermostability: folding kinetics of a hyperthermophilic cold shock protein at high temperatures using <sup>19</sup>F NMR. *Biochemistry* 41:11670-11680.
212. Settanni G, Fersht AR (2009) Downhill versus barrier-limited folding of BBL 3. Heterogeneity of the native state of the BBL peripheral subunit binding domain and its implications for folding mechanisms. *J Mol Biol* 387:993-1001.
213. Sharpe T, Jonsson AL, Rutherford TJ, Daggett V, Fersht AR (2007) The role of the turn in beta-hairpin formation during WW domain folding. *Protein Sci* 16:2233-2239.
214. Shaw AZ, Martin-Malpartida P, Morales B, Yraola F, Royo M, Macias MJ (2005) Phosphorylation of either Ser16 or Thr30 does not disrupt the structure of the Itch E3 ubiquitin ligase third WW domain. *Proteins* 60:558-560.
215. Shi H, Asher C, Chigaev A, Yung Y, Reuveny E, Seger R, Garty H (2002) Interactions of beta and gamma ENaC with Nedd4 can be facilitated by an ERK-mediated phosphorylation. *J Biol Chem* 277:13539-13547.
216. Sigurskjold BW (2000) Exact analysis of competition ligand binding by displacement isothermal titration calorimetry. *Anal Biochem* 277:260-266.
217. Skorski T, Nieborowska-Skorska M, Wlodarski P, Wasik M, Trotta R, Kanakaraj P, Salomoni P, Antonyak M, Martinez R, Majewski M, Wong A, Perussia B, Calabretta B (1998) The SH3 domain contributes to BCR/ABL-dependent

- leukemogenesis in vivo: role in adhesion, invasion, and homing. *Blood* 91:406-418.
218. Socolich M, Lockless SW, Russ WP, Lee H, Gardner KH, Ranganathan R (2005) Evolutionary information for specifying a protein fold. *Nature* 437:512-518.
219. Sreerama N, Woody RW (2003) Structural composition of beta1- and beta11-proteins. *Protein Sci* 12:384-388.
220. Stauffer TP, Martenson CH, Rider JE, Kay BK, Meyer T (1997) Inhibition of Lyn function in mast cell activation by SH3 domain binding peptides. *Biochemistry* 36:9388-9394.
221. Stein R (1998) SH2 and SH3 domains. Unraveling signaling networks with peptide antagonists. *Methods Mol Biol* 88:187-195.
222. Strano S, Monti O, Pediconi N, Baccarini A, Fontemaggi G, Lapi E, Mantovani F, Damalas A, Citro G, Sacchi A, Del Sal G, Levrero M, Blandino G (2005) The transcriptional coactivator Yes-associated protein drives p73 gene-target specificity in response to DNA Damage. *Mol Cell* 18:447-459.
223. Strano S, Munarriz E, Rossi M, Castagnoli L, Shaul Y, Sacchi A, Oren M, Sudol M, Cesareni G, Blandino G (2001) Physical interaction with Yes-associated protein enhances p73 transcriptional activity. *J Biol Chem* 276:15164-15173.
224. Streicher WW, Makhatadze GI (2007) Unfolding thermodynamics of Trp-cage, a 20 residue miniprotein, studied by differential scanning calorimetry and circular dichroism spectroscopy. *Biochemistry* 46:2876-2880.
225. Sturtevant JM (1977) Heat capacity and entropy changes in processes involving proteins. *Proc Natl Acad Sci U S A* 74:2236-2240.
226. Sudarman E, Bollati-Fogolin M, Hafner M, Muller W, Scheller J, Rose-John S, Eichler J (2008) Synthetic mimetics of the gp130 binding site for viral interleukin-6 as inhibitors of the vIL-6-gp130 interaction. *Chem Biol Drug Des* 71:494-500.
227. Sudol M (1994) Yes-associated protein (YAP65) is a proline-rich phosphoprotein that binds to the SH3 domain of the Yes proto-oncogene product. *Oncogene* 9:2145-2152.

228. Sudol M (1996a) Structure and function of the WW domain. *Prog Biophys Mol Biol* 65:113-132.
229. Sudol M (1996b) The WW module competes with the SH3 domain? *Trends Biochem Sci* 21:161-163.
230. Sudol M, Chen HI, Bougeret C, Einbond A, Bork P (1995) Characterization of a novel protein-binding module--the WW domain. *FEBS Lett* 369:67-71.
231. Sudol M, Hunter T (2000) NeW wrinkles for an old domain. *Cell* 103:1001-1004.
232. Sudol M, Recinos CC, Abraczinskas J, Humbert J, Farooq A (2005) WW or WoW: the WW domains in a union of bliss. *IUBMB Life* 57:773-778.
233. Sudol M, Sliwa K, Russo T (2001) Functions of WW domains in the nucleus. *FEBS Lett* 490:190-195.
234. Sze CI, Su M, Pugazhenti S, Jambal P, Hsu LJ, Heath J, Schultz L, Chang NS (2004) Down-regulation of WW domain-containing oxidoreductase induces Tau phosphorylation in vitro. A potential role in Alzheimer's disease. *J Biol Chem* 279:30498-30506.
235. Tao W, Zhang S, Turenchalk GS, Stewart RA, St John MA, Chen W, Xu T (1999) Human homologue of the *Drosophila melanogaster* lats tumour suppressor modulates CDC2 activity. *Nat Genet* 21:177-181.
236. Tian Y, Kolb R, Hong JH, Carroll J, Li D, You J, Bronson R, Yaffe MB, Zhou J, Benjamin T (2007) TAZ promotes PC2 degradation through a SCFbeta-Trcp E3 ligase complex. *Mol Cell Biol* 27:6383-6395.
237. Timmins J, Schoehn G, Ricard-Blum S, Scianimanico S, Vernet T, Ruigrok RW, Weissenhorn W (2003) Ebola virus matrix protein VP40 interaction with human cellular factors Tsg101 and Nedd4. *J Mol Biol* 326:493-502.
238. Toepert F, Pires JR, Landgraf C, Oschkinat H, Schneider-Mergener J (2001) Synthesis of an Array Comprising 837 Variants of the hYAP WW Protein Domain This work was supported by the DFG (INK 16/B1-1), by the Fonds der Chemischen Industrie, and by the Universitätsklinikum Charite Berlin. *Angew Chem Int Ed Engl* 40:897-900.
239. Urata S, Yasuda J (2010) Regulation of Marburg virus (MARV) budding by Nedd4.1: a different WW domain of Nedd4.1 is critical for binding to MARV and Ebola virus VP40. *J Gen Virol* 91:228-234.

240. Varela L, Morel B, Azuaga AI, Conejero-Lara F (2009) A single mutation in an SH3 domain increases amyloid aggregation by accelerating nucleation, but not by destabilizing thermodynamically the native state. *FEBS Lett* 583:801-806.
241. Vassilev A, Kaneko KJ, Shu H, Zhao Y, DePamphilis ML (2001) TEAD/TEF transcription factors utilize the activation domain of YAP65, a Src/Yes-associated protein localized in the cytoplasm. *Genes Dev* 15:1229-1241.
242. Velazquez-Campoy A, Freire E (2006) Isothermal titration calorimetry to determine association constants for high-affinity ligands. *Nat Protoc* 1:186-191.
243. Velazquez Campoy A, Freire E (2005) ITC in the post-genomic era...? Priceless. *Biophys Chem* 115:115-124.
244. VerPlank L, Bouamr F, LaGrassa TJ, Agresta B, Kikonyogo A, Leis J, Carter CA (2001) Tsg101, a homologue of ubiquitin-conjugating (E2) enzymes, binds the L domain in HIV type 1 Pr55(Gag). *Proc Natl Acad Sci U S A* 98:7724-7729.
245. 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.
246. Wang C, Pawley NH, Nicholson LK (2001a) The role of backbone motions in ligand binding to the c-Src SH3 domain. *J Mol Biol* 313:873-887.
247. Wang K, Degerny C, Xu M, Yang XJ (2009) YAP, TAZ, and Yorkie: a conserved family of signal-responsive transcriptional coregulators in animal development and human disease. *Biochem Cell Biol* 87:77-91.
248. Wang NX, Lee HJ, Zheng JJ (2008) Therapeutic use of PDZ protein-protein interaction antagonism. *Drug News Perspect* 21:137-141.
249. Wang W, Lim WA, Jakalian A, Wang J, Luo R, Bayly CI, Kollman PA (2001b) An analysis of the interactions between the Sem-5 SH3 domain and its ligands using molecular dynamics, free energy calculations, and sequence analysis. *J Am Chem Soc* 123:3986-3994.
250. Warren JR, Gordon JA (1966) On the Restrictive Indices of Aqueous Solutions of Urea. *Journal of Physical Chemistry* 70:297-300.
251. Wiesner S, Stier G, Sattler M, Macias MJ (2002) Solution structure and ligand recognition of the WW domain pair of the yeast splicing factor Prp40. *J Mol Biol* 324:807-822.

252. Wiseman T, Williston S, Brandts JF, Lin LN (1989) Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal Biochem* 179:131-137.
253. Wishart DS, Sykes BD (1994) Chemical shifts as a tool for structure determination. *Methods Enzymol* 239:363-392.
254. Woody RW (1994) Contributions of tryptophan side chains to the far-ultraviolet circular dichroism of proteins. *Eur Biophys J* 23:253-262.
255. Wu J, Yang JT, Wu CS (1992) Beta-II conformation of all-beta proteins can be distinguished from unordered form by circular dichroism. *Anal Biochem* 200:359-364.
256. Wüthrich K. 1986. *NMR of Proteins and Nucleic Acids.*, New York.
257. Yagi R, Chen LF, Shigesada K, Murakami Y, Ito Y (1999) A WW domain-containing yes-associated protein (YAP) is a novel transcriptional co-activator. *EMBO J* 18:2551-2562.
258. Yan KS, Kuti M, Zhou MM (2002) PTB or not PTB -- that is the question. *FEBS Lett* 513:67-70.
259. Yasuda J, Hunter E (1998) A proline-rich motif (PPPY) in the Gag polyprotein of Mason-Pfizer monkey virus plays a maturation-independent role in virion release. *J Virol* 72:4095-4103.
260. Yasuda J, Nakao M, Kawaoka Y, Shida H (2003) Nedd4 regulates egress of Ebola virus-like particles from host cells. *J Virol* 77:9987-9992.
261. Yuan B, Campbell S, Bacharach E, Rein A, Goff SP (2000) Infectivity of Moloney murine leukemia virus defective in late assembly events is restored by late assembly domains of other retroviruses. *J Virol* 74:7250-7260.
262. Zarrinpar A, Bhattacharyya RP, Lim WA (2003a) The structure and function of proline recognition domains. *Sci STKE* 2003:RE8.
263. Zarrinpar A, Lim WA (2000) Converging on proline: the mechanism of WW domain peptide recognition. *Nat Struct Biol* 7:611-613.
264. Zarrinpar A, Park SH, Lim WA (2003b) Optimization of specificity in a cellular protein interaction network by negative selection. *Nature* 426:676-680.
265. Zhang L, Yue T, Jiang J (2009) Hippo signaling pathway and organ size control. *Fly (Austin)* 3:68-73.

266. Zhang YL, Zhang ZY (1998) Low-affinity binding determined by titration calorimetry using a high-affinity coupling ligand: a thermodynamic study of ligand binding to protein tyrosine phosphatase 1B. *Anal Biochem* 261:139-148.
267. Zhao B, Wei X, Li W, Udan RS, Yang Q, Kim J, Xie J, Ikenoue T, Yu J, Li L, Zheng P, Ye K, Chinnaiyan A, Halder G, Lai ZC, Guan KL (2007) Inactivation of YAP oncoprotein by the Hippo pathway is involved in cell contact inhibition and tissue growth control. *Genes Dev* 21:2747-2761.
268. Zhou YQ, Hall CK, Karplus M (1999) The calorimetric criterion for a two-state process revisited. *Protein Sci* 8:1064-1074.



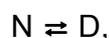
## APÉNDICES



## APÉNDICES

### APÉNDICE I. Modelo de equilibrio de dos estados utilizado para el análisis de las trazas calorimétricas de DSC

El modelo de dos estados considera que durante el proceso de desplegamiento de una proteína sólo existen dos estados poblados en un equilibrio químico, el estado nativo (N) y el estado desnaturalizado o desplegado (D):



A continuación se describe este modelo que ha sido el utilizado para el análisis de los termogramas que se presentan en esta Tesis. En este tipo de equilibrio, la constante de equilibrio aparente para el desplegamiento viene dada por la expresión:

$$K_D = \frac{[D]}{[N]} \quad (I.I)$$

La función de partición del sistema se define como:

$$q = 1 + K_D \quad (I.II)$$

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

$$x_N = \frac{1}{q} \quad (I.III)$$

$$x_D = \frac{K_D}{q} \quad (I.IV)$$

La entalpía molar parcial del sistema será:

$$\langle H \rangle = x_N \cdot H_N + x_D \cdot H_D = H_N + x_D \cdot \Delta H_D = H_N + \frac{K_D}{1+K_D} \cdot \Delta H_D \quad (I.V)$$

donde  $\Delta H_D$  es el cambio de entalpía de desplegamiento,  $\Delta H_D = H_D - H_N$ . De esta manera, la capacidad calorífica molar parcial,  $C_p$ , se define como:

$$C_p = \left[ \frac{\partial \langle H \rangle}{\partial T} \right]_P = C_{p,N} + \frac{\Delta H_D}{(1+K_D)^2} \left[ \frac{\partial K_D}{\partial T} \right]_P + \frac{K_D}{(1+K_D)} \left[ \frac{\partial \Delta H_D}{\partial T} \right]_P \quad (I.VI)$$

donde  $C_{p,N}$  es la capacidad calorífica molar parcial del estado nativo. Si ahora expresamos las dependencias de  $K_D$ ,  $\Delta H_D$  y  $\Delta S_D$  con la temperatura empleando las relaciones de Kirchhoff y van't Hoff:

$$\left[ \frac{\partial \ln K_D}{\partial T} \right]_P = \frac{\Delta H_D}{RT^2} = \frac{1}{K_D} \left[ \frac{\partial K_D}{\partial T} \right]_P \quad (I.VII)$$

$$\left[ \frac{\partial \Delta H_D}{\partial T} \right]_P = \Delta C_p \quad (I.VIII)$$

$$\left[ \frac{\partial \Delta S_D}{\partial T} \right]_P = \frac{\Delta C_p}{T} \quad (I.IX)$$

donde  $K_D = \exp\left(-\frac{\Delta G_D}{RT}\right)$ , y siendo  $\Delta G_D$  la diferencia de energía de Gibbs entre el estado desplegado y el nativo. Entonces, la ecuación 6.3.10 puede escribirse como:

$$C_p = C_{p,N} + \frac{(\Delta H_D)^2}{RT^2} \frac{K_D}{(1+K_D)^2} + x_D \Delta C_p \quad (I.X)$$

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. La función capacidad calorífica de una proteína desplegada,  $C_{p,D}(T)$ , puede describirse como un polinomio de segundo grado y puede estimarse a partir de su secuencia de aminoácidos (Makhatadze & Privalov 1990; Privalov & Makhatadze 1990). Por otro lado, la función que describe a la capacidad calorífica molar parcial del estado nativo,  $C_{p,N}(T)$ , es prácticamente lineal en el intervalo antes de la transición (Privalov 1989). Así las funciones  $C_{p,D}$ ,  $C_{p,N}$  y  $\Delta C_p$  pueden describirse de acuerdo con la siguientes expresiones:

$$C_{p,D} = a + b \cdot T + c \cdot T^2 \quad (I.XI)$$

$$C_{p,N} = d + e \cdot T \quad (I.XII)$$

$$\Delta C_p = (a - d) + (b - e) \cdot T + c \cdot T^2 \quad (I.XIII)$$

Para obtener las dependencias de  $\Delta H_D$  y  $\Delta S_D$  con la temperatura integramos las ecuaciones I.VIII y I.IX:

$$\Delta H_D(T) = \Delta H_m + \int_{T_m}^T \Delta C_p \cdot dT \quad (I.XIV)$$

$$\Delta S_D(T) = \Delta S_m + \int_{T_m}^T \frac{\Delta C_p}{T} \cdot dT \quad (I.XV)$$

donde  $\Delta H_m$  y  $\Delta S_m$  son los incrementos de entalpía y entropía de desplegamiento a la temperatura de la transición ( $T_m$ ), en la que  $x_D$  es 0.5,  $\Delta G_D$  se anula y por lo tanto  $K_D$  se hace igual a 1. En ese punto, el cambio de entropía,  $\Delta S_m$ , se puede obtener según la ecuación siguiente:

$$\Delta S_m = \frac{\Delta H_m}{T_m} \quad (I.XVI)$$

Si integramos las ecuaciones de  $\Delta H_D(T)$  y de  $\Delta S_D(T)$  con respecto a la temperatura, obtenemos:

$$\Delta H_D = \Delta H_m + (a - d) \cdot (T - T_m) + \frac{(b - e)}{2} \cdot (T^2 - T_m^2) + \frac{1}{3} \cdot c \cdot (T^3 - T_m^3) \quad (\text{I.XVII})$$

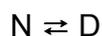
$$\Delta S_D = \frac{\Delta H_m}{T_m} + (a - d) \cdot \ln\left(\frac{T}{T_m}\right) + (b - e) \cdot (T - T_m) + \frac{1}{2} \cdot c \cdot (T^2 - T_m^2) \quad (\text{I.XVIII})$$

$$\Delta G_D = \Delta H_D - T \cdot \Delta S_D \quad (\text{I.XIX})$$

De acuerdo con este modelo, las curvas experimentales de  $C_p$  se analizaron mediante ajustes no lineales por mínimos cuadrados a la ecuación I.X, implementada en el programa SigmaPlot2000 versión 6.10 (*SPSS Inc*). Pero antes es preciso definir las magnitudes de las que depende  $C_p$  en dicha ecuación en función de la temperatura,  $\Delta H_D$ ,  $\Delta S_D$  y  $\Delta G_D$ , mediante las ecuaciones I.XVII, I.XVIII y I.XIX.

## APÉNDICE II. Modelo de equilibrio de dos estados utilizado para el análisis de la desnaturalización térmica seguida por CD

Las curvas de desplegamiento térmico se analizaron de acuerdo al modelo de equilibrio de dos estados. Para ello, se define el equilibrio entre el estado nativo (N) y el estado desplegado (D), de manera análoga a los datos de DSC en el apéndice I:



Es necesario definir las líneas base correspondientes a las funciones lineales del cambio de elipticidad de los estados nativo y desnaturalizado con la temperatura,  $\theta_N$  y  $\theta_D$ :

$$\theta_N = \theta_{N,0} + m_N(T - T_r) \quad (\text{II.I})$$

$$\theta_D = \theta_{D,0} + m_D(T - T_r) \quad (\text{II.II})$$

donde  $T_r$  es la temperatura de referencia (25 °C),  $\theta_{N,0}$  y  $m_N$ , por una parte, y  $\theta_{D,0}$  y  $m_D$  por otra, representan las respectivas ordenadas en el origen y pendientes. La elipticidad total,  $\theta$ , puede expresarse de acuerdo con la una combinación lineal de las funciones de elipticidad de los estados nativo y desplegado:

$$\theta = \theta_N X_N + \theta_D X_D \quad (\text{II.III})$$

en la que, las fracciones  $X_N$  y  $X_D$ , están determinadas por el valor de la constante de equilibrio,  $K_D$ , entre los dos estados definidos. Sustituyendo, se obtiene la siguiente expresión para la elipticidad:

$$\theta = \theta_N \frac{1}{1 + K_D} + \theta_D \frac{K_D}{1 + K_D} = \frac{(\theta_N + \theta_D K_D)}{1 + K_D} \quad (\text{II.IV})$$

A su vez, la  $K_D$  está determinada por la diferencia de energía libre de Gibbs entre el estado desnaturalizado y el nativo  $\Delta G_D(T)$ :

$$K_D = \exp\left(-\frac{\Delta G_D(T)}{RT}\right) \quad (\text{II.V})$$

donde  $R$  es la constante universal de los gases ideales. Por otra parte, los cambios en entalpía,  $\Delta H_D$ , y entropía,  $\Delta S_D$ , asociados al proceso de desplegamiento se pueden escribir mediante las siguientes ecuaciones que tienen en cuenta su dependencia con la temperatura,

$$\Delta H_D(T) = \Delta H_m + \Delta C_{p,D}(T - T_m) \quad (\text{II.VI})$$

$$\Delta S_D(T) = \Delta S_m + \Delta C_{p,D} \ln\left(\frac{T}{T_m}\right) \quad (\text{II.VII})$$

donde  $T_m$  es la temperatura de transición, en la que la población del estado nativo es igual que la población del estado desnaturalizado,  $\Delta H_m$  y  $\Delta S_m$  son los cambios de entalpía y entropía a la temperatura de la transición, y  $\Delta C_{p,D}$  es el cambio de la capacidad calorífica del proceso de desnaturalización que, como aproximación más sencilla, puede considerarse independiente de la temperatura y de la concentración de urea. Teniendo en cuenta que  $\Delta S_m = \Delta H_m / T_m$ , y sustituyendo las funciones de entalpía y entropía de desplegamiento en la ecuación general  $\Delta G_D = \Delta H_D(T) - T \cdot \Delta S_D(T)$ , se obtiene la siguiente expresión para la energía libre de Gibbs:

$$\Delta G_D(T) = \Delta H_m \left(1 - \frac{1}{T_m}\right) - \Delta C_{p,D} \left[ T_m - T + T \ln\left(\frac{T}{T_m}\right) \right] \quad (\text{II.VIII})$$

Finalmente, sustituyendo todos los términos de la ecuación II.IV para expresarlo en función del menor número posible de parámetros haciendo uso de las expresiones II.I, II.II, II.V y II.VIII, obtenemos la siguiente ecuación para la elipticidad:

$$\theta = \frac{[\theta_{N,0} + m_N(T - T_r)] + [\theta_{D,0} + m_D(T - T_r)] \exp\left\{-\left(\Delta H_m\left(1 - \frac{1}{T_m}\right) - \Delta C_{p,D}\left[T_m - T + T \ln\left(\frac{T}{T_m}\right)\right]\right)\right\} / R \cdot T}{1 + \exp\left\{-\left(\Delta H_m\left(1 - \frac{1}{T_m}\right) - \Delta C_{p,D}\left[T_m - T + T \ln\left(\frac{T}{T_m}\right)\right]\right)\right\} / R \cdot T}$$

(II.IX)

De acuerdo con este modelo, las curvas experimentales de  $\theta$  se analizaron mediante ajustes no lineales por mínimos cuadrados a la ecuación II.IX, implementada en el programa SigmaPlot2000 versión 6.10 (SPSS Inc). Los parámetros que se obtienen de estos ajustes son  $\theta_{N,0}$ ,  $m_N$ ,  $\theta_{D,0}$ ,  $m_D$ ,  $T_m$ ,  $\Delta H_m$  y  $\Delta C_{p,D}$ .

### APÉNDICE III. Modelo de unión de un ligando a una macromolécula con “n” sitios idénticos e independientes para ITC

El término “idénticos” implica la misma constante de unión microscópica a cualquiera de los “n” sitios, y el término “independientes” que la unión a un sitio cualquiera no modifica la afinidad de los restantes. La formulación y el desarrollo matemático correspondiente para este modelo de unión se detallan a continuación:



Función de partición de unión (Z) de los diferentes estados moleculares accesibles del sistema:

$$Z = (1 + K \cdot [L])^n \quad (\text{III.I})$$

donde  $K$  es la constante de equilibrio por sitio de unión,  $[L]$  la concentración de ligando libre y  $n$  es el número de sitios de unión de la macromolécula para el ligando.

Parámetro de unión ( $\bar{\nu}$ ) o relación de la concentración de ligando unido,  $[L]_b$ , a la concentración de macromolécula total,  $[M]_T$ , está relacionada con la función de partición de unión por:

$$\bar{\nu} = \frac{[L]_b}{[M]_T} = \frac{\partial \ln Z}{\partial \ln [L]} = \frac{n \cdot K \cdot [L]}{1 + K \cdot [L]} \quad (\text{III.II})$$

Calor liberado o absorbido en una inyección cualquiera “i”:

$$q_i = \Delta H_{ap} \left( \frac{kJ}{mol L_b} \right) \Delta(\text{moles } L_b) \quad (\text{III.III})$$

donde  $\Delta H_{ap}$  es el cambio de entalpía aparente por mol de ligando unido. Entonces, se puede escribir como:

$$q_i = \Delta H_{ap} \cdot V_c \cdot ([L]_{b,i} - [L]_{b,i-1}) = \Delta H_{ap} \cdot V_c \cdot (\bar{v}_i \cdot [M]_{T,i} - \bar{v}_{i-1} \cdot [M]_{T,i-1}) \quad (\text{III.IV})$$

donde  $V_c$  representa el volumen de la célula o volumen inicial de la muestra y  $[M]_T$  la concentración de proteína en la célula. Considerando el volumen efectivo de la célula, el volumen de inyección  $V_{in}$  y la concentración en jeringa del ligando  $[L]_0$ , se corrigen las concentraciones de macromolécula y ligando,  $[M]_{T,i}$  y  $[L]_{T,i}$ , mediante las ecuaciones:

$$[M]_{T,i} = [M]_{T,i-1} \frac{V_c - V_{in}}{V_c} \quad (\text{III.V})$$

$$[L]_{T,i} = \frac{(V_c - V_{in}) \cdot [L]_{T,i-1} + V_{in} \cdot [L]_0}{V_c} \quad (\text{III.VI})$$

Calor total acumulado después de N inyecciones:

$$Q = \sum_{i=1}^N q_i = V_c \cdot [M]_T \cdot \Delta H_{ap} \cdot \bar{v}_N = V_c \cdot [M]_T \cdot \Delta H_{ap} \cdot \frac{n \cdot K \cdot [L]}{1 + K \cdot [L]} \quad (\text{III.VII})$$

Durante el experimento el valor de la variable  $[L]$  no es conocida por lo que operativamente conviene expresarla en función de las variables experimentales  $[L]_T$  y  $Q$ :

$$[L] = [L]_T - [L]_b = [L]_T - \frac{Q}{V_c \cdot \Delta H_{ap}} \quad (\text{III.VIII})$$

Sustituyendo la ecuación anterior en la ecuación II.VII, y despejando  $Q$  se obtiene una ecuación de segundo grado, cuya solución es:

$$Q = \frac{V_c \cdot \Delta H_{ap}}{2 \cdot K} \left[ 1 + K \cdot [L]_T + n \cdot K \cdot [M]_T - \sqrt{(1 + K \cdot [L]_T + n \cdot K \cdot [M]_T)^2 - 4 \cdot n \cdot K^2 \cdot [M]_T \cdot [L]_T} \right] \quad (\text{III.IX})$$

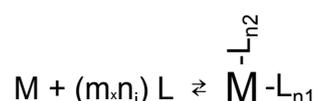
Derivando esta expresión respecto a  $[L]_T$  y escribiendo el cociente  $[L]_T/[M]_T$  como  $\theta$ , obtendremos una expresión para el calor por mol de ligando añadido asociado a cada inyección:

$$\frac{1}{V_c} \cdot \frac{dQ}{d[L]_T} \approx \frac{1}{V_c} \cdot \frac{\Delta Q}{\Delta[L]_T} = \frac{\Delta H_{ap}}{2} \left[ 1 - \frac{1 + [M]_T \cdot \theta - n \cdot K \cdot [M]_T}{\sqrt{(1 + K \cdot [M]_T \cdot \theta + n \cdot K \cdot [M]_T)^2 - 4 \cdot n \cdot K^2 \cdot [M]_T^2 \cdot \theta}} \right]$$

(III.X)

**APÉNDICE IV. Modelo de unión de un ligando a una macromolécula con “m” clases de sitios diferentes e independientes para ITC**

Cada clase de sitios de unión se definen como sitios diferentes (no idénticos) e independiente de las otras clases para el mismo ligando. El término “sitios diferentes” implica una constante microscópica en el equilibrio para cada clase sitios, y el término “independiente” que la unión a un sitio cualquiera no modifica la afinidad de los restantes. La formulación y el desarrollo matemático correspondiente para este modelo de unión, considerando que sólo existen dos clases diferentes de sitios ( $m = 2$ ),  $n_1$  y  $n_2$ , se describe a continuación (Brandts et al. 1990):



Parámetro de unión ( $\bar{\nu}$ ) o relación de la concentración de ligando unido en cada sitio,  $[L]_{b,i}$ , respecto a la concentración de macromolécula total,  $[M]_T$ :

$$\bar{\nu} = \sum_{i=1}^{m-2} \nu_i = \frac{[L]_{b,i}}{[M]_T} = \sum_{i=1}^{m-2} \frac{n_i \cdot K_i \cdot [L]}{1 + K_i \cdot [L]} \quad (IV.I)$$

Calor liberado o absorbido en una inyección cualquiera “i”:

$$q_j = \sum_{i=1}^{m-2} \Delta H_{ap,i} \left( \frac{kJ}{mol L_{b,i}} \right) \cdot \Delta(moles L_{b,i}) \quad (IV.II)$$

donde  $\Delta H_{ap,i}$  es el cambio de entalpía aparente por mol de ligando unido para una clase de sitio, i. Entonces, se puede escribir como:

$$q_j = \sum_{i=1}^{m-2} \Delta H_{ap,i} \cdot V_c \cdot ([L]_{b,i,j} - [L]_{b,i,j-1}) = \sum_{i=1}^{m-2} \Delta H_{ap,i} \cdot V_c \cdot (\bar{v}_j \cdot [M]_{T,j} - \bar{v}_{i-1} \cdot [M]_{T,j-1}) \quad (IV.III)$$

donde  $V_c$  representa el volumen de la célula o volumen inicial de la muestra y  $[M]_{T,j}$  la concentración de proteína en la célula en la inyección "j". Considerando el volumen efectivo de la célula, el volumen de inyección  $V_{in}$  y la concentración en jeringa del ligando  $[L]_0$ , entonces se corrigen la concentraciones de macromolécula,  $[M]_{T,j}$ , y ligando,  $[L]_{T,i}$ , mediante las siguientes ecuaciones:

$$[M]_{T,j} = [M]_{T,j-1} \frac{V_c - V_{in}}{V_c} \quad (IV.IV)$$

$$[L]_{T,i} = \frac{(V_c - V_{in}) \cdot [L]_{T,i-1} + V_{in} \cdot [L]_0}{V_c} \quad (IV.V)$$

Calor total acumulado después de N inyecciones:

$$Q(N) = \sum_{j=1}^N q_j = V_c \cdot [M]_T \cdot \sum_{i=1}^{m-2} \Delta H_{ap,i} \cdot \bar{v}_{i,N} = V_c \cdot [M]_T \cdot \sum_{i=1}^{m-2} \Delta H_{ap,i} \frac{n_i \cdot K_i \cdot [L]}{1 + K_i \cdot [L]} \quad (IV.VI)$$

$$Q(N) = V_c \cdot [M]_T \left[ \Delta H_{ap,1} \frac{n_1 \cdot K_1 \cdot [L]}{1 + K_1 \cdot [L]} + \Delta H_{ap,2} \frac{n_2 \cdot K_2 \cdot [L]}{1 + K_2 \cdot [L]} \right] \quad (IV.VII)$$

Durante el experimento el valor de la variable  $[L]$  no es conocida por lo que operativamente conviene expresarla en función de las variables experimentales ligando unido total  $[L]_{b,T} = [\bar{v}_1 + \bar{v}_2] \cdot [M]_T$  y el calor total acumulado  $Q(N)$ :

$$[L] = [L]_T - [L]_{b,T} = [L]_T - [M]_T \cdot \left[ \frac{n_1 \cdot k_1 \cdot [L]}{1 + k_1 \cdot [L]} + \frac{n_2 \cdot k_2 \cdot [L]}{1 + k_2 \cdot [L]} \right] \quad (IV.VIII)$$

Sustituyendo y ordenando se obtiene la siguiente ecuación cúbica:

$$[L]^3 + a_2[L]^2 + a_1[L] + a_0 = 0 \quad (IV.IX)$$

donde se definen:  $a_0 = -\frac{[L_T]}{k_1 k_2}$ ,  $a_1 = \left(\frac{n_1}{k_2} + \frac{n_2}{k_1}\right)[M_T] - \left(\frac{1}{k_1} + \frac{1}{k_2}\right)[L_T] + \frac{1}{k_1 k_2}$  y

$a_2 = \frac{1}{k_1} + \frac{1}{k_2} + (n_1 + n_2)[M_T] - [L_T]$ ; y cuya solución es:

$$[L] = \sqrt[3]{A + \sqrt{A^2 + B^3}} + \sqrt[3]{A - \sqrt{A^2 + B^3}} + C \quad (IV.X)$$

ahora, las expresiones A, B y C agrupan de nuevo los términos a, b y c como:

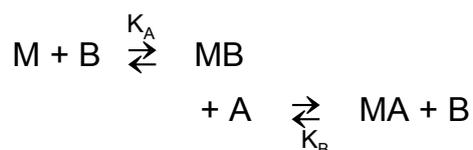
$$A = \frac{-a_2^3}{27} + \frac{a_1 a_2}{6} - \frac{a_0}{2}, \quad B = \left(\frac{a_1}{3} - \frac{a_2^2}{9}\right) \text{ y } C = -\frac{a_2}{3}.$$

La solución de la ecuación cúbica IV.X nos permite calcular el calor total acumulado tras un número dado de inyecciones. Sustituyendo en la ecuación IV.VII y haciendo uso de la expresión propuesta que incluye la corrección del calor generado por la fracción de volumen desplazado en cada inyección, se calculan los calores asociados a cada inyección:

$$\Delta Q(j) = Q_T(j) - Q(j-1) + \frac{V_{in}}{V_c} \left( \frac{Q_T(j) + Q_T(j-1)}{2} \right) \quad (IV.XI)$$

**APÉNDICE V. Modelo de unión a una macromolécula (con un solo sitio de unión) de un ligando A con afinidad moderada por desplazamiento de otro B de afinidad baja**

Para formular este modelo de unión por desplazamiento, se suponen dos ligandos, A y B, que pueden unir al mismo sitio de unión de una proteína, M, con diferentes constantes de afinidad ( $K_A \gg K_B$ ), siguiendo el esquema:



Es necesario asumir que la unión de dos ligandos, A y B, se produce en el mismo sitio de unión de la macromolécula, M. Entonces, se consideran los equilibrios de unión,  $M + A \rightleftharpoons MA$  y  $M + B \rightleftharpoons MB$ , con sus respectivas constantes de asociación:

$$K_A = \frac{[MA]}{[M][A]} \quad (V.I)$$

$$K_B = \frac{[MB]}{[M][B]} \quad (V.II)$$

Sustituyendo en función de las constantes iniciales,  $[A]_0 = [MA] + [A]$  y  $[B]_0 = [MB] + [B]$ , se obtienen:

$$[MA] = \frac{[M][A]_0}{1/K_A + [M]} \quad (V.III)$$

$$[MB] = \frac{[M][B]_0}{1/K_B + [M]} \quad (V.IV)$$

Considerando que  $K_A \gg K_B$ , durante la titulación del ligando A sobre la macromolécula M en presencia del ligando B, el ligando A desplazará parcialmente al ligando B del complejo MB dependiendo de los valores de sus constantes y de  $[A]_0$  y  $[B]_0$ .

Primero, se definen las fracciones molares de todas las especies que contengan a la macromolécula como:  $x_M = [M]/[M]_T$ ,  $x_{MA} = [MA]/[M]_T$ , y  $x_{MB} = [MB]/[M]_T$ , donde  $[M]_T$  es la concentración de macromolécula total. Si además, se escriben las relaciones molares entre las cantidades iniciales de A y B respecto de las concentración de macromolécula total,  $r_A = [A]_0/[M]_T$  y  $r_B = [B]_0/[M]_T$ . Y por último, si se expresan las concentraciones de A y B durante la interacción como productos de las constantes de asociación  $c_A = K_A [M]_T$  y  $c_B = K_B [M]_T$ . Entonces, sustituyendo estas expresiones en las ecuaciones V.III y V.IV, se llega a las siguientes expresiones para todas las especies que forma la macromolécula:

$$x_M + x_{MA} + x_{MB} = 1 \quad (\text{V.V})$$

$$x_{MA} = \frac{r_A \cdot x_M}{1/c_A + x_M} \quad (\text{V.VI})$$

$$x_{MB} = \frac{r_B \cdot x_M}{1/c_B + x_M} \quad (\text{V.VII})$$

Sustituyendo las ecuaciones V.VI y V.VII en la ecuación V.V; y ordenando, se obtiene la siguiente ecuación cúbica:

$$x_M^3 + a \cdot x_M^2 + b \cdot x_M + c = 0 \quad (\text{V.VIII})$$

donde se definen:  $a = \frac{1}{c_A} + \frac{1}{c_B} + r_A + r_B - 1$ ,  $b = \frac{r_A - 1}{c_A} + \frac{r_B - 1}{c_B} + \frac{1}{c_A \cdot c_B}$  y

$c = -\frac{1}{c_A \cdot c_B}$ . Si se resuelve la ecuación cúbica, se encuentra la siguiente solución

real:

$$x_M = \frac{2 \cdot (\sqrt{a^2 - 3b}) \cdot \cos(\theta/3) - a}{3} \quad (\text{V.IX})$$

En la cual,  $\theta$  se puede escribir como:

$$\theta = \arccos \frac{-2 \cdot a^3 - 9 \cdot a \cdot b - 27 \cdot c}{2 \cdot (\sqrt{a^2 - 3 \cdot b})^3} \quad (\text{V.X})$$

Una vez se determina la fracción molar de macromolécula libre,  $x_M$ , también se conocen las fracciones molares de las otras especies que forma la macromolécula,  $x_{MA}$  y  $x_{MB}$ , despejando las ecuaciones V.VI y V.VII.

Calor liberado o absorbido en una inyección: El calor liberado tras cada inyección será proporcional a los cambios en [MA] y [MB] y sus respectivas entalpías molares de unión. Por tanto, se puede escribir como:

$$\Delta Q = V_c \cdot (\Delta H_A \cdot \Delta[MA] + \Delta H_B \cdot \Delta[MB]) = V_c \cdot [M]_T \cdot (\Delta H_A \cdot \Delta x_{MA} + \Delta H_B \cdot \Delta x_{MB}) \quad (\text{V.XI})$$

siendo  $V_c$  el volumen de la célula. Para corregir las concentraciones de macromolécula y ligandos, se define el siguiente cambio infinitesimal:

$$d[X] = -\frac{dV_i}{V_0} \cdot [X] \quad (\text{V.XII})$$

donde [X] representa la concentración de cualquiera de las especies. De la integración de la expresión anterior entre los límites:  $[X]_i$  y  $[X]_{i-1}$ ; y  $V_0$  y  $V_i$ ; resulta:

$$[X]_i = [X]_{i-1} \exp\left(-\frac{V_i}{V_0}\right) = [X]_{i-1} f_i \quad (\text{V.XIII})$$

donde  $f_i$ , es el factor de dilución que permiten definir las relaciones molares tras cada inyección como  $[A]_i = [A]_{i-1}(1 - f_i)$ ;  $[M]_i = f_i [M]_{i-1}$ ; y  $[B]_i = f_i [B]_{i-1}$ . Ahora el calor liberado tras cada inyección puede expresarse como:

$$\Delta Q_i = V_c \cdot [M]_T \cdot \left\{ \Delta H_A \cdot (\Delta x_{MA,i} - f_i \cdot x_{MA,i-1}) + \Delta H_B \cdot (\Delta x_{MB,i} - f_i \cdot x_{MB,i-1}) \right\} \quad \text{V.XIV}$$