

**UNIVERSIDAD DE GRANADA**

**FACULTAD DE CIENCIAS**



**DEPARTAMENTO DE QUÍMICA FÍSICA**

**RECONOCIMIENTO DE SECUENCIAS RICAS EN PROLINA  
POR DOMINIOS MODULARES DE INTERACCIÓN  
PROTEÍNA-PROTEÍNA**

**ANDRÉS PALENCIA CARRILERO**

**Tesis Doctoral  
GRANADA  
2008**

Editor: Editorial de la Universidad de Granada  
Autor: Andrés Palencia Carrilero  
D.L.: GR. 1206-2008  
ISBN: 978-84-691-4092-5



UNIVERSIDAD DE GRANADA  
FACULTAD DE CIENCIAS  
DEPARTAMENTO DE QUÍMICA FÍSICA

**Memoria presentada para aspirar al grado de Doctor en Química**

Fdo: **Andrés Palencia Carrilero**

**Licenciado en Química por la Universidad de Granada**

Granada, a 28 de Marzo de 2008

VºBº  
DIRECTORES DE LA TESIS

Fdo: **Irene Luque Fernández**  
Profesora Contratada Doctora  
Departamento de Química Física  
Universidad de Granada

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



## AGRADECIMIENTOS

En primer lugar me gustaría agradecer a mis padres, Emilio y Carmen, su amor, su fe y su ejemplo, sin ellos no hubiera sido posible alcanzar esta meta. Agradezco a mis hermanos, Herminio y Mari Carmen, su apoyo en todo momento, aunque es muy fácil para un hermano decir esto, sinceramente, creo que ellos son de las personas más maravillosas que conozco.

A Jose, el haberme introducido en el mundo de la investigación y por ofrecerme la posibilidad de aprender tanto en este grupo. A Irene, el haberme acogido como su primer becario y por ofrecerme sus conocimientos en todo momento, ha sido mucho lo que he aprendido de ti, no sólo ciencia, sino también otros menesteres de la vida. Ambos son, de entre las personas jóvenes que se dedican a la ciencia, de las más brillantes que he conocido.

A mi compañera, amiga, maestra y no se que más ... Adela, simplemente GRACIAS. Al resto de compañeros de departamento, algunos, que me ayudaron mucho al principio, como Eva y Salva, porque la ayuda en los primeros pasos favorece el caminar firme en el futuro; a Bertrand, compañero de trabajo, hogar, alegría y lamentos, gracias por todo; a Carles, Jose Manuel, Jose Luis, Lorena, Sisi, Javi, Raquel, Raúl, David y otros muchos más, de todos vosotros he aprendido algo, y gracias por ser siempre tan maravillosos conmigo.

A todos mis amigos RODENSES: Berna, Chema, David, Fran, Josete, Juanjo, Juanra, Matías, Miguel Angel, Oscar, Paquito, Paulino, Rafa, Ruben, aunque alguno no os lo puedo agradecer ahora, ya tendré ocasión. A todos los que he conocido en "Granada", Jose Antonio, Mariano y Nico la palabra amigo es muy pobre para expresar todo lo que siento por vosotros, y a otros muchos más que por problemas de espacio es difícil de citar, GRACIAS.

A Pedro Luis Mateo, por ofrecerme su ayuda y opinión tan acertada en todo momento, creo que eres una de las personas más sabias que he conocido.

A Ana Cámara, por haberme introducido en el mundo de la Cristalografía, gracias por haber hecho que me guste tanto esta disciplina de la que me queda aún todo por aprender. A Jose Antonio Gavira por su generosa ayuda en todo momento y a Juan Manuel García por ofrecernos la posibilidad de medir los cristales a "tantos kilómetros" de distancia.

A Mayte, por haberme dado la oportunidad de trabajar en su grupo de investigación y concederme tanta libertad para elegir como hacerlo, y a todos los miembros de su grupo, me ha gustado mucho lo mucho que me habéis enseñado, aunque me habéis creado serias dudas para elegir mi tema de investigación Postdoctoral. A mucha gente en Dresden, especialmente a Mónica y Joan, y a todos los que conocí por ellos, GRACIAS por haberme hecho sentir como en casa.

Por último, pero no menos importante, a Claudia, GRACIAS, por se como eres, por haberme animado/aguantado tanto, por lo feliz que he sido durante estos meses y por haber hecho que todo esto no fuera tan difícil.



***A mis padres***

***A mis hermanos***



*“La suerte no dura demasiado tiempo  
porque no depende de ti,  
la Buena Suerte la crea uno mismo,  
por eso dura siempre”.*



# ÍNDICE



<b>CAPÍTULO 1: INTRODUCCIÓN</b>	<b>1</b>
<b>1.1. ANTECEDENTES</b>	<b>7</b>
1.1.2. Dominios modulares para el reconocimiento de secuencias ricas en prolina	9
1.1.3. Dominios SH3	15
1.1.3.1. Dominio SH3 de Abl ( <i>Abelson leukemia virus</i> )	23
1.1.4. Dominios WW	25
1.1.5. Dominios UEV ( <i>Ubiquitin E2 Variant</i> )	33
1.1.6. Dominios L víricos	36
<b>CAPÍTULO 2: MATERIALES Y MÉTODOS</b>	<b>45</b>
<b>2.1. MATERIALES</b>	<b>47</b>
<b>2.2. INSTRUMENTACIÓN</b>	<b>48</b>
<b>2.3. MÉTODOS</b>	<b>51</b>
2.3.1. Clonado, expresión y purificación del dominio SH3 de Abl ( <i>virus leucémico de Abelson</i> )	51
2.3.2. Obtención de variantes del dominio SH3 de Abl por mutagénesis dirigida	52
2.3.3. Expresión y purificación del dominio UEV de Tsg101 ( <i>Tumor Susceptibility Gene 101</i> )	53
2.3.4. Clonado, expresión y purificación de dominios WW (WW3 y WW4 de Nedd4 humana y WW de Yap65 humana)	54
2.3.4.1. Clonado <i>de novo</i> de dominios WW	54
2.3.4.2. Expresión y purificación de dominios WW	58
2.3.5. Preparación de disoluciones tampón	59
2.3.6. Diálisis de proteínas	60
2.3.7. Determinación de las concentraciones de proteína y ligando	60
2.3.8. Espectroscopia de fluorescencia	61
2.3.8.1. Introducción	61
2.3.8.2. Obtención de espectros de fluorescencia	62
2.3.8.3. Análisis de los espectros de fluorescencia	63
2.3.9. Calorimetría isotérmica de titulación	64
2.3.9.1. Introducción	64
2.3.9.2. Diseño del experimento de ITC	69
2.3.9.3. Formulación y análisis de un experimento directo de ITC	70
2.3.9.4. Formulación y análisis de un experimento de competición	75
2.3.9.5. Interpretación de los parámetros termodinámicos obtenidos	80
2.3.10. Calorimetría diferencial de barrido	83
2.3.10.1. Diseño del experimento de calorimetría diferencial de barrido	84
2.3.10.2. Cálculo de la capacidad calorífica molar parcial de la	

proteína	85
2.3.10.3. Análisis de las trazas calorimétricas según la termodinámica de equilibrio	86
2.3.10.3.1. Modelo de dos estados	86
2.3.11. Dicroísmo circular	90
2.3.11.1. Espectros de dicroísmo circular de proteínas	92
2.3.11.2. Diseño y ejecución de los experimentos de CD	96
2.3.12. Cristalización de proteínas	96
2.3.12.1. Introducción	96
2.3.12.2. Preparación de las muestras de cristalización	97
2.3.12.3. Técnicas de cristalización: difusión de vapor y contradifusión en capilares	98
2.3.12.4. Seguimiento del crecimiento cristalino	101
2.3.12.5. Metodología utilizada para la difracción de rayos X	102
2.3.13. Dinámica molecular de proteínas	104
2.3.13.1. Formulación de las simulaciones de dinámica molecular	106
2.3.13.2. Simulaciones de dinámica molecular para los complejos del dominio SH3 de Abl, y los mutantes N114A y N114Q con el ligando p41	110
2.3.13.3. Análisis y visualización de trayectorias de dinámica molecular	115
<b>CAPÍTULO 3: RESULTADOS</b>	<b>121</b>
<b>3.1. ESTUDIO DEL RECONOCIMIENTO DE LIGANDOS RICOS EN PROLINA POR EL DOMINIO SH3 DE ABL</b>	<b>123</b>
3.1.1. Análisis termodinámico-estructural de la interacción de Abl-SH3 con el ligando de diseño p41	123
3.1.1.1. La interacción de Abl-SH3 con p41 presenta un patrón termodinámico incongruente con la elevada hidrofobicidad de la interfase de unión	123
3.1.1.2. La termodinámica de unión de Abl-SH3 con p41 no refleja la influencia de equilibrios de ionización acoplados	127
3.1.1.3. No existen cambios conformacionales significativos asociados a la unión de Abl-SH3 y p41	130
3.1.1.4. La interacción entre Abl-SH3 y p41 está mediada por un conjunto de moléculas de agua ocluidas en la interfase de unión	132
3.1.1.5. Diseño de mutaciones conservativas en Abl-SH3 para perturbar la red de interacciones mediadas por moléculas de agua	139
3.1.1.6. Las mutaciones en posición N114 no perturban las propiedades conformacionales del dominio Abl-SH3	140
3.1.1.7. La sustitución de las cadenas laterales que coordinan las moléculas de agua interfaciales produce cambios significativos en la energética de unión	146
3.1.1.8. Las mutaciones en el resto N114 producen efectos locales, conservando la estructura e interacciones del dominio silvestre	151

3.1.1.9. Cuatro de las cinco moléculas de agua interfaciales observadas en Abl-SH3 se conservan en las estructuras cristalográficas de los mutantes N114 unidos a p41	160
3.1.1.10. La dinámica molecular de las moléculas de agua interfaciales permite explicar los efectos termodinámicos observados en la interacción de p41 con Abl-SH3 y sus mutantes N114A y N114Q	167
3.1.2. Análisis termodinámico-estructural de la interacción de Abl-SH3 con derivados mono-sustituídos de p41 y con el ligando 3BP1	187
3.1.2.1. Efecto de la sustitución P8L: comparación de los ligandos p40 frente a p17 y p7 frente a p0	188
3.1.2.2. Efecto de la sustitución S5P: comparación de los ligandos p40 frente a p7 y p17 frente a p0	191
3.1.2.3. La sustitución por Tyr del cuarto resto en el ligando natural 3BP1 resulta en un aumento de afinidad por Abl-SH3	192
<b>3.2. RECONOCIMIENTO DE LIGANDOS RICOS EN PROLINA POR DOMINIOS WW</b>	<b>194</b>
3.2.1. Interacción del ligando M1 con Yap65-WW y Nedd4-WW3	196
3.2.2. Interacción de dominios L víricos tipo PPxY con Nedd4-WW3 y Yap65-WW	200
<b>3.3. RECONOCIMIENTO DE LIGANDOS RICOS EN PROLINA POR DOMINIOS UEV</b>	<b>205</b>
3.3.1. Caracterización termodinámica de la interacción entre el dominio L de VIH y Tsg101-UEV	205
3.3.2. Caracterización termodinámica de las interacciones de los dominios L de ébola y leucemia humana tipo I con Tsg101-UEV	210
3.3.3. Determinación estructural del dominio UEV de Tsg101 mediante difracción de rayos X	216
3.3.3.1. Cristalización y experimentos de difracción del dominio UEV de Tsg101	216
3.3.3.2. Descripción de la estructura del dominio UEV de Tsg101	219
<b>CAPÍTULO 4: DISCUSION</b>	<b>225</b>
<b>4.1. RECONOCIMIENTO DE LIGANDOS RICOS EN PROLINA POR DOMINIOS SH3</b>	<b>227</b>
4.1.1. La interacción del dominio Abl-SH3 con ligandos ricos en prolina está caracterizada por un patrón termodinámico incongruente con la naturaleza hidrofóbica de la interacción	227
4.1.2. Los sitios de hidratación observados en el complejo Abl-SH3/p41 están conservados en otros complejos de Abl-SH3 y presentan una elevada ocupancia en disolución	231
4.1.3. Las interacciones mediadas por moléculas de agua contribuyen significativamente al elevado carácter exotérmico de la interacción entre Abl-SH3 y p41	234
4.1.4. La presencia de moléculas de agua en la interfase de unión es un fenómeno muy extendido entre los distintos dominios SH3 y puede ser en parte responsable de la especificidad de unión	240
4.1.4.1. Moléculas de agua en la región de la hélice 3 <sub>10</sub>	242
4.1.4.2. Moléculas de agua en el bolsillo de especificidad (lazos n- <i>Src</i> y RT)	248
4.1.5. La incorporación de las interacciones mediadas por moléculas de agua a estrategias de diseño racional permitiría optimizar la afinidad	

y especificidad de unión en los complejos SH3	249
4.1.6. Efectos inesperados en la energética de unión inducidos por modificaciones en la secuencia del ligando ponen de manifiesto la elevada complejidad del reconocimiento de secuencias ricas en prolina por dominios SH3	252
<b>4.2. RECONOCIMIENTO DE SECUENCIAS RICAS EN PROLINA POR OTROS DOMINIOS MODULARES</b>	<b>257</b>
4.2.1. El reconocimiento de ligandos ricos en prolina por parte de dominios WW y UEV presenta el mismo patrón que los dominios SH3, incongruente con la elevada hidrofobicidad de las interacciones	257
4.2.2. Reconocimiento de dominios L víricos tipo PPxY por dominios WW	261
4.2.3. Reconocimiento de dominios L víricos tipo PTAP por el dominio UEV de Tsg101	263
<b>CAPÍTULO 5: RESUMEN Y CONCLUSIONES</b>	<b>267</b>
<b>CAPÍTULO 6: APÉNDICES</b>	<b>275</b>
<b>CAPITULO 7: BIBLIOGRAFÍA</b>	<b>285</b>

## **ABREVIATURAS**



## **ABREVIATURAS Y SIMBOLOS**

**ADN:** ácido desoxirribonucleico

**ARN:** ácido ribonucleico

**A<sub>d</sub>:** absorbancia de luz polarizada circularmente hacia la derecha

**A<sub>i</sub>:** absorbancia de luz polarizada circularmente hacia la izquierda

**[A]:** concentración de ligando A

**[A]<sub>0</sub>:** concentración inicial de ligando A

**[B]:** concentración de ligando B

**[B]<sub>0</sub>:** concentración inicial de ligando B

**CD (o DC):** dicroísmo circular

**DSC (o CDB):** calorimetría diferencial de barrido

**EM:** espectrometría de masas

**E<sub>covalente</sub>:** potencial o campo de fuerzas debido a la parte covalente de las moléculas

**E<sub>nocovalente</sub>:** potencial o campo de fuerzas debido a la parte nocovalente de las moléculas

**E<sub>enlace</sub>:** potencial o campo de fuerzas debido a los enlaces químicos de las moléculas

**E<sub>ángulo</sub>:** potencial o campo de fuerzas debido a los ángulos de los enlaces de las moléculas

**E<sub>diedro</sub>:** potencial o campo de fuerzas debido a los diedros de los enlaces de las moléculas

**E<sub>electrostático</sub>:** potencial o campo de fuerzas debido a la contribución electrostática de las moléculas

**E<sub>vanderWaals</sub>:** potencial o campo de fuerzas debido a las fuerzas de van der Waals de las moléculas

**E<sub>ca</sub>:** potencial de corto alcance en interacciones de tipo no covalente

**E<sub>ia</sub>:** potencial de corto alcance en interacciones de tipo no covalente

**E<sub>±</sub>:** vector eléctrico de luz polarizada circularmente

**fs:** femtosegundos

**I<sub>d</sub>:** intensidad de luz polarizada circularmente hacia la derecha

**I<sub>i</sub>:** intensidad de luz polarizada circularmente hacia la izquierda

**I<sub>o</sub>:** intensidad de luz polarizada circularmente que incide sobre la muestra

**ITC (o CIT):** calorimetría isotérmica de titulación

**K:** constante de equilibrio de unión

**$K_a$** : constante de equilibrio de asociación  
 **$K_d$** : constante de equilibrio de disociación  
 **$K_D$** : constante de equilibrio de desplegamiento  
**LB**: medio de cultivo Luria Bertani  
**[L]**: concentración de ligando libre  
**[L]<sub>b</sub>**: concentración de ligando unido  
**[L]<sub>0</sub>**: concentración de ligando en la jeringa  
 **$m_p$** : cantidad de proteína en la célula calorimétrica  
**[M]<sub>T</sub>**: concentración de macromolécula total  
**[M]<sub>i</sub>**: concentración de macromolécula total en tras cada inyección "i"  
**n**: número de sitios de unión de un ligando a una proteína  
 **$n_H$** : número de protones intercambiados con el tampón  
**ns**: nanosegundos  
**PCR**: reacción en cadena de la polimerasa  
**[P]**: concentración de proteína  
**[P]<sub>0</sub>**: concentración de proteína inicial  
**ps**: picosegundos  
**RMSD**: desviación cuadrática media  
**RB**: medio de cultivo enriquecido (del inglés Rich Broth)  
**RMN**: resonancia magnética nuclear  
 **$T_a$** : temperatura de anillamiento de oligonucleótidos en la PCR  
 **$T_f$** : temperatura de fusión de oligonucleótidos en la PCR  
 **$T_m$** : temperatura de la transición a la que  $\Delta G_D = 0$   
 **$t_r$** : tiempo de residencia de una molécula de agua  
 **$\langle t_r \rangle$** : tiempo de residencia promedio de una molécula de agua  
 **$V_C$** : volumen de célula del calorímetro  
 **$V_{in}$** : volumen de inyección  
 **$V_p$** : volumen específico parcial de la proteína  
 **$V_s$** : volumen específico parcial del disolvente  
**wt**: proteína silvestre (del inglés wild type)  
**z**: carga de un ión  
**Z**: función de partición de unión  
 **$\bar{v}$** : parámetro de unión  
 **$C_P$** : capacidad calorífica molar parcial de una proteína

$C_{P,S}$  : capacidad calorífica molar parcial del disolvente  
 $C_{P,N}$  : capacidad calorífica molar parcial del estado nativo  
 $C_{P,D}$  : capacidad calorífica molar parcial del estado desnaturalizado  
 $\Delta ASA$ : cambio de área accesible al disolvente  
 $\Delta C_p$ : cambio de capacidad calorífica de desnaturalización a presión constante  
 $\Delta C_{p, ap}$ : cambio de capacidad calorífica aparente de unión  
 $\Delta C_{p, int}$ : cambio de capacidad calorífica intrínseco de unión  
 $\Delta C_{p, ion}$ : cambio de capacidad calorífica característico del tampón  
 $\Delta S_{ap}$ : entropía aparente de unión  
 $\Delta S_D$ : entropía de desplegamiento de una proteína  
 $\Delta m_S$ : cantidad de agua desplazada por una proteína al desplegarse  
 $\Delta G_{ap}$ : energía libre aparente de Gibbs asociada a un proceso de unión  
 $\Delta G_D$ : energía libre de Gibbs de desplegamiento  
 $\Delta H_{ap}$ : entalpía aparente de unión  
 $\Delta H_{conf}$ : entalpía de unión asociada a cambios conformacionales en proteína y/o ligando  
 $\Delta H_D$ : entalpía de desplegamiento de una proteína  
 $\Delta H_{ap}$ : entalpía aparente de unión  
 $\Delta H_{ion}$ : entalpía de ionización característica del tampón  
 $\Delta H_{int}$ : entalpía intrínseca de unión considerando que ligando y proteína se unen sin cambio conformacional asociado  
 $\Delta H_m$ : cambio de entalpía a la temperatura  $T_m$   
 $\Delta H_{vH}$ : entalpía de van't Hoff  
 $\langle H \rangle$  : entalpía molar parcial del sistema  
 $\Delta S_{ap}$ : entropía aparente de unión  
 $\Delta S_D$ : entropía de desplegamiento de una proteína  
 $\Delta S_m$ : cambio de entropía a la temperatura  $T_m$   
 $\epsilon'$  : mínimo de potencial  
 $\epsilon_0$  : constante dieléctrica del medio  
 $\epsilon_D$  : coeficiente de extinción molar para luz polarizada hacia la derecha  
 $\epsilon_L$  : coeficiente de extinción molar para luz polarizada hacia la izquierda  
 $\lambda$ : longitud de onda de luz polarizada  
 $\nu$ : frecuencia de la luz polarizada  
 $[\theta]$ : elipticidad

$[\theta]_{\text{MMR}}$ : elipticidad molar media por aminoácido

$\Phi$ : potencial genérico de interacción

$\Phi_{\text{ca}}$ : potencial genérico para interacciones de corto alcance

$\Phi_{\text{elect}}$ : potencial electrostático

$\Phi_{\text{la}}$ : transformada de Fourier para el potencial de largo alcance

$\rho'(\mathbf{k})$ : densidad de carga

$\xi$ : constante de fricción para las partículas

# **INTRODUCCIÓN**



## **1. INTRODUCCIÓN**

El origen de la palabra proteína parece provenir del griego “prota”, que significa lo primero, o del dios “proteo” por la cantidad de formas diferentes que puede tomar. Las proteínas, o mejor dicho los aminoácidos que las componen, son uno de los elementos básicos e imprescindibles que conforman las células, junto con el ADN, ARN, glúcidos y lípidos. Casi la mayoría de procesos biológicos celulares dependen de las proteínas, que tienen funciones muy diversas: son elementos estructurales que ayudan a mantener la integridad de las células, conectan diferentes elementos dentro de ellas, son responsables de la regulación y transporte de nutrientes, generan, transforman y transfieren impulsos nerviosos, amén de otras muchas más funciones (Pawson 1995).

Dentro del contexto funcional, las proteínas establecen una compleja red de interacciones que es crucial para el correcto funcionamiento celular. Cualquier distorsión de esta red de interacciones puede traducirse en serias alteraciones de las funciones celulares que originan el desarrollo de muchas enfermedades humanas, algunas del calibre del cáncer o VIH. Por este motivo, una profunda y detallada comprensión de estas interacciones nos permitirá predecir y modular la respuesta fisiológica, o incluso patológica, de la célula a un estímulo específico y facilitará el desarrollo de nuevas estrategias de diseño molecular que puedan ser utilizadas en el tratamiento de algunas de estas enfermedades.

Con la revolución en el campo de la genómica y proteómica experimentada durante la última década se ha hecho aún más patente la necesidad de profundizar en el estudio del reconocimiento molecular por parte de numerosas proteínas. La resolución del genoma humano, junto con la de otros organismos modelo, ha proporcionado miles de nuevos genes y proteínas que deberán caracterizarse estructural y funcionalmente. Sus interacciones con otras moléculas en la célula deberán ser identificadas y analizadas con el fin de avanzar hacia la comprensión y el control de los mecanismos de regulación celular. A su vez, muchas de las proteínas identificadas en los genomas son dianas potenciales para el desarrollo de nuevos agentes terapéuticos contra una gran variedad de enfermedades que habrá que validar y caracterizar.

## **Introducción**

Las interacciones proteína-proteína se establecen frecuentemente a través de pequeños dominios modulares que se encuentran muy conservados y están presentes en distintas proteínas, los cuales reconocen secuencias específicas en sus dianas (Pawson 1995; Pawson 2004; Pawson *et al.* 2002). Dentro de los diferentes tipos de dominios modulares, uno de los más ampliamente representados en el genoma humano son los que reconocen secuencias ricas en prolina (MRP) (Macias *et al.* 2002; Mayer 2001; Zarrinpar *et al.* 2003a; Zarrinpar & Lim 2000). Actualmente se conocen hasta cinco familias de este tipo de dominios, que son los dominios SH3, WW, EVH1, GYF y UEV (Ball *et al.* 2005). Bastantes de los miembros de estas familias están implicados en el desarrollo de enfermedades como la leucemia humana, alzheimer, ébola y VIH, entre otras, por lo que el diseño de ligandos que interfieran en el establecimiento de las interacciones con sus dianas se ha consolidado en los últimos años como una estrategia viable para el desarrollo de nuevos fármacos específicos contra estas enfermedades (Kardinal *et al.* 2000; Lee *et al.* 2002; Oneyama *et al.* 2002).

En todos los casos, estos dominios modulares reconocen secuencias consenso específicas, caracterizadas por contener al menos un resto prolina, que favorecería la formación de la hélice de poliprolina II (PPII) necesaria para la unión al dominio (Macias *et al.* 2002; Sudol 1998). A pesar de las grandes diferencias estructurales entre las distintas familias y la diversidad en las secuencias que reconocen, todas las familias comparten un mecanismo común para el reconocimiento de sus ligandos muy similar al establecido en la última década para los dominios SH3, que se basa en la inserción de las prolinas principales del ligando en bolsillos hidrofóbicos del dominio formados por restos aromáticos (Ball *et al.* 2005; Li 2005). Aunque los dominios SH3 son los más ampliamente estudiados, los determinantes de la afinidad y especificidad de unión de las interacciones que establecen siguen sin estar lo suficientemente claros hoy día, debido en gran medida a la escasa información disponible sobre la naturaleza y magnitud de las fuerzas que dirigen estas interacciones. Así, mientras que para los dominios SH3 existen algunos trabajos que muestran un patrón termodinámico incongruente con la naturaleza hidrofóbica del sitio de unión y de los ligandos, para otras familias de MRP no existe ningún trabajo de este tipo. El objetivo que se pretende con este trabajo es llevar a cabo una caracterización detallada de las interacciones establecidas por los

dominios SH3, así como de otras dos familias de dominios modulares, WW y UEV, para las que apenas existe información de carácter biofísico. Para ello, hemos abordado el estudio de estas interacciones haciendo uso de diversas técnicas experimentales, desde la clonación, expresión y purificación de ejemplos representativos de cada una de estas familias de dominios, hasta su estudio con técnicas experimentales espectroscópicas, calorimétricas, y en algunos casos hasta su determinación estructural mediante difracción de rayos X, así como con técnicas computacionales, mediante simulaciones de dinámica molecular de algunos de sus complejos con ligandos.

El trabajo realizado en esta Memoria se ha llevado a cabo en las siguientes etapas:

- I. Caracterización termodinámica y estructural de las interacciones establecidas por el dominio Abl-SH3 con ligandos ricos en prolina.
- II. Evaluación del papel de las moléculas de agua en las interacciones que establece Abl-SH3 con el ligando rico en prolina p41 mediante técnicas calorimétricas, cristalográficas y computacionales.
- III. Caracterización termodinámica de las interacciones establecidas por dominios WW (Yap65-WW y Nedd4-WW3) con dominios L tipo PPxY (secuencias ricas en Pro presentes en las poliproteínas gag de algunos virus como el ébola y de la leucemia humana).
- IV. Caracterización termodinámica y estructural de las interacciones establecidas por el dominio UEV de Tsg101 con dominios L tipo PTAP.

El objetivo final que pretendemos sería, por consiguiente, obtener información energética, estructural y dinámica que sea útil para el diseño de nuevos ligandos con mejores propiedades (afinidad y especificidad) que puedan ser utilizados con fines terapéuticos en el tratamiento de algunas de estas enfermedades.



## **ANTECEDENTES**



## **1.1. ANTECEDENTES**

### **1.1.2. Dominios modulares de reconocimiento de secuencias ricas en prolina**

Las interacciones proteína-proteína presentan un elevado interés ya que juegan un papel central en la regulación de muchos procesos celulares (rutas metabólicas, progresión del ciclo celular, síntesis de proteínas, replicación del ADN, etc.). Cualquier distorsión de esta compleja red de interacciones entre proteínas puede traducirse en serias alteraciones de las funciones celulares que constituyen el origen de algunas enfermedades humanas como VIH, cáncer, leucemia, etc. Con frecuencia, las interacciones proteína-proteína están mediadas por pequeños dominios modulares, que reconocen secuencias específicas en sus dianas y están altamente conservados y presentes en las diferentes familias de proteínas en número variable y en distintas combinaciones, (Pawson 1995; Pawson *et al.* 2002; Pawson & Scott 1997). Estos dominios han proporcionado a la naturaleza un mecanismo muy eficiente para el control de la actividad celular, ya que los distintos módulos contribuyen a la localización de enzimas y sustratos para crear cascadas de señalización específicas, así como al reconocimiento de las diferentes modificaciones post-transcripcionales o incluso a la intercomunicación entre distintas rutas de señalización.

Los dominios modulares son generalmente de pequeño tamaño (entre 30 y 150 restos) y se encuentran plegados en estructuras compactas y estables caracterizadas por uno o más sitios de unión. Las familias de dominios son estructuralmente diferentes y reconocen pequeñas secuencias de entre 3 y 6 aminoácidos, específicas y muy conservadas, en sus proteínas diana. Así, los dominios SH3 se unen a secuencias ricas en prolina, los dominios SH2 reconocen restos de tirosina fosforilada y los dominios PDZ interaccionan con secuencias correspondientes a extremos carboxilo terminal (Ball *et al.* 2005; Jelen *et al.* 2003; Kuriyan & Cowburn 1997). Estas secuencias consenso están flanqueadas por restos adicionales que interaccionan con elementos variables en el sitio de unión del dominio y que parecen ser responsables de la especificidad de unión dentro de cada familia de dominios. De este modo, las interacciones mediadas por dominios

## **Antecedentes**

modulares presentan dos características especialmente interesantes: por un lado reconocen epítomos que constituyen secuencias continuas en las proteínas que los contienen, lo que simplifica notablemente la identificación "*in silico*" de dianas fisiológicamente relevantes para cada dominio a nivel de los distintos proteomas (Brannetti & Helmer-Citterich 2003; Neduva *et al.* 2005; O'Leary *et al.* 2004; Reiss & Schwikowski 2004); por otro, estos dominios presentan sitios de unión cóncavos con superficies de interacción pequeñas (en torno a 1000 Å<sup>2</sup>), más similares a las observadas en los complejos entre pequeños ligandos y enzimas que las descritas para la mayoría de las interacciones proteína-proteína, lo que facilita el diseño de moléculas de pequeño tamaño que bloqueen estas interacciones y hace de estos dominios atractivas dianas para el desarrollo de agentes terapéuticos (Kardinal *et al.* 2000; Lawrence 2005; Lee *et al.* 2002).

De entre las distintas familias de dominios modulares, los módulos de reconocimiento de secuencias ricas en prolina (MRP) son los más ampliamente representados en metazoos. Estos dominios parecen haber coevolucionado con las secuencias ricas en prolina y éstas son, a su vez, los motivos más ampliamente distribuidos en los distintos proteomas, desde procariotas a eucariotas (Rubin *et al.* 2000), debido en gran medida a las propiedades específicas que el aminoácido prolina les confiere (Creamer 1998; Creamer & Campbell 2002; Ferreon & Hilser 2003a; Kay *et al.* 2000; Stapley & Creamer 1999; Vijayakumar *et al.* 1999; Zarrinpar *et al.* 2003a). El carácter cíclico de su cadena lateral restringe considerablemente el espacio conformacional accesible a este aminoácido, lo que se traduce en una baja propensión a formar parte de estructuras secundarias como hélices alfa o láminas beta y hace que las secuencias ricas en prolina se encuentren con frecuencia en la superficie molecular, requisito imprescindible para cualquier secuencia implicada en interacciones proteína-proteína. La estructura adoptada espontáneamente con mayor frecuencia cuando se presentan dos o más restos de prolina consecutivos es la denominada hélice poliprolina II (PPII) que consiste en una hélice a izquierdas con tres aminoácidos por vuelta. Esta es la misma conformación adoptada por las secuencias ricas en prolina en los complejos, lo que reduce notablemente el coste entrópico de la interacción. La hélice PPII es además muy estable y resiste la intercalación de otros aminoácidos permitiendo cierta variabilidad en la secuencia de los epítomos. En la conformación de PPII los carbonilos de las prolinas se proyectan

hacia fuera de la hélice desde el eje y, por lo tanto, están bien situados para establecer interacciones con otras moléculas. En concreto, la ausencia del protón amida proporciona unas características electrónicas particulares para este carbonilo favoreciendo su participación como grupo aceptor en el establecimiento de puentes de hidrógeno.

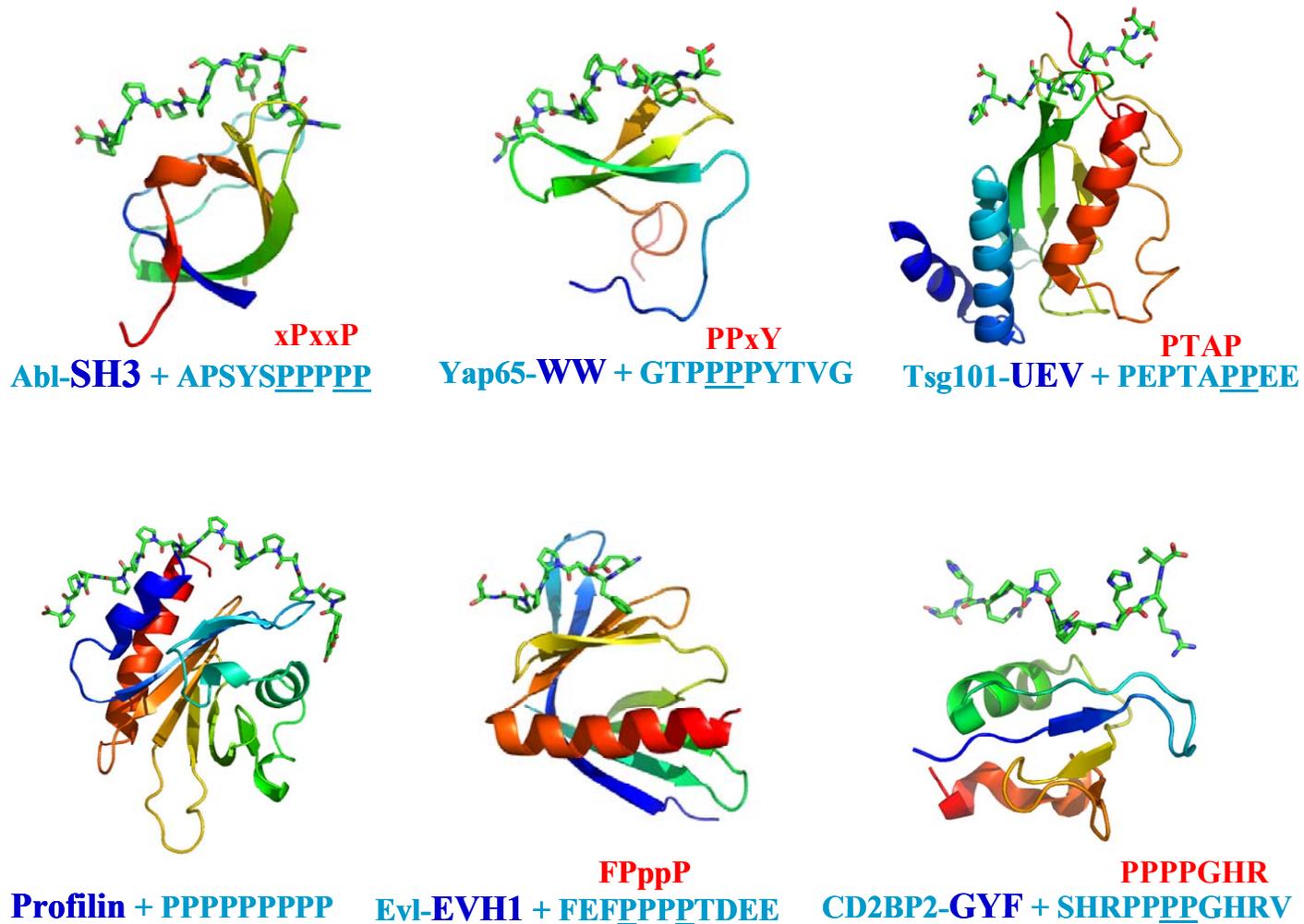
Actualmente, se conocen cinco familias de MRP: los dominios SH3 (*Src Homology 3*), los dominios WW (que toman su nombre de dos restos Trp muy conservados), los dominios EVH1 (*Enabled Vasodilator-stimulated-protein Homology*), los dominios GYF (denominados así porque contienen la tríada característica Gly-Tyr-Phe) y los dominios UEV (*Ubiquitin E2 Variant*) (Ball *et al.* 2005; Li 2005). Los MRP presentan tamaños que varían entre los 30-35 aminoácidos de los dominios WW hasta los 150 restos de los UEV, forman parte de proteínas muy diversas y frecuentemente actúan como sitios de anclaje para el reclutamiento de sustratos y la formación de complejos supramoleculares que conducen a la modificación enzimática de algunos de sus componentes. En ocasiones, como es el caso de algunos dominios SH3, estos módulos son también piezas esenciales en la regulación de la actividad enzimática de las proteínas que los contienen mediante el establecimiento de interacciones intramoleculares con otros elementos de la molécula (Arold *et al.* 1998; Barila & Superti-Furga 1998; Brabek *et al.* 2002; Brasher *et al.* 2001). Los dominios SH3 y WW son, con diferencia, los más abundantes en vertebrados con un número estimado de 409 y 125 sólo en el proteoma humano, así como los más extensamente estudiados {Zarrinpar, 2000 #74; Mayer, 2001 #59; Macías, 2002 #73; Zarrinpar, 2003 #100}. Las demás familias de MRP son menos abundantes, lo que quizá refleja una función más especializada (Ball *et al.* 2002; Freund *et al.* 1999; Holt & Koffer 2001).

Cada una de estas familias de MRP reconoce secuencias canónicas diferentes que se caracterizan por contener al menos un resto de prolina {Sudol, 1998 #545; Macías, 2002 #73}. Estas secuencias favorecen la formación de la conformación PPII y juegan un papel muy importante para la interacción, aunque la selectividad y especificidad de unión parecen radicar en los aminoácidos que las flanquean. Algunas familias como los dominios GYF interactúan sólo con patrones de secuencia muy definidos, mientras que otras como los dominios SH3, WW o

## **Antecedentes**

EVH1 son mucho menos selectivas y pueden subdividirse en diferentes clases de acuerdo a su especificidad (Cesareni *et al.* 2002; Guo *et al.* 2003; Schleinkofer *et al.* 2004). Aunque se ha avanzado notablemente en el establecimiento de las secuencias consenso para cada familia de dominios, hay algunos aspectos que todavía no están lo suficientemente claros. Así, existen familias de dominios como los SH3 y WW que interaccionan con secuencias consenso muy similares (Cesareni *et al.* 2002; Otte *et al.* 2003; Schleinkofer *et al.* 2004). Paradójicamente, aunque se postula que las diferencias entre los miembros de una misma familia son suficientes para establecer redes de interacción específicas (Zarrinpar *et al.* 2003b), se ha detectado un elevado grado de promiscuidad en las interacciones que establecen muchos de estos dominios.

A pesar de las grandes diferencias estructurales existentes entre las cinco familias de MRP y la diversidad en las secuencias que reconocen (ilustradas en la Figura 1.1), todas las familias interaccionan con su dianas con afinidades moderadas ( $K_d$  entre 1 y 500  $\mu\text{M}$ ) y comparten un mecanismo de unión que se asemeja al establecido para los dominios SH3 durante la última década (Ball *et al.* 2005; Li 2005), lo que podría ser el origen de la promiscuidad y reactividad cruzada existente entre las distintas clases de dominios (Hu *et al.* 2004; Landgraf *et al.* 2004; Tong *et al.* 2002). Las superficies de interacción de todos los dominios están enriquecidas en restos aromáticos, muy conservados dentro de los miembros de una misma familia, cuya disposición tridimensional es crítica para el reconocimiento del ligando que, como ya se ha mencionado, se une en todos los casos adoptando una conformación de hélice PPII. En general, los distintos MRP contienen al menos un bolsillo hidrofóbico formado por restos aromáticos muy conservados dentro de cada familia en el que se acomoda un dipéptido del tipo xP, donde x es frecuentemente un resto hidrofóbico (Figura 1.1). Contiguo a estos bolsillos suele encontrarse otro bolsillo más variable que es el que interacciona con los restos del ligando adyacentes a la secuencia consenso central, que serían los responsables de dictar la especificidad en las interacciones entre los distintos miembros de una misma familia (Ball *et al.* 2005; Zarrinpar *et al.* 2003a).



**Figura 1.1.** Ejemplos representativos de las familias de dominios modulares que reconocen secuencias poliprolina para SH3, WW, GYF, EVH1 y UEV, y la familia de proteínas de profilina, formando complejo con ligandos tipo PPII. En todos los casos, los ligandos (representados en código de colores universal, en la parte superior de cada complejo) se unen a cada dominio a través de un sitio de unión hidrofóbico formado por restos aromáticos muy conservados entre los miembros de cada familia de dominios. La secuencia consenso descrita originalmente para las diferentes familias aparece en color rojo sobre la secuencia completa de cada ligando, donde las prolinas que se insertan en los bolsillos hidrofóbicos aparecen subrayadas.

## **Antecedentes**

Como consecuencia de la importancia de su papel en la regulación de numerosas funciones celulares los MRP están asociados al desarrollo de importantes patologías humanas como cáncer, distrofia muscular y procesos inflamatorios entre otros (Brabek *et al.* 2002; Dalgarno *et al.* 1997; Skorski *et al.* 1998; Sudol *et al.* 2001). Además muchos virus han evolucionado para explotar el reconocimiento de secuencias ricas en prolina por parte de algunos MRP para su proliferación. En este contexto, el diseño de moléculas que puedan interferir en el establecimiento de las interacciones entre estos dominios y sus dianas se ha consolidado como una estrategia viable para el desarrollo de nuevos fármacos específicos para el tratamiento de numerosas enfermedades (Kardinal *et al.* 2000; Lee *et al.* 2002; Oneyama *et al.* 2002; Stauffer *et al.* 1997; Vidal *et al.* 2001). Sin embargo, a pesar de la abundante información estructural y funcional acumulada sobre el reconocimiento de secuencias ricas en prolina por los distintos MRP y del interés que la inhibición de las interacciones entre MPR y sus dianas ha cobrado en los últimos años, el diseño de ligandos de alta afinidad y especificidad se ha revelado como un reto especialmente complejo, debido en parte a la baja afinidad de las interacciones naturales y también en gran medida a que no se dispone de un conocimiento profundo y detallado de los determinantes de la afinidad y especificidad de unión en estos sistemas. Muchos han sido los esfuerzos dirigidos a la identificación y diseño de ligandos de alta afinidad para dominios SH3 y otros MRP, (Dalgarno *et al.* 1997; Garbay *et al.* 2000) que incluyen estrategias diversas como el uso de bibliotecas de compuestos sintéticos, la sustitución sistemática de las distintas posiciones del ligando por alanina (Feng *et al.* 1994; Ren *et al.* 1993; Yu *et al.* 1994), técnicas de expresión en fagos (Panni *et al.* 2002; Tong *et al.* 2002), la utilización de elementos no peptídicos en los ligandos (Combs *et al.* 1996; Feng *et al.* 1996; Panni *et al.* 2002) o incluso la incorporación de D-aminoácidos (Schumacher *et al.* 1996). En todos los casos los avances han sido muy limitados, aunque el uso de métodos de química combinatorial combinados con bibliotecas de peptoides donde las prolinas clave han sido modificadas ha proporcionado algunos ligandos de muy alta afinidad para dominios SH3 (Aghazadeh & Rosen 1999; Nguyen *et al.* 2000; Nguyen *et al.* 1998). La aplicación de técnicas de diseño racional ha proporcionado también algunos resultados satisfactorios, como los obtenidos por el grupo del Dr. Luís Serrano (EMBL, Heidelberg), donde han conseguido varias secuencias naturales de alta afinidad y especificidad para el

dominio SH3 de Abl (Pisabarro *et al.* 1994; Pisabarro & Serrano 1996). Sin embargo, el éxito de estas metodologías de diseño racional ha sido modesto en general.

Tradicionalmente, los procesos de identificación y optimización de ligandos se han dirigido únicamente a incrementar la afinidad de unión, bien mediante el cribado de extensas bibliotecas de compuestos, o bien mediante el uso de algoritmos computacionales sustentados exclusivamente en información estructural. Sin embargo, considerar la afinidad de unión, determinada por la energía de Gibbs como único criterio, constituye una visión muy limitada del proceso de unión, ya que no proporciona información alguna sobre la naturaleza y magnitud de las fuerzas que lo dirigen (es decir, de las contribuciones entálpicas y entrópicas) y, en muchos casos, impide un avance significativo en el desarrollo de nuevos ligandos o en la comprensión de la especificidad de unión. Trabajos recientes en varios grupos de investigación, incluyendo el grupo del Dr. Freire y el nuestro propio, han puesto de manifiesto la importancia de incorporar estas consideraciones termodinámicas al proceso de diseño (Luque & Freire 1998; Luque & Freire 2002; Luque *et al.* 2002; Luque *et al.* 1998; Velazquez Campoy & Freire 2005; Velazquez-Campoy & Freire 2006; Velazquez-Campoy *et al.* 2001; Velazquez-Campoy *et al.* 2004). Es en este contexto en el que se enmarca el trabajo que se presenta en esta Memoria que tiene como objetivo general proporcionar una descripción termodinámico-estructural profunda y detallada de los procesos de reconocimiento de secuencias ricas en prolina por distintos MRP, con el fin de avanzar en la comprensión de los determinantes de la afinidad y especificidad de unión y en el desarrollo de estrategias racionales de diseño de inhibidores específicos para estas interacciones.

### **1.1.3. Dominios SH3**

Como hemos comentado, los dominios SH3 son los MRP más ampliamente representados en los distintos proteomas, con más de 400 dominios diferentes identificados hasta el momento (Macias *et al.* 2002; Mayer 2001; Zarrinpar *et al.* 2003a). Estos módulos forman parte de proteínas muy diversas presentes en eucariotas, tales como las tirosina quinasas Abl, Yes, Src, Fyn, Lyn, Hck, Lck, Nck, proteínas de unión a actina, Bem1, cdc25, cortactina, la subunidad del canal de

## **Antecedentes**

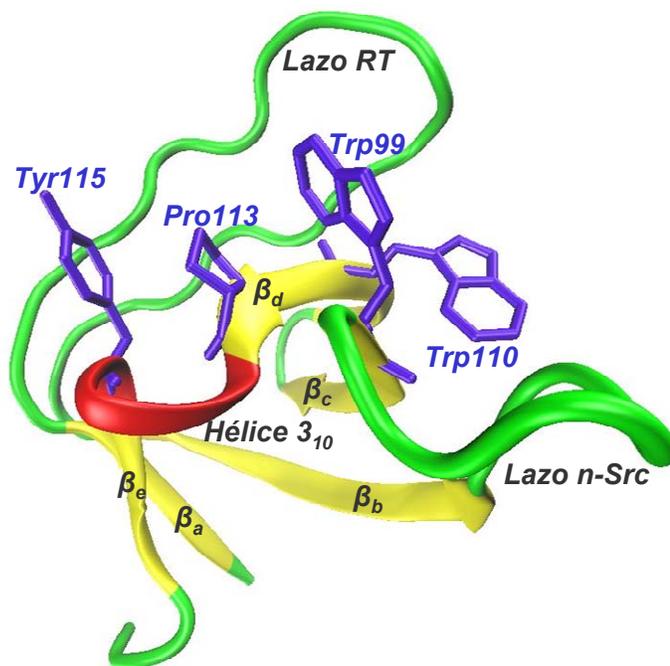
calcio beta1B2, Grb2, miosina, la subunidad reguladora de PI3K, el activador de GTPasa ras, espectrina, la proteína ZO-1, etc. (Kay *et al.* 2000; Mayer 2001; Zarrinpar *et al.* 2003a). Los dominios SH3 actúan generalmente como sitios de anclaje para el reclutamiento de sustratos y la formación de complejos supramoleculares que conducen a la modificación enzimática de algunos de sus componentes. En ocasiones desempeñan también un papel esencial para la regulación de la actividad enzimática de las proteínas que los contienen mediante el establecimiento de interacciones intramoleculares con otros elementos de la molécula (Arold *et al.* 2001; Barila & Superti-Furga 1998; Brabek *et al.* 2002; Brasher *et al.* 2001); este es el caso de las quinasas de tirosina de la familia Src, algunos de cuyos miembros, especialmente Src y Yes, están claramente relacionados con el desarrollo de procesos cancerígenos. Además del cáncer, los dominios SH3 están implicados en otras importantes patologías como el SIDA (Lyn, Hck, Lck), leucemia (Abl), osteoporosis (Src), o procesos inflamatorios, alérgicos y asmáticos (Dalgarno *et al.* 1997; Skorski *et al.* 1998).

Los primeros estudios estructurales con dominios SH3 fueron realizados por los grupos de los Doctores Saraste y Schreiber, que obtuvieron las primeras estructuras resueltas por rayos X y RMN respectivamente y proporcionaron las primeras pistas para identificar el sitio de unión para los ligandos ricos en prolina (Musacchio *et al.* 1992; Yu *et al.* 1992). Al mismo tiempo estos estudios estructurales proporcionaron una nueva visión de las interacciones proteína-proteína y condujeron a la idea de que éstas podían estar mediadas por estos dominios de naturaleza modular (Yu *et al.* 1994). Su importancia biológica, además de su pequeño tamaño (en torno a 60 restos) y a la relativa simplicidad de sus protocolos de expresión y purificación los han convertido en candidatos ideales para su estudio por parte de numerosos grupos de biología estructural. Gracias a esto, el número de estructuras de dominios SH3 ha crecido exponencialmente durante la última década (Arold *et al.* 1998; Bauer *et al.* 2005; Blanco *et al.* 1997; Borchert *et al.* 1994; Martin-Garcia *et al.* 2007; Musacchio *et al.* 1992; Noble *et al.* 1993; Ortega Roldan *et al.* 2007; Schweimer *et al.* 2002), también de sus complejos con ligandos peptídicos (Camara-Artigas *et al.* 2007; Candel *et al.* 2007; Feng *et al.* 1994; Feng *et al.* 1995; Lim & Richards 1994; Morton *et al.* 1996; Musacchio *et al.* 1994; Pisabarro *et al.* 1998; Renzoni *et al.* 1996; Wittekind *et al.* 1994; Wu *et al.* 1995; Yu *et al.* 1994), así como

de las proteínas que los contienen (Burchat *et al.* 2006; Cowan-Jacob *et al.* 2005; Nagar *et al.* 2006; Nagar *et al.* 2003; Sicheri *et al.* 1997; Xu *et al.* 1997), pudiendo citar ejemplos hasta las casi 400 estructuras depositadas en el PDB. El plegamiento característico de los dominios SH3 se muestra en la Figura 1.2 y consiste en una estructura de tipo barril  $\beta$  constituido por cinco hebras antiparalelas, denominadas  $\beta_a$ ,  $\beta_b$ ,  $\beta_c$ ,  $\beta_d$ , y  $\beta_e$ , dispuestas en dos láminas  $\beta$  ortogonales entre sí, una de menor tamaño constituida por las hebras  $\beta_b$ - $\beta_a$ - $\beta_e$  y otra mayor integrada por las hebras  $\beta_b$ - $\beta_c$ - $\beta_d$ . Las cinco hebras están conectadas por tres giros variables que históricamente se han denominado RT, n-Src y distal, de modo que el lazo RT conecta las dos primeras hebras  $\beta_a$  y  $\beta_b$ , el lazo n-Src las hebras  $\beta_b$  y  $\beta_c$  y el giro distal conecta las hebras  $\beta_c$  y  $\beta_d$ . Finalmente la conexión entre las hebras  $\beta_d$  y  $\beta_e$  se realiza mediante una vuelta de hélice  $3_{10}$ .

El equilibrio conformacional y la estabilidad de los dominios SH3 han sido caracterizados de forma extensiva por varios grupos de investigación, incluido el nuestro, que mediante el uso de técnicas espectroscópicas y calorimétricas han establecido que el equilibrio de plegamiento de estos dominios puede describirse adecuadamente según el modelo de dos estados (Casares *et al.* 2007b; Casares *et al.* 2004; Filimonov *et al.* 1999; Martinez *et al.* 1998; Martinez *et al.* 1999; Sadqi *et al.* 1999). De igual forma, estudios cinéticos llevados a cabo con dominios SH3 aplicando la técnica de flujo detenido también han establecido la ausencia de intermedios cinéticos significativamente poblados durante los procesos de plegamiento y desplegamiento. Según estos estudios, el núcleo de plegamiento se encuentra localizado en la horquilla beta que forman las hebras  $\beta_c$  y  $\beta_d$  en torno al giro distal y en la hélice  $3_{10}$  (Viguera *et al.* 1996), de modo que la formación de este giro es el factor limitante en la velocidad de plegamiento de estos dominios, la cual es menor que la observada para otras proteínas de estructura beta del mismo tamaño (Perl *et al.* 1998). En este contexto, los estudios cinéticos realizados con permutantes circulares del dominio Spc-SH3 (Martinez *et al.* 1999; Viguera *et al.* 1995) han demostrado la relevancia que tiene la topología de la cadena polipeptídica durante el plegamiento, destacando la importancia que desempeñan en este proceso las horquillas beta presentes en la estructura de dicho dominio.

## Antecedentes



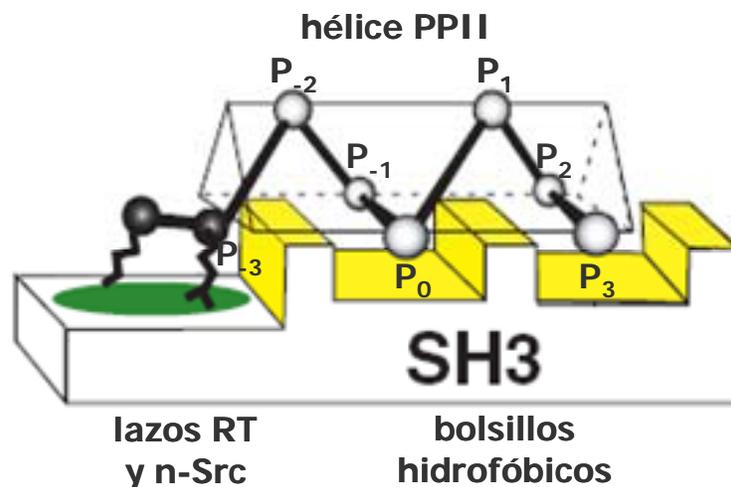
**Figura 1.2.** Estructura del dominio SH3 de Abl. Las cinco láminas beta se representan como  $\beta_a$ ,  $\beta_b$ ,  $\beta_c$ ,  $\beta_d$  y  $\beta_e$  (amarillo) y la hélice  $3_{10}$  que conecta las hebras  $\beta_d$  y  $\beta_e$  aparece en rojo. El lazo n-Src que conecta las hebras  $\beta_b$  y  $\beta_c$ , el lazo RT que conecta las hebras  $\beta_a$  y  $\beta_b$ , también se indican (verde). Los restos aromáticos más conservados que son importantes para la unión a los ligandos se representan en violeta. También se indica el resto aromático Trp110 que, en el caso de la proteína Abl-SH3 tomada como ejemplo, forma parte del bolsillo hidrofóbico para la unión a los ligandos. Código Pdb: 1abq.

A pesar de la aparente simplicidad macroscópica de los procesos de plegamiento/desplegamiento de estos pequeños dominios, estudios recientes de intercambio hidrógeno-deuterio (IHD) seguido por RMN realizados con el dominio Spc-SH3 indican que en condiciones nativas el dominio se encuentra en un equilibrio conformacional de naturaleza estadística que implica conformaciones muy variadas, que abarcan desde el estado plegado hasta estados altamente desordenados, pasando por conformaciones con desorden local o parcial. Los resultados más recientes sugieren que el dominio SH3 presenta una elevada, aunque no completa, cooperatividad estructural y, como consecuencia, las fluctuaciones conformacionales de los distintos elementos estructurales están parcialmente acopladas entre sí. Esta cooperatividad estructural implica que cualquier cambio energético o interacción producidos a escala local se puede transmitir de forma eficiente al resto de la estructura (Casares *et al.* 2007b; Casares *et al.* 2003; Sadqi *et al.* 1999; Sadqi *et al.* 2002).

Las interacciones entre dominios SH3 y sus ligandos naturales se caracterizan por afinidades moderadas con constantes de disociación generalmente en el intervalo entre 1 y 200  $\mu\text{M}$ , lo que refleja la necesidad de establecer interacciones dinámicas de carácter transitorio en el contexto de las redes celulares de transducción de señales. Los ligandos de dominios SH3 contienen generalmente una secuencia canónica del tipo xPxxP, donde x es frecuentemente un aminoácido hidrofóbico, y se unen adoptando una conformación PPII en el complejo. Como se ilustra en la Figura 1.2, el sitio de unión consiste en una hendidura poco profunda que contiene dos bolsillos hidrofóbicos constituidos por restos aromáticos muy conservados entre los distintos dominios SH3 (Trp99 situado en la hebra  $\beta_c$ , Pro113 en la hebra  $\beta_d$  y Tyr115 de la secuencia SNY en la hélice  $3_{10}$ ). Cada uno de estos bolsillos hidrofóbicos acomoda uno de los dipéptidos de tipo xP de la secuencia consenso, como se muestra en la Figura 1.3, de modo que cada resto prolina de la secuencia canónica se intercala entre dos o más restos aromáticos. Los restos adyacentes a la secuencia central xPxxP interactúan con un tercer bolsillo responsable de la especificidad en la unión delimitado por los lazos RT y n-Src, más variables en secuencia entre los miembros de la familia de dominios SH3. (Ball *et al.* 2005; Zarrinpar *et al.* 2003a). Este modo de unión se caracteriza por la escasez de interacciones polares en los complejos, ya que generalmente éstas se limitan a dos puentes de hidrógeno muy conservados, establecidos entre el Trp99 en la hebra  $\beta_c$  y la Tyr115 en la hélice  $3_{10}$  y los carbonilos de algunos restos prolina de los ligandos, y a un puente salino presente en algunos tipos de complejos establecido entre un resto Aspártico (Asp79) conservado en el dominio SH3 y un resto Arginina en el ligando (Figura 1.4).

Debido a la pseudosimetría de la conformación PPII los ligandos ricos en prolina pueden unirse al dominio con dos orientaciones diferentes: a) N a C Terminal en ligandos de tipo I caracterizados por una secuencia consenso R/Kx#PxxP o b) C a N Terminal en ligandos de tipo II con secuencia consenso xPx#PxR/K, donde # suele ser con frecuencia un resto hidrofóbico y x puede ser cualquier aminoácido (Feng *et al.* 1994; Lim *et al.* 1994). Así, la orientación del ligando se determina por las interacciones que establecen los restos que flanquean a la secuencia consenso PxxP, en particular por la posición del aminoácido básico implicado en el puente salino situado dos posiciones anterior al motivo xPxxP en los ligandos de tipo I y dos

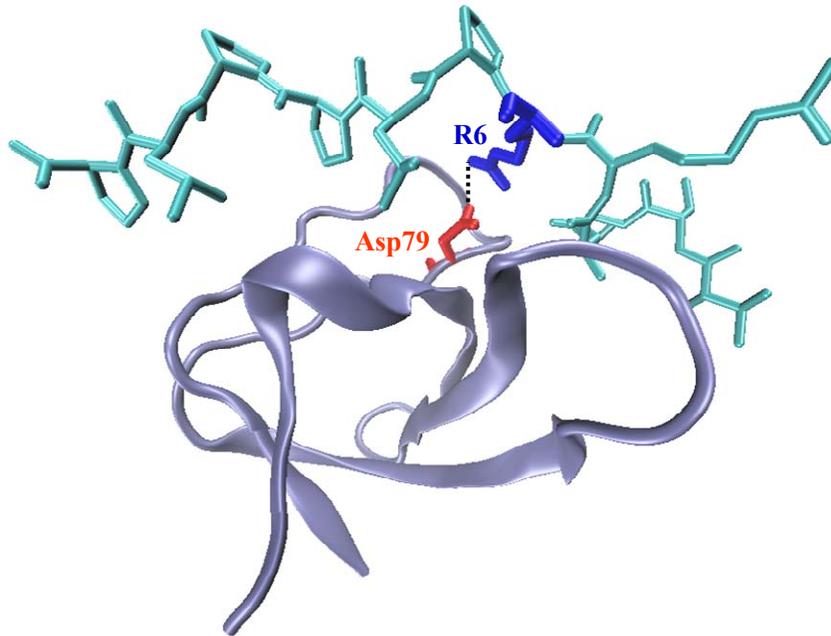
## Antecedentes



**Figura 1.3.** Representación esquemática del modo de unión mostrado por dominios SH3. En la figura se representa el sitio de unión 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<sub>0</sub>x<sub>1</sub>x<sub>2</sub>P<sub>3</sub>.

restos posterior a este motivo en los ligandos de tipo II (Feng *et al.* 1994; Lim *et al.* 1994). Estudios recientes sugieren además que la orientación del ligando puede también venir dictada por la conformación de la cadena lateral de un resto triptófano altamente conservado de la hebra  $\beta_c$  del dominio SH3 (Fernandez-Ballester *et al.* 2004).

Los dominios SH3 se pueden clasificar en tres tipos diferentes de acuerdo a su especificidad de unión: 1) dominios SH3 tipo Src-SH3 que se caracterizan por tener un puente salino entre un resto del ligando cargado positivamente (generalmente arginina) con un resto cargado negativamente en la superficie del dominio (Asp en el caso de Src). En estos casos se observan con frecuencia complejos tipo I y II, como es el caso del dominio SH3 de Src que puede unir tanto el péptido RPPPLP como el reverso PPVPPR (Kang *et al.* 2000; Sparks *et al.* 1996); 2) dominios tipo Abl-SH3 que interaccionan con el motivo canónico xPxxP pero que no presentan el puente salino (Musacchio *et al.* 1994; Pisabarro *et al.* 1998); y 3) dominios descubiertos recientemente que interaccionan con secuencias no canónicas que no contienen el motivo consenso xPxxP y que presentan modos de unión alternativos o incluso no contienen ningún resto prolina.



**Figura 1.4.** Estructura en solución del complejo formado por la proteína Src-SH3 y el ligando VSL12 (). El dominio SH3 de Src se muestra en estructura tipo cinta azul claro, y el ligando VSL12 en estructura de varillas en color turquesa. Los restos cargados Asp79 en proteína y R6 en el ligando, se representan en estructura de varillas en rojo y azul respectivamente, y el puente salino que forman como una línea discontinua. Pdb: 1qwf.

Así, por ejemplo, el dominio SH3 de la proteína p53BP2 reconoce una serie de restos en la estructura terciaria de la molécula con la que une y que representan un epítipo comprendido por restos no contiguos (Gorina & Pavletich 1996). Otro tipo de superficie de interacción discontinua es la mostrada por el dominio SH3 de Csk, el cual une la secuencia PEST presente en la proteína PEP (Ghose *et al.* 2001). Un tercer ejemplo de este tipo de especificad anómala es el mostrado por la proteína Eps8, la cual presenta una dimerización a través de las interacciones establecidas en el sitio de unión, mediante el reconocimiento de la secuencia PxxDY en lugar de la secuencia canónica PxxP (Kishan *et al.* 1997; Mongiovi *et al.* 1999). Incluso han llegado a describirse interacciones mediadas por dominios SH3 a través de secuencias que no contienen secuencias de prolina, como por ejemplo del tipo RxxK (Liu *et al.* 2003);Kang, 2000 #18}. Lo cierto es que cada vez es mayor el número de epítopos que continen secuencias diferentes a la secuencia canónica xPxxP descrita para los dominios SH3, lo que nos indica la promiscuidad presentada por este tipo

## **Antecedentes**

de dominios, y por tanto la posibilidad de su implicación en determinados procesos celulares que probablemente todavía no conozcamos.

La implicación de los dominios SH3 en el desarrollo de importantes patologías los hace, por tanto, atractivas dianas para el diseño de fármacos. De hecho, estudios recientes demuestran que compuestos con la capacidad para bloquear o modular interacciones mediadas por estos dominios ostentan buen potencial como agentes terapéuticos (Kardinal *et al.* 2000; Lee *et al.* 2002; Oneyama *et al.* 2002; Stauffer *et al.* 1997; Vidal *et al.* 2001). Así, por ejemplo, se ha descrito que determinados péptidos que penetran en la célula y unen con alta afinidad al dominio SH3 de CRKL inhiben la interacción de éste con la proteína Bcr-Abl, lo que afecta negativamente a la proliferación de células con leucemia crónica mielagínosa en pacientes infectados (Kardinal *et al.* 2000). Otro caso es el observado con la proteína Lck para la que se ha observado que el bloqueo de las interacciones que establece su dominio SH3 con un péptido derivado de ácido palmítico reduce el crecimiento de tumores cancerígenos, tanto *in vitro* como *in vivo* (Lee *et al.* 2002).

La mayor limitación a estas aplicaciones surge porque el diseño de ligandos de alta afinidad y especificidad para dominios SH3 (fármacos selectivos) se ha revelado como un reto especialmente complejo, debido a la baja afinidad de los ligandos naturales y al complicado balance entre especificidad y promiscuidad que caracteriza las interacciones mediadas por estos dominios (Cesareni *et al.* 2002; Landgraf *et al.* 2004; Tong *et al.* 2002). Esta situación se complica aún más por la escasez de información termodinámica, necesaria para conocer los determinantes de la afinidad y especificidad de unión, o dicho en otras palabras, la naturaleza y magnitud de las fuerzas que dirigen la interacción. Los estudios termodinámicos sobre las interacciones mediadas por MRP se reducen a unos pocos trabajos que presentan algunos experimentos puntuales con un grupo limitado de sistemas (Arold *et al.* 1998; Ferreon & Hilser 2004; Renzoni *et al.* 1996; Wang *et al.* 2001). En este contexto, de acuerdo con la información estructural, el reconocimiento de ligandos ricos en prolina por dominios SH3 está basado fundamentalmente en la intercalación de los restos prolina del ligando entre aminoácidos aromáticos e hidrofóbicos del sitio de unión del dominio, con muy escasas interacciones polares. Es de esperar que este tipo de interacciones presente un patrón termodinámico dominado por el

efecto hidrofóbico, que vendría determinado por el componente entrópico. La entropía de unión favorable sería debida al incremento en los grados de libertad de las moléculas de agua que se liberan cuando las superficies hidrofóbicas del ligando y la proteína entran en contacto en el complejo. En estos ejemplos, la entalpía de unión, que refleja el balance entre las interacciones establecidas entre el ligando y la proteína con respecto a las que ambas establecen con el disolvente, es generalmente desfavorable, o en el mejor de los casos ligeramente favorable, a la unión para este tipo de interacción. Sin embargo, como se obtiene para todos los dominios SH3 estudiados por el momento, cuyos resultados se presentarán en la Tabla 3.1 de la sección de resultados, el reconocimiento de ligandos ricos en prolina por dominios SH3 está gobernado por una entalpía de unión favorable a la que se opone parcialmente una contribución entrópica desfavorable. Este comportamiento termodinámico no se puede racionalizar fácilmente considerando exclusivamente las interacciones directas entre interfases hidrofóbicas y apunta hacia una mayor complejidad en estas interacciones. En este contexto, en esta Memoria de Tesis Doctoral se presenta un estudio termodinámico-estructural del reconocimiento de ligandos ricos en prolina por dominios SH3, indispensable para la racionalización de este patrón termodinámico, difícilmente interpretable en términos exclusivamente estructurales a tenor de lo expuesto. Para abordar este estudio termodinámico hemos escogido como objeto de estudio la interacción del dominio SH3 de Abl con un ligando de diseño (Pisabarro *et al.* 1998).

#### **1.1.3.1. Dominio SH3 de Abl (*Abelson Leukemia virus*)**

La forma celular del virus leucémico de Abl es una tirosina quinasa constituida por un dominio con actividad quinasa, un dominio SH2, un dominio SH3 y una secuencia sin estructura secundaria que conecta el dominio SH2 con el dominio quinasa. En la forma inactiva de la quinasa esta secuencia, que adopta una conformación PPII, interacciona con el dominio SH3 análogamente a como lo hacen los ligandos ricos en prolina (Nagar *et al.* 2006). Se ha demostrado que algunas mutaciones y deleciones en el dominio SH3 de Abl, el cual juega un papel importante en la regulación de la actividad quinasa, conducen a la activación constitutiva de la enzima. Es necesario comentar la importancia farmacológica de esta funcionalidad, puesto que la desregulación de la actividad quinasa de Abl

## **Antecedentes**

conduce al desarrollo de la leucemia crónica mielaginososa (Cirinna *et al.* 2000; Skorski *et al.* 1998). En este contexto, estudios recientes han descrito que la utilización de ligandos de alta afinidad, que compiten con las interacciones intramoleculares establecidas por el dominio SH3 en el contexto de la quinasa, aumenta el efecto antitumoral del fármaco Gleevec, dirigido hacia la inhibición de la actividad quinasa. (Azam *et al.* 2003; Barila & Superti-Furga 1998).

El dominio SH3 de Abl ha sido extensamente estudiado, tanto desde el punto de vista estructural como funcional. La estabilidad y el equilibrio conformacional de este dominio se ha estudiado en nuestro grupo mediante calorimetría diferencial de barrido y se ha establecido que su desplegamiento se describe a través de un modelo de dos estados (Filimonov *et al.* 1999), siendo los parámetros termodinámicos los típicos para proteínas globulares y similares a los obtenidos para otros dominios SH3 (Martinez 1995; Martinez *et al.* 1998; Martinez *et al.* 1999).

Desde principios de la década de los 90 varios grupos han trabajado en la identificación y optimización de ligandos para Abl-SH3. El grupo de investigación del Dr. Baltimore (Centro Rockefeller, Nueva York, USA), utilizando este dominio marcado con biotina y unido a GST frente a una librería de expresión de ADN bacteriófago (Cicchetti *et al.* 1992), consiguió identificar dos secuencias como ligandos de Abl-SH3, que se denominaron 3BP-1 y 3BP-2, y que contenían sendas zonas ricas en prolina. Mediante una mutagénesis sistemática por alanina establecieron que estos ligandos tenían en común la secuencia PxxP, que define hoy día la secuencia consenso para la unión a la mayoría de dominios SH3. El grupo del Dr. Luis Serrano (EMBL, Heidelberg, Alemania) ha trabajado extensamente en el diseño mediante técnicas de modelado molecular de ligandos de alta afinidad y especificidad para este dominio, basados en las secuencias de los ligandos 3BP-1 y 3BP-2 (Pisabarro *et al.* 1994; Pisabarro & Serrano 1996). Como resultado de estos trabajos se dispone de un amplio conjunto de ligandos con afinidades optimizadas, entre los que destacan los péptidos p40 (APTYSPPPPP) y p41 (APSYSPPPPP) como los de mayor afinidad ( $K_d$  en torno a 2  $\mu$ M) y especificidad por Abl-SH3 (Pisabarro & Serrano 1996). Este estudio supuso un ejemplo de diseño satisfactorio ya que se consiguió aumentar afinidad por el dominio más de 30 veces con respecto al ligando natural 3BP-1 (APTMPPPLPP), que presenta un valor de  $K_d$  de 34  $\mu$ M.

Hoy día se dispone de estructuras cristalográficas de alta calidad para la quinasa completa (Nagar *et al.* 2003), para el dominio Abl-SH3 aislado y unido a varios ligandos ricos en prolina. La estructura del dominio SH3 fue resuelta por el grupo del Dr. Saraste (Musacchio *et al.* 1994), y presenta el plegamiento típico de estos dominios descrito anteriormente. Las estructuras cristalográficas de los complejos con los ligandos 3BP1 y p41 (Musacchio *et al.* 1994; Pisabarro *et al.* 1998) muestran el modo de unión característico de los dominios SH3, donde cada uno de los dipéptidos xP de la secuencia consenso PxxP se inserta en uno de los dos bolsillos hidrofóbicos. En el caso del complejo con p41, también se establecen dos nuevos puentes de hidrógeno mediados por el resto Y4 con los restos Ser75 y Asp77 en el lazo -RT del bolsillo de especificidad.

Su relevancia biológica, la abundancia de información estructural de alta calidad, la existencia de un amplio juego de ligandos relacionados entre sí en cuanto a secuencia y caracterizados por constantes de unión en el intervalo óptimo para la caracterización de la energética de unión mediante calorimetría isotérmica de titulación, hacen del dominio SH3 de Abl el sistema ideal para llevar a cabo un estudio termodinámico-estructural detallado de las interacciones establecidas por los dominios SH3.

#### **1.1.4. Dominios WW**

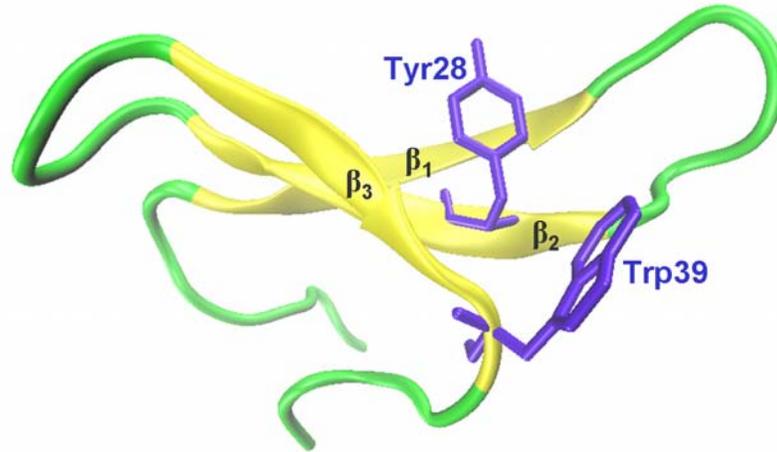
Los dominios WW son módulos de reconocimiento de secuencias ricas en prolina que forman parte de muchas proteínas del citoplasma y del núcleo celular. Desempeñan diversas funciones involucradas en la regulación de la transcripción, procesamiento de ARN, reconocimiento de ubiquitina, tráfico de proteínas, receptores de señalización y control del citoesqueleto (Kay *et al.* 2000; Macias *et al.* 2002; Sudol & Hunter 2000; Sudol *et al.* 2001). Estos dominios están así implicados en el desarrollo de algunas enfermedades humanas que incluyen el síndrome de hipertensión de Liddle, la distrofia muscular, las enfermedades de Alzheimer y Huntington, el cáncer, e infecciones víricas como el Ebola (Sudol & Hunter 2000).

Los dominios WW fueron identificados por primera vez como una repetición imperfecta de 38 restos intercalada en una proteína asociada a la quinasa Yes, a la

## **Antecedentes**

que se llamó Yap (Bork & Sudol 1994; Sudol 1994). El análisis de la secuencia reveló la presencia de dos triptófanos altamente conservados, espaciados entre 20 y 22 restos en las secuencias, que dieron nombre a estos dominios. Al igual que los dominios SH3, desde un punto de vista estructural están caracterizados por una estructura común y muy compacta formada por tres hebras beta, denominadas  $\beta_1$ ,  $\beta_2$  y  $\beta_3$ , que constituyen una lámina beta antiparalela ligeramente curvada, ilustrada en la Figura 1.5 (Ball *et al.* 2005; Macias *et al.* 2002; Sudol & Hunter 2000; Zarrinpar & Lim 2000). Estas tres hebras beta se encuentran conectadas por dos lazos variables en extensión entre los diferentes dominios, los cuales, aunque parece que en menor medida que en los dominios SH3, contribuyen a la especificidad y afinidad de unión a los ligandos (Ball *et al.* 2005; Chong *et al.* 2006; Macias *et al.* 2002; Zarrinpar & Lim 2000).

Existen algunos estudios sobre la estabilidad y el equilibrio de plegamiento/desplegamiento de algunos dominios WW, como Yap (Crane *et al.* 2000; Ferguson *et al.* 2001; Karanicolas & Brooks 2004; Koepf *et al.* 1999a; Koepf *et al.* 1999b), PIN1 (Jager *et al.* 2001; Jager *et al.* 2007; Karanicolas & Brooks 2004) y FBP28 (Ferguson *et al.* 2001; Karanicolas & Brooks 2004; Mu *et al.* 2006; Petrovich *et al.* 2006), que indican que estos dominios presentan una naturaleza monomérica en solución, salvo algunas excepciones (Ohnishi *et al.* 2007), y una elevada reversibilidad en la mayoría de condiciones experimentales utilizadas. En la mayoría de estos trabajos, basados en la utilización de técnicas espectroscópicas, el equilibrio se describe según modelos clásicos de dos estados, aunque estas metodologías no permiten discernir con certeza entre distintos modelos de plegamiento. De hecho, algunos estudios, como los realizados con el dominio WW de Yap revelan diferencias significativas entre las curvas de desplegamiento de fluorescencia y dicroísmo circular, que podrían ser indicativas de un desplegamiento que no es de dos estados (Koepf *et al.* 1999b). De los estudios cinéticos se desprende, como característica general, que el plegamiento de estos dominios ocurre a partir de la formación del primer giro beta entre las hebras  $\beta_1$  y  $\beta_2$ , para las que observa el mayor valor de  $\phi$  en los análisis cinéticos.



**Figura 1.5.** Estructura del dominio WW de la proteína Yap65, formada por las tres láminas beta  $\beta_1$ ,  $\beta_2$ , y  $\beta_3$  (amarillo). Los restos aromáticos altamente conservados que definen el bolsillo hidrofóbico tipo xP donde se insertan los ligandos ricos en Pro se muestran en color violeta. Figura obtenida a partir del código Pdb: 1jmq.

Actualmente, nuestro grupo de investigación está embarcado en el estudio calorimétrico del equilibrio conformacional de varios dominios WW, con el fin de proporcionar una caracterización más profunda de su plegamiento.

En cuanto a las interacciones establecidas por dominios WW, los primeros estudios se basaron en el cribado de librerías de ADN frente al dominio WW de Yap. Esto permitió identificar dos fragmentos que codificaban por dos secuencias diferentes, las cuales tenían en común la secuencia PPPPY (Chen & Sudol 1995). Mediante la sustitución de cada uno de estos restos por alanina se identificó que la secuencia consenso requerida para la unión al dominio WW era PPxY, donde x podría ser cualquier resto. Por esta razón se pensó que el dominio WW era muy parecido al dominio SH3 en cuanto a su habilidad para reconocer secuencias ricas en prolina. Posteriormente, numerosos estudios estructurales y funcionales han mostrado que, según el modo de reconocimiento de los ligandos ricos en prolina, los dominios WW pueden clasificarse en cuatro clases : clase I) los que como Yap65 reconocen la secuencia consenso PPxY, clase II) aquéllos que como FE65 interactúan con secuencias consenso PPLP, muy parecidas a las características de dominios SH3, clase III) los que como Npw38 reconocen secuencias PPII flanqueadas o interrumpidas por los aminoácidos Arg o Lys, del tipo R/LPPP(R), y

## **Antecedentes**

clase IV) los dominios WW, como Pin1, que interaccionan con secuencias que contienen una prolina precedida por fosfoserina o fosfotreonina, del tipo po(S/T)P (Huang *et al.* 2000; Komuro *et al.* 1999; Macias *et al.* 2002; Meiyappan *et al.* 2007). En la Figura 1.6 se muestra un esquema con las distintas especificidades y algunos ejemplos representativos. Las clases II) y III) son las que presentan menos especificidad en las secuencias que reconocen, por lo que existen algunos trabajos en los que se ha propuesto una reclasificación en nuevas subclases, de modo que la clase II se dividiría a su vez en II-a si la especificidad está determinada por restos situados exclusivamente en el extremo N-terminal del ligando, y II-b si la especificidad depende de aminoácidos en ambos extremos; y por otro lado la clase III se divide a su vez en III-L, si los ligandos contienen el motivo PPLP, o III-poli-P, si además contienen otra secuencia poliprolina después de la secuencia consenso PPLP (Otte *et al.* 2003). Además, de acuerdo con la similitud estructural observada entre las clases II y III y el parecido en las secuencias que reconocen, opuestamente también se ha propuesto que estas dos clases podrían agruparse en una mayor denominada clase II/III (Meiyappan *et al.* 2007).

El sitio de unión de los dominios WW, representado esquemáticamente en la Figura 1.7, es una superficie cóncava que contiene una serie de restos aromáticos expuestos, cuya disposición origina un bolsillo hidrofóbico donde se acomodan los péptidos con conformación PPII. Este bolsillo está presente en todas las clases de especificidad y está formado por dos restos aromáticos altamente conservados (Tyr28 en la hebra  $\beta_2$  y Trp39 en la hebra  $\beta_3$ ; Figura 1.5) entre los que se inserta el dipéptido xP, de un modo similar al observado en los dominios SH3 (Sudol & Hunter 2000; Zarrinpar & Lim 2000). Además, existe otro bolsillo de especificidad más variable, que se localiza contiguamente al bolsillo hidrofóbico xP, y que en la clase I está optimizado para acomodar el resto Tyr en el motivo PPxY. Además de estos bolsillos, las clases II y III, presentan un tercer bolsillo hidrofóbico menos profundo, denominado como xP2, que está formado por un resto Tyr conservado en la hebra  $\beta_2$  y otros restos aromáticos más variables presentes en esta misma hebra (Kato *et al.* 2006; Pires *et al.* 2005). Así se justifica que esta clase de dominios sea la que presente más similitud con los dominios SH3, también caracterizados por dos bolsillos hidrofóbicos parecidos, existiendo ejemplos de dominios SH3 y WW no relacionados evolutivamente que han convergido hacia un mecanismo común

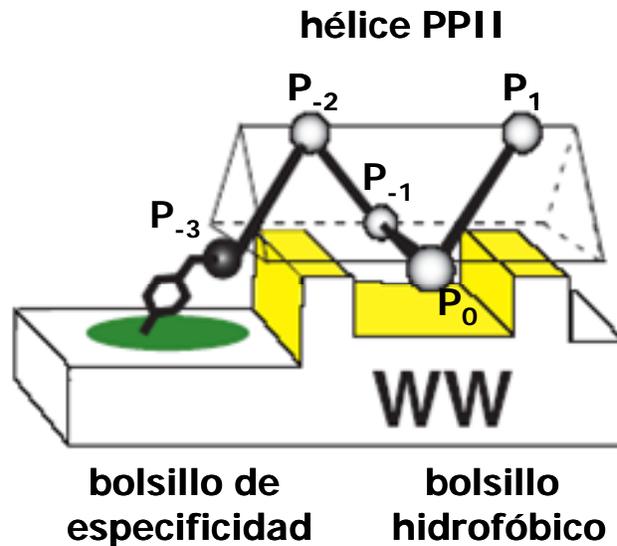
Dominio	ID	Nr*	Secuencia**
Prototipo	IE0M		...GSMGLPPGWDEYKTHN.GKTYYYNHNTKTSWLDPRMSS..
			11                    21 24ab                    31                    41
h WWP3-1	O00309	17	AEDNLGGLPENWEMAYTEN.GEVYFIDHNTKTTSLDPRSLNKQ
y RSP5-3	P39940	40	PVSQGLGLPSGWEMRLTNT.ARUYFVDHNTKTTTWDDPRLPSSL
h NEDD4-2	P46934	12	LLPTSSGLPPGWEEKQDER.GRSYYVDHNSRTTWTKPTVQATY
y RSP5-2	P39940	9	TTSGLGELPSGWQRFTPE.GRAYFVDHNTKTTTWDDPRRQOYI
h NEDD4-3	P46934	13	SEISQGLPKGWVVRHAPN.GRPFYIDHNTKTTTWEDPRLKIPA
h YAP65	P46937	10	EIPDDVPLPAGWEMAKTSS.GQRYFLNHIDQTTTWQDPRKAMLS
y RSP5-1	P39940	8	FEDQYGRLLPPGWERRDNF.GRTYYVDHNTKTTWKRPDLQTE
m ITCHY-1	O54971	15	NTVSQAPLPPGWQRVDQH.GRVYYVDHVEKRTTWDRPEPLPPG
m ITCHY-3	O54971	16	EFDPLGLPLPPGWKRTDSN.GRVYFVNHNTKTTTWEDPSSQQL
h WWP3-2	O00309	18	ELDSELELPAGWEEKIEDPVY.G.YYYVDHINRKTQYENPVLEAKR
h SMURF1-2	Q9UJT8	37	NSDELGLPLPPGWVSRVSVS.GRIYFVDHNTKTTTQFTDPRLHHIM
h NEDD4-4	P46934	14	TSNDLGLPLPPGWERTHTD.GRIYFINHNTKTTTWEDPRLNVA
h NEDD4-1	P46934	11	QQEQSPPLPPGWERRQDIL.GRTYYVNHESRTTQWKRPTPQDNL
h WWOX-2	Q9NZC7	32	RKRVAADLPYGWEQETDEN.GQVFFVDHINKRTTYLDPRLAFT
h DYSTRO	P11532	39	QHFLSTSVQGFWRERAIISP.N.KVPPYINHETQTTWCDHPKNTELY
h PEPP2-2	Q9HAU0	33	HRRQSTDLPTGWEEAYTPE.GARYYINHNERKVTSKHPVTGQPS
h SMURF2-1	Q9H260	38	SRLFDNDLPDGWEERTAS.GRIQYLNHITRTTQWERPTRPASE
h PEPP2-1	Q9HAU0	34	LNLEWISLPSWYTYGITRG.GRVFFINEEAKSTWLRHPVTGEAV
h WWOX-1	Q9NZC7	31	DTDSEDELPPGWERTTKD.GWVYVANHTEEKTOWEHPKTKGRK
y YJ08	P46995	3	IDLRRVRLPPGWETIHEN.GRPLYYNAEQTKLHYPPSSGSSKVF
<b>Clase I</b>	<b>Cons.</b>		.....LPQWE.....G.YFV.H.R.T.H.P.....
h HYP109-1	Q9NT81	26	SLAGVGIEMGDWQEVMDENTGSYYWNTQTNVTVWELPQYLATQ
h PQBP1	O60828	22	EATRLEGLPPSMYKVFDPSSGLPYYWNIADTDLVSWLSPHDPSV
h FE65	O00213	20	AFETDSDLPAGWVRVQDTS.G.TYYWNIPTGTTQWEPGRASPSQ
y YFB0	P43582	5	SKSNPQVPSGWKAVFDDEYQTYYYVDLSTNSSQWEPFRGTWP
<b>Clase IIa</b>	<b>Cons.</b>		.....P..W..V.D...G..YYW...T..S..WE.P.....
h FBP21-1	O75554	27	KKRKKDPSKGRWVEGITSE.GYHYYYDLISGASQWEKPEGFQGD
h FBP21-2	O75554	28	GDLKKTAVKTVWVWVGLSED.GPTYYNTETGESRWEKPDFIPH
<b>Clase IIb</b>			
h FBP11-2	O75400	24	TPAEQLLSKSPWKEYKSDS.GKPYYYNSQTKESRWAKPKLEDL
h FBP11-1	O75400	23	AAGTASGAKSNWTEHKSPD.GRTYYNTETKQSTWEKPDCLKTP
<b>Clase III L</b>			
y PRP40-2	P33203	7	SQEELLRENGWKAAKTAD.GKYYYNPTTRETSTWIPAFKVK
h GAS7	O60861	21	BESQTVILPPGWQSYLSPQ.GRRYYVNTTINETTWERPSSSPI
<b>Clase III poli-P</b>			
h PIN1	Q13526	1	MA.DEEKLPPGWKRMRSRSGRVYYFNHITNASQWERPSGNSSS
y ESS1	P22696	2	DVASRTGLPTFWTVRYSKSKKREYFFNPETKHSQWEEPBTNK
<b>Clase IV</b>			
h HYP109-2	Q9NT81	25	KQYEINATPKWSSHWDRDHRRYFYVNBQSGESQWEPDGESEE
y PRP40-1	P33203	6	.....MSIWKEAKDAS.GRIYYYNLTKKSTWEKPNELISQ
y YPR152C	Q06525	4	.....MRGEMQEPKTPA.GKYYYNKNTKQSRWEKPNLKKGS
h CA150-1	O14776	29	GTPALPPTTEIHWENKTPD.GKYYYNARTRESAWTKPDGVKV
h IQGAP1	P46940	19	KKLAVGDNNSKWKHWVKG.GYHYHNLETQEGGWDEPPNFVN
h CA150-3	O14776	30	PVATAPIPTPMSVWVWVTD.ERVFFYNPTIRLSHWDRPDDLIGR
h SMURF1-1	Q9UJT8	36	HGHQSPPELPEGYEORTTVQ.GQVYFLHTQTGVSTWHDPRIPRDL
y SSM4	P40318	41	PEAYKPTSIISWKNPNTILL..TLYFTKRILESSSYVVKPLLERYW
h KIAA1052	Q9UPV0	35	REGIVAPLPGWKPQDIT.GDIYYFNFANGQSNWHDHPDEHYR
y TIN1	P53076	42	AINIETESRNLWDTSDKNS.GLOYYPDQSPSSSFSSPRVSSGN

**Figura 1.6.** Clasificación de dominios WW en las clases I, II, III, y IV, de acuerdo con el tipo de secuencias que reconocen. La clase I aparece en el recuadro rojo, la II en verde, la III en amarillo, y la IV en azul turquesa. La clase II se puede dividir, a su vez, en IIa y IIb, mientras que la III, en los que reconocen únicamente secuencias PPLP (L), y en los dominios que además contienen una secuencia contigua poliprolina (poli-P). En la parte inferior de cada recuadro se indican los restos más conservados (Cons.) para las clases I y II. ID: representa el código de acceso para el banco de Proteínas (Pdb). \*Nr: número de domino WW en la clasificación realizada por Otte et al. (Otte et al. 2003. \*\* Secuencia numerada respecto a la proteína Yap-WW {Pires, 2001 #514}.

## **Antecedentes**

para el reconocimiento de secuencias en conformación PPII (Macias *et al.* 2002; Zarrinpar & Lim 2000).

Estudios estructurales de varios dominios WW y sus complejos (distrofina (Huang *et al.* 2000), Yap65 (Macias *et al.* 1996), Nedd4 (Kanelis *et al.* 2006) y Smurf2 (Chong *et al.* 2006) de la clase I; Pin 1 (Verdecia *et al.* 2000) de la clase IV; y FBP11 de la clase II/III (Kato *et al.* 2006; Pires *et al.* 2005)) han proporcionado detalles importantes acerca del modo de unión que caracteriza a estos dominios modulares. En todos los casos, los ligandos peptídicos adoptan una conformación PPII que se inserta en los bolsillos hidrofóbicos xP y xP2 según la clase de dominio, mientras que, al igual que ocurría con dominios SH3, algunos carbonilos de las prolinas ocasionalmente establecen puentes de hidrógeno con algunos restos del dominio. También se ha descrito que las interacciones que se establecen con los lazos que conectan las hebras beta del dominio WW, variables en secuencia, y, en ocasiones, con restos fuera del sitio de unión u otros dominios adyacentes, contribuyen adicionalmente a la afinidad y especificidad (Chong *et al.* 2006; Macias *et al.* 2002; Zarrinpar & Lim 2000). En la Figura 1.7 se ilustra el modo de unión característico de los dominios WW de clase I, los más abundantes y mejor estudiados, (Kanelis *et al.* 2006; Macias *et al.* 1996). En estos casos, las dos prolinas de la secuencia consenso PPxY del ligando se acomodan en el bolsillo hidrofóbico xP frente a los restos aromáticos conservados Tyr28 y Trp39, como se ilustran para en la Figura 1.8, donde se representa un complejo del dominio Yap65-WW con un ligando rico en prolina. La Tyr de la secuencia PPxY se acomoda en el bolsillo de especificidad, que presenta más variabilidad entre los diferentes dominios, y establece un puente de hidrógeno con la His32 (Macias *et al.* 1996). La interacción se caracteriza por la ausencia de interacciones polares, a excepción del puente de hidrógeno anterior, y los establecidos por la Thr/Ser en la posición 37 del dominio con el carbonilo de segunda Prolina de la secuencia PPxY y otro entre el Trp39 y el carbonilo del resto precedente a la secuencia PPxY, ambos altamente conservados (Macias *et al.* 1996; Meiyappan *et al.* 2007; Pires *et al.* 2001). Al igual que en el caso de los dominios SH3, los dominios WW pueden unir los ligandos con dos orientaciones.

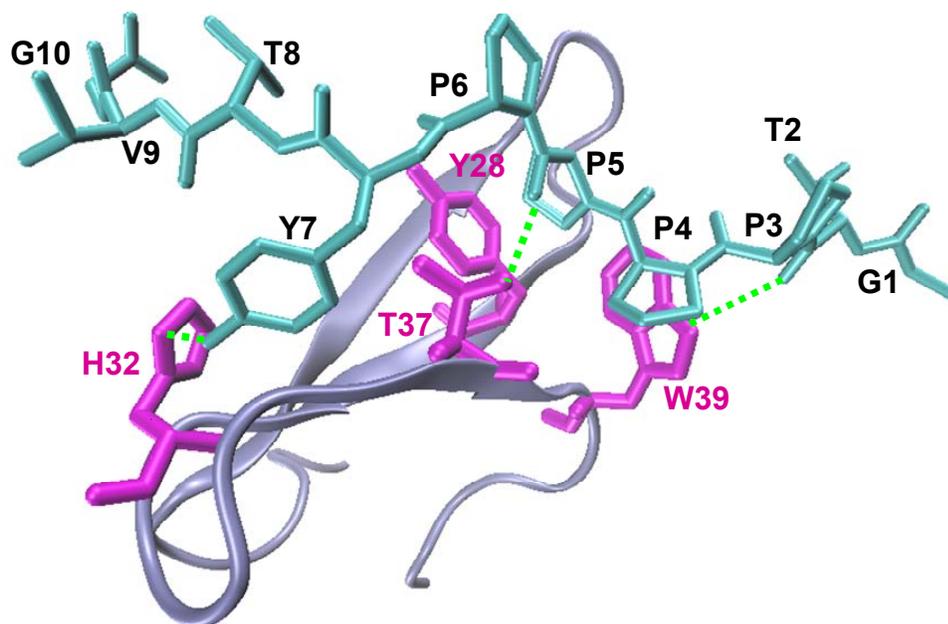


**Figura 1.7.** Representación esquemática del modo de unión mostrado por dominios WW de la clase I. En la figura se representa el sitio de unión modelado con un ligando tipo PPxY. El sitio de unión contiene un bolsillo hidrofóbico xP formado por restos aromáticos conservados (amarillo) y el bolsillo de especificidad adyacente que es más variable (verde). Las letras P en el ligando indican las posiciones en su secuencia,  $P_0P_{-1}x_2Y_3$ .

Por ejemplo, el dominio WW de Pin1 de la clase IV une ligandos de tipo I (orientación N-C) y el dominio WW de distrofina de la clase I a ligandos de tipo II (orientación C-N). La orientación del ligando parece venir determinada por restos que no son Pro dentro de la secuencia consenso que reconoce cada clase, como por ejemplo el resto Tyr en la secuencia consenso PPxY de la clase I. Sin embargo, y a diferencia de otros dominios modulares como EVH1 y los propios SH3, los restos que componen la secuencia consenso parecen ser más importantes tanto para la afinidad como para la especificidad de unión, lo que indica que estos dominios tienen un epítipo de reconocimiento muy corto. Incluso un ligando del tipo Ac-PPPPY-NH<sub>2</sub> parece unir con casi la misma afinidad que otros ligandos mayores al dominio WW de Yap65 (Pires *et al.* 2001). Estructuralmente, esta característica podría justificarse en base al pequeño tamaño de los dominios WW y a la reducida superficie que exponen para interactuar con los ligandos.

Estudios de unión de ligandos ricos en prolina, realizados mediante fluorescencia, indican que el valor de las constantes de disociación varía generalmente entre 3 y 350  $\mu\text{M}$  {Macias, 1996 #155;Nguyen, 1998 #6;Pires, 2001

## Antecedentes



**Figura 1.8.** Estructura en solución del complejo formado por el dominio WW de Yap65 con el ligando rico en prolina GTPPPPYTVG. El ligando se representa en color turquesa en estructura de varillas, y la proteína en estructura tipo cinta en azul claro. Los restos de la proteína que establecen puentes de hidrógeno con el ligando (líneas verdes discontinuas) se muestran en fucsia y estructura de varillas. Pdb: 1jmq

#514;Pires, 2005 #287;Kanelis, 2006 #286}. En el caso del dominio WW de Yap65, considerado como prototipo de los dominios WW de clase I y que es el más ampliamente estudiado (Macias *et al.* 1996), ha sido posible obtener algunos ligandos con alta afinidad mediante estrategias de diseño racional (Pires *et al.* 2001). Concretamente, el péptido que denominaremos como **M1** a lo largo de esta Memoria, de secuencia **EYPPYPPPPYPSG**, representa el ligando de mayor afinidad descrito hasta la fecha para el dominio WW de Yap65 ( $K_d = 6 \mu\text{M}$ ). Como ya hemos comentado al inicio de esta Memoria, la información termodinámica sobre el reconocimiento de secuencias ricas en prolina por estos dominios es prácticamente inexistente. Sólo muy recientemente se han publicado algunos experimentos puntuales con el dominio WW de FE65 y el ligando rico en prolina de la proteína Mena (PPPPPLPP) (Meiyappan *et al.* 2007) y con el dominio WW3 de Itch con péptido tipo PPxY (Morales *et al.* 2007).

### **1.1.5. DOMINIOS UEV (UBIQUITIN E2 VARIANT)**

Los dominios UEV son mucho menos abundantes que las familias de dominios WW y SH3, lo que posiblemente responda a una función más especializada (Martin-Serrano *et al.* 2001). De hecho, se han identificado hasta el momento únicamente en las proteínas Tsg101 y Mms2 y presentan gran similitud con el dominio catalítico UBCc (catalítico conjugado de ubiquitina) de enzimas E2, tanto en secuencia como en organización estructural. Las enzimas tipo E2, junto con las enzimas tipo E1 y E3, forman parte de un complejo enzimático presente en la ruta de conjugación de ubiquitina, necesaria para el marcaje de determinados sustratos para su degradación (Pickart 2004). Concretamente, la enzima E2 recibe una molécula de ubiquitina a través de la enzima E1, y después es transportada hasta la enzima E3. Una vez formado el complejo ternario E2-E3-ubiquitina, el sustrato es reconocido y marcado con esta molécula de ubiquitina mediante el sitio activo de cisteína de la enzima E3 (Hershko & Ciechanover 1998).

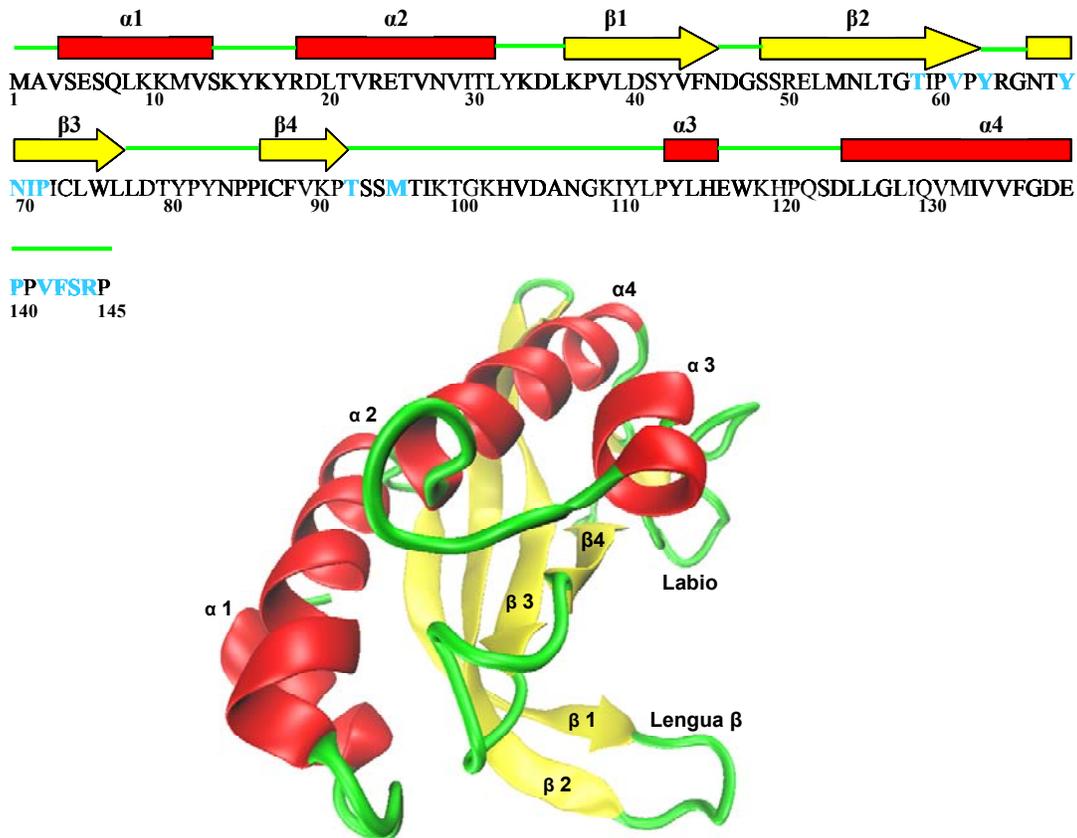
Las proteínas tipo Mms2 sirven como cofactores en reacciones de ubiquitinación que llevan a la formación de cadenas de poliubiquitina tipo K63 (Moraes *et al.* 2001; VanDemark *et al.* 2001). Mms2 forma un heterodímero con la enzima Ubc13, única enzima tipo E2 a la que une, que es fundamental para la formación de la cadena de poliubiquitina (Hofmann & Pickart 1999; VanDemark *et al.* 2001). La proteína Tsg101 (*Tumour Susceptibility Gene 101*), cuyo dominio UEV estudiaremos en este trabajo, es una proteína clasificadora del sistema vacuolar que forma parte del complejo ESCRT-I, que juega un papel fundamental para la inclusión de proteínas en vesículas vacuolares y en algunos compartimentos del endosoma (Katzmann *et al.* 2001; Lemmon & Traub 2000).

La información estructural para los dominios tipo UEV es mucho más limitada. Se dispone de las estructuras en solución del dominio UEV de Tsg101 humana, libre, unido a un péptido del dominio L del virus del SIDA (PEPTAPPEE) (Owen Pornillos *et al.* 2002a) y unido a ubiquitina (Sundquist *et al.* 2004); también tenemos las estructuras cristalográficas del dominio Mms2-UEV humano (Moraes *et al.* 2001) y del dominio UEV de Vps23 (homóloga de Tsg101) de levadura (Teo *et al.* 2004). La estructura de los dominios UEV es similar a la adoptada por proteínas del tipo E2

## Antecedentes

salvo por la ausencia de las dos hélices alfa C-terminales características de éstas. De este modo, se genera un sitio de unión adicional en los dominios UEV para el reconocimiento de secuencias ricas en prolina además del sitio de unión para ubiquitina, común a ambos tipos de proteínas (Koonin & Abagyan 1997).

La estructura del dominio Tsg101-UEV, que se muestra en la Figura 1.9, está constituida por cuatro hélices alfa y una lámina beta antiparalela formada por cuatro hebras (Owen Pornillos *et al.* 2002a). Existen algunas diferencias estructurales entre los dominios UEV de Tsg101 y Mms2, de modo que el dominio UEV de Tsg101 contiene dos hélices en el extremo N-terminal mientras que Mms2 sólo presenta una. Además, la disposición de las hebras beta 1 y 2 es diferente en ambos dominios, lo que contribuye a que Tsg101-UEV presente un empaquetamiento mucho más compacto de las hélices alfa contra la lámina beta antiparalela. En ambos dominios, el sitio de unión para ubiquitina está delimitado por el lazo que conecta las hebras  $\beta 3$  y  $\beta 4$ , denominado lengua beta, la hélice  $\alpha 4$  y algunos restos en la hebra  $\beta 4$ , y el labio, que forma parte del lazo que conecta la hélice  $\alpha 3$  y la hebra  $\beta 4$ . El sitio de unión para  ${}^7\text{PTAP}^{-10}$  consiste en un pequeño bolsillo hidrofóbico donde se inserta la P7 del ligando, constituido por los restos Thr58, Pro71, Thr92 y Met95; y otro bolsillo hidrofóbico de tipo xP donde se inserta el dipéptido  ${}^9\text{AP}^{-10}$  del ligando, constituido por los restos apolares Ile70, Val89 y Pro91 (que pertenecen al lazo que conecta las hebras  $\beta 3$  y  $\beta 4$  y se localizan en el fondo del bolsillo) junto los restos Gly126 y Gly136 en el extremo C-terminal (O. Pornillos *et al.* 2002). Este último bolsillo es equivalente a los bolsillos xP descritos para los dominios SH3 y WW (Ball *et al.* 2005; Macias *et al.* 2002). A diferencia de los dominios SH3 y WW en los que la parte superior del ligando se encuentra expuesta al disolvente, en los complejos con dominios UEV el ligando se encuentra cubierto por parte del dominio, estableciendo interacciones adicionales con los restos Thr58, Asn69, Ser143, Arg144 (O. Pornillos *et al.* 2002), como se ilustra en la Figura 1.10. Adyacentes a este bolsillo se encuentran dos restos básicos (Arg64 y Arg144) que parecen ser importantes para determinar la orientación del péptido en el sitio de unión, mediante la interacción con los aminoácidos de carácter ácido que frecuentemente flanquean la secuencia consenso PTAP en el ligando (Owen Pornillos *et al.* 2002a). De hecho, se ha demostrado que mutaciones puntuales a alanina de los restos Glu6 y Glu13 en el dominio L del virus del VIH ( ${}^5\text{PEPTAPPEE}^{-13}$ ) hacen que su afinidad disminuya

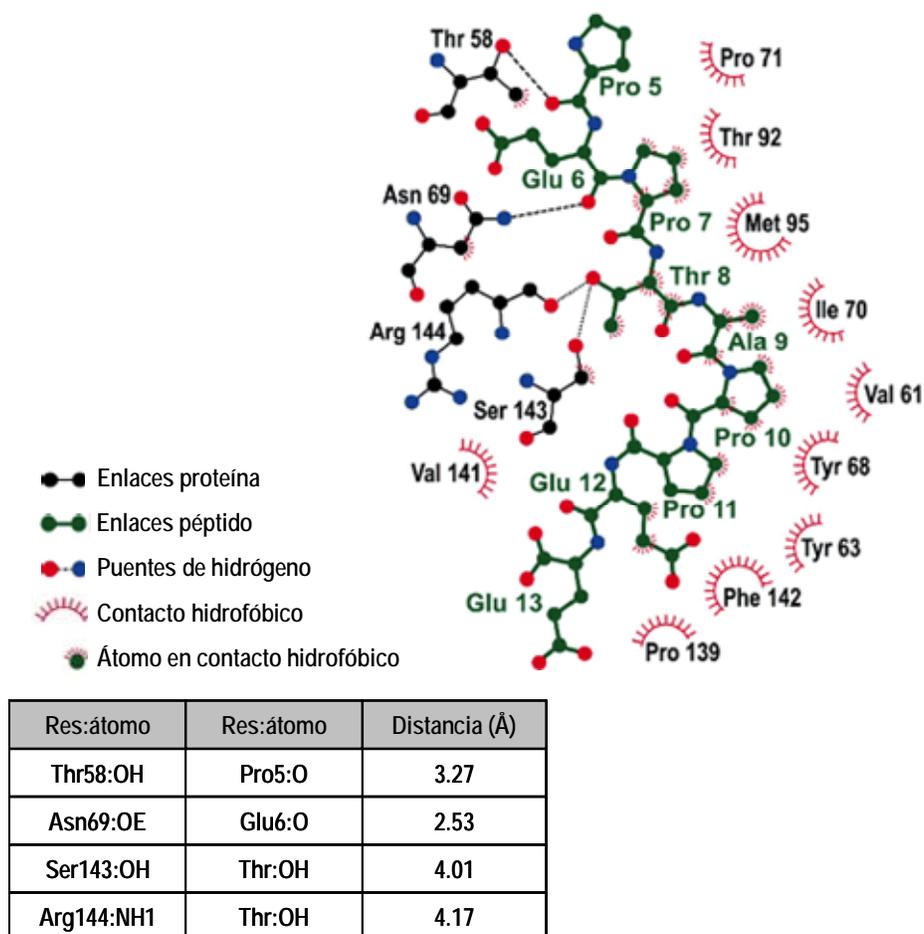


**Figura 1.9.** Secuencia y estructura secundaria del dominio UEV de Tsg101, donde los restos más importantes para la interacción con PTAP se resaltan en turquesa en la secuencia. Las hélices alfa se representan como  $\alpha 1$ -  $\alpha 4$  en rojo y las hebras beta como  $\beta 1$ -  $\beta 4$  en amarillo. En la parte inferior de la figura se representa la estructura en solución de Tsg101-UEV (Pdb: 1kpq) con el mismo código de colores que en la secuencia.

significativamente por Tsg101-UEV (Garrus *et al.* 2001). Por el contrario, la mutación de Glu12 a Ala no produce ningún efecto sobre la afinidad de unión. Además de los restos ya descritos, experimentos mediante RMN indican que hay otros aminoácidos, como Tyr63, Met95 y Val141 que bordean el bolsillo hidrofóbico y que presentan los mayores desplazamientos químicos, junto con Val89, como consecuencia de la unión del péptido PEPTAPPEE. Otros restos que también experimentan desplazamiento químico tras la adición de ligando, y que podrían ser importantes para la interacción, son Val61, Tyr68, Pro139, Phe142 (Figura 1.10).

Estos dominios, menos abundantes y descubiertos recientemente, aún no han sido estudiados desde el punto de vista de su plegamiento, estabilidad y equilibrio

## Antecedentes



**Figura 1.10.** Representación esquemática de las interacciones establecidos entre Tsg101-UEV (esqueleto peptídico en negro) y el ligando PEPTAPPEE del dominio L del virus VIH (esqueleto peptídico en verde). Los átomos de oxígeno y nitrógeno se representan como esferas rojas y azules, respectivamente, los puentes de hidrógeno como líneas negras discontinuas, y las interacciones de tipo hidrofóbico como semiesferas rayadas rojas. En la figura se representan las interacciones más importantes en el complejo resuelto mediante RMN (O. Pornillos et al. 2002). Los puentes de hidrógeno mediados por el oxígeno y de Thr8 del péptido no se observan en algunas de las 15 estructuras finales propuestas. En la tabla inferior se representa la distancia para los puentes de hidrógeno formados entre proteína y ligando.

conformacional. De igual modo, no se dispone hasta la fecha de información termodinámica sobre sus interacciones con ligandos.

### 1.1.6. Dominios L víricos

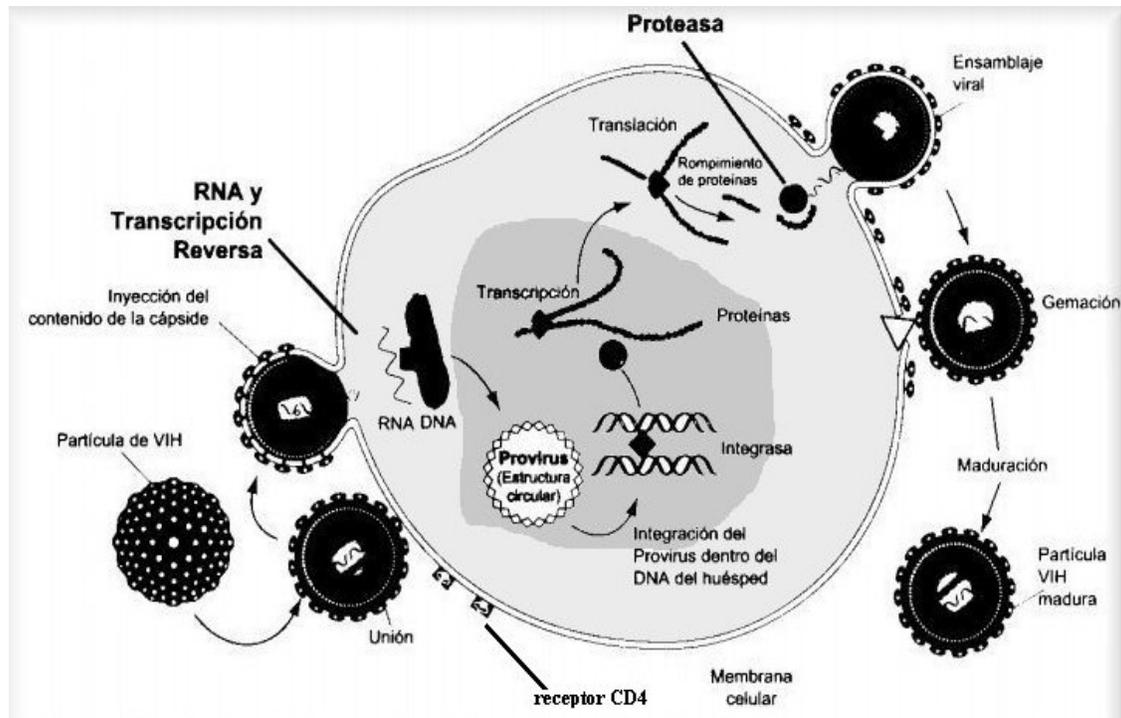
El principal objetivo del trabajo que se presenta en esta Memoria es proporcionar una mejor comprensión de los procesos de reconocimiento de secuencias ricas en prolina por MRP. En este sentido, si profundizar en el estudio de dominios SH3 es importante, también lo es abordar el estudio termodinámico de las

interacciones con ligandos ricos en prolina con otras familias de MRP, como son los dominios WW y UEV, con el fin de establecer si el comportamiento termodinámico anómalo descrito para dominios SH3 es una característica universal del reconocimiento de secuencias ricas en prolina e investigar posibles diferencias entre las distintas familias que pudieran explotarse para optimizar la especificidad de los ligandos.

En este trabajo hemos seleccionado como sistemas representativos de las familias de dominios WW y UEV el tercer dominio WW de la ubiquitina ligasa humana Nedd4 y el dominio UEV de Tsg101 respectivamente, que han sido identificados recientemente como las dianas celulares de algunos dominios L víricos (Demirov & Freed 2004; Freed 2002; Martin-Serrano *et al.* 2001). Estos dominios L (del inglés “*late domains*”) son pequeñas secuencias muy conservadas que se localizan en las poliproteínas gag de algunos retrovirus como el de la leucemia humana o el VIH, en proteínas estructurales de algunos filovirus como el del ébola y radovirus como el de la estomatitis vesicular, cuya presencia e integridad son imprescindibles para el correcto desarrollo de las últimas etapas en el ciclo de vida vírico (Owen Pornillos *et al.* 2002b). El ciclo de vida de todos los virus encapsulados consiste en varias etapas, denominadas como unión de la partícula viral, inyección del RNA en la célula, transcripción reversa e integración en el caso de retrovirus, transcripción, translación, ensamblaje, y gemación, la última de las cuales implica la separación de la membrana del virion de la membrana celular. En la Figura 1.11 se muestra como ejemplo el ciclo de vida del virus VIH, para el que se ha observado que la eliminación o mutación de su dominio L conduce a la congelación del proceso de gemación en sus últimas etapas, haciendo que, como se ilustra en la Figura 1.12, las partículas víricas o bien permanezcan unidas a la membrana celular o bien sean defectuosas, impidiendo la progresión de la infección (Demirov *et al.* 2002; Goila-Gaur *et al.* 2003; Yasuda *et al.* 2003).

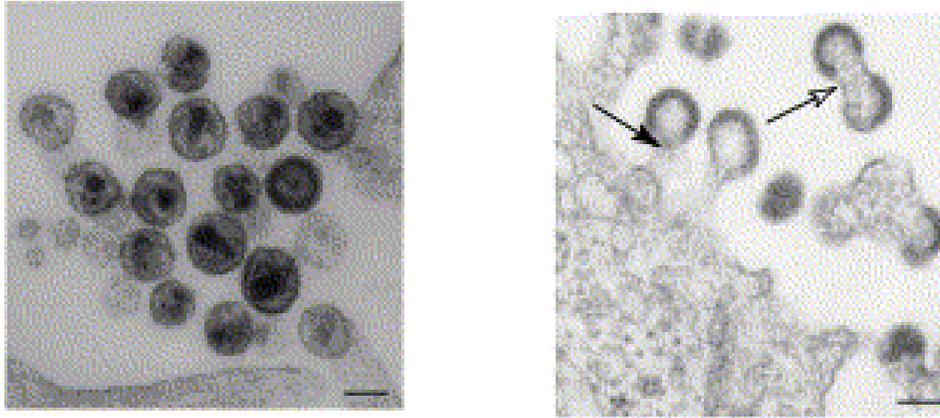
Los distintos tipos de dominios L contienen distintas secuencias conservadas que hasta la fecha se han identificado como tres clases principales de dominios L, denominados tipo PTAP, PPxY y YxxL, que se muestran en la Figura 1.13.

## Antecedentes



*Figura 1.11. Representación simplificada del ciclo de vida del virus 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, transcripción reversa del ARN en ADN, 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.*

Los distintos dominios L parecen funcionar de forma sinérgica e incluso en ocasiones son intercambiables, lo que es indicativo de cierto nivel de redundancia en sus funciones. Las secuencias consenso características de los distintos dominios L coinciden con las propias de los ligandos de distintos dominios modulares mediadores de interacciones proteína-proteína (por ejemplo, el motivo PTAP se corresponde con la secuencia consenso PxxP típica de dominios SH3 y el PPxY con la secuencia característica de ligandos de dominios WW de tipo I). De hecho, hoy en día está bien establecido que el mecanismo de acción de los dominios L se basa en el secuestro de la maquinaria celular encargada de la gestión de las vesículas en la célula hacia las zonas de gemación de los viriones en la membrana celular. Las distintas familias de dominios L reclutan secciones diferentes de la maquinaria celular necesaria para la liberación del virus mediante el establecimiento de interacciones directas con distintas proteínas celulares que participan en estos procesos.

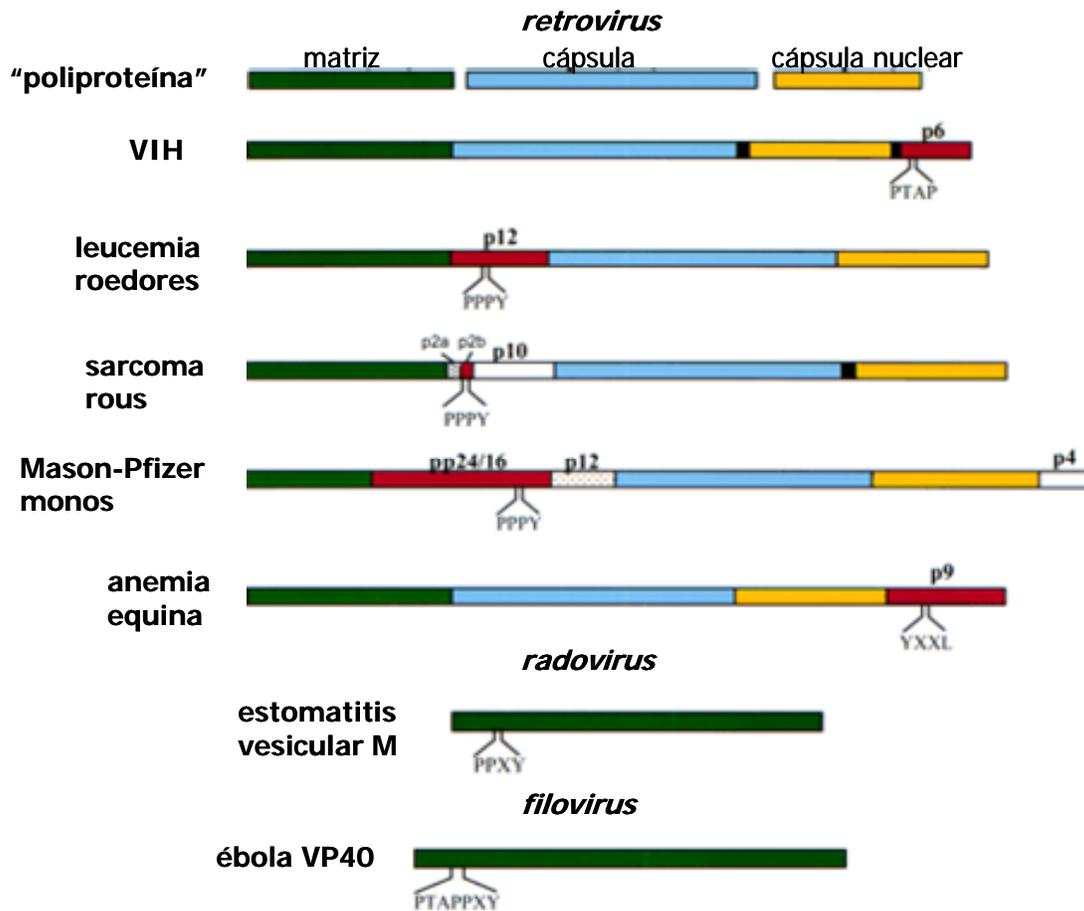


**Figura 1.12.** Fotografías de microscopía electrónica para el proceso de liberación de partículas virales del VIH desde la célula infectada. En el panel de la izquierda se representa la pared de la membrana celular y las partículas virales escapando de la célula como ocurre normalmente, mientras que el panel de la izquierda se representa la misma situación tras su sobreexpresión de la proteína celular Tsg101. En este caso, se puede apreciar como algunas partículas virales permanecen unidas a la célula infectada y no pueden escapar (flecha abierta), o bien la liberación es defectuosa, ya que se aparecen partículas virales que no están aisladas (flecha cerrada). Tomadas de (Freed 2002).

Así, se ha demostrado que los dominios de tipo PTAP interaccionan con el dominio UEV de Tsg101 (Freed 2002; Freed 2003; Owen Pornillos *et al.* 2002a) y que los dominios de tipo PPxY se unen al tercer dominio WW de la ubiquitina ligasa humana de Nedd4. En este sentido, se ha propuesto que compuestos capaces de inhibir o bloquear las interacciones entre estos dominios virales y sus dianas celulares podrían constituir nuevos agentes antivirales de amplio espectro.

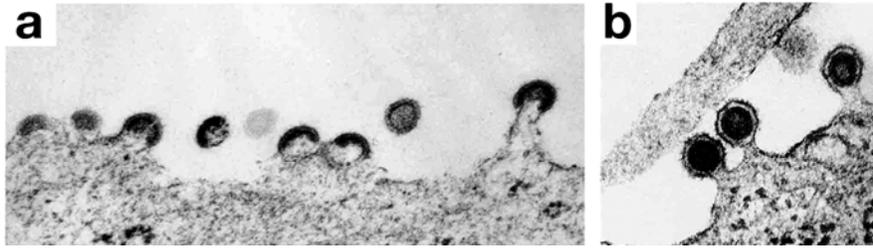
Tsg101 (Tumour Susceptibility Gene 101) es una proteína clasificadora del sistema vacuolar que forma parte del complejo ESCRT-I, que juega un papel fundamental para la inclusión de proteínas en vesículas vacuolares y en algunos compartimentos del endosoma (Katzmann *et al.* 2001; Lemmon & Traub 2000). El complejo ESCRT-I, presente en numerosos organismos, desde levadura o drosófila hasta humanos, parece consistir en un heterotrímero que, por ejemplo, en levadura está constituido por las proteínas Vps23 (homóloga de Tsg101 en humanos), Vps28 y Vsp37 (Katzmann *et al.* 2001; Kostelansky *et al.* 2006). El complejo ESCRT-I está relacionado con otros complejos celulares conocidos como ESCRT-II y ESCRT-III, que complementan las funciones de transporte y formación de vesículas de proteínas en el endosoma para su degradación en el lisosoma.

## Antecedentes



**Figura 1.13.** Localización de dominios L tipo PTAP, PPxY e YxxL en poliproteínas de algunos retrovirus y en proteínas estructurales (matriciales) de radovirus y filovirus. Las poliproteínas retrovirales que contienen dominios L se muestran en rojo; matriz, cápsula, y cápsula nuclear se muestran en verde, azul y amarillo, respectivamente. Los dominios L encontrados en cada caso se indican debajo de cada proteína.

Como ya hemos comentado, el complejo ESCRT-I es usurpado por algunos virus como VIH y ébola para poder escapar de la célula infectada (Goila-Gaur *et al.* 2003; Martin-Serrano *et al.* 2001; Owen Pornillos *et al.* 2002b; Strack *et al.* 2000). En primer lugar, el virus necesita dirigir la maquinaria celular endosomal hasta la pared de la membrana celular, y una vez allí formará una invaginación en ésta, la cual una vez que se seccione permitirá liberar la partícula viral de la célula infectada (Figura 1.14). Es precisamente en el transporte de la maquinaria celular donde las interacciones establecidas por la proteína celular Tsg101 con los dominios L tipo PTAP de algunos virus parecen ser esenciales para la liberación de éstos de la célula infectada (Garrus *et al.* 2001).

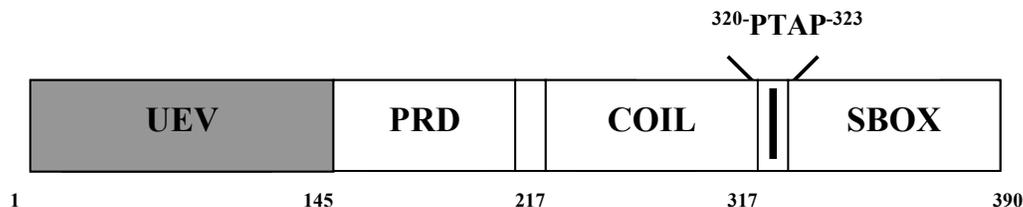


**Figura 1.14.** Proceso de liberación de partículas virales de la célula infectada, a) formación de vesículas conteniendo las partículas virales y transporte de éstas hasta la pared de la membrana celular; b) formación de invaginaciones en la pared de la membrana celular y escape de las partículas virales de la célula. Fotografías tomadas de (Demirov *et al.* 2002; Freed 2003).

Existen algunos trabajos que indican que la mutación o sobreexpresión del dominio L del VIH produce un bloqueo de la eliminación de partículas virales de la célula infectada o en todo caso una liberación defectuosa, lo que confirma la importancia de su interacción con Tsg101 (Demirov *et al.* 2002; Garrus *et al.* 2001).

Desde un punto de vista estructural, Tsg101, representada esquemáticamente en la Figura 1.15, es una proteína multidominio, que consiste en un dominio UEV, un dominio rico en prolina denominado PRD, una zona tipo “coil-coiled” y otro dominio en el extremo N-terminal denominado SBOX. La función del dominio UEV está relacionada con el reconocimiento de ubiquitina, mientras que las zonas “coil-coiled” y SBOX son requeridas para la interacción con otras subunidades del complejo ESCRT-I, además SBOX también parece ser importante para mantener la integridad de la proteína (Katzmann *et al.* 2001). Tsg101 interacciona directamente con proteínas marcadas con ubiquitina para su degradación a través de su dominio UEV, el cual es capaz de unir ubiquitina aunque carece de actividad ligasa, ya que la posición característica de la cisteína catalítica del sitio activo de la ligasa está ocupada por un resto tirosina en el dominio UEV (Garrus *et al.* 2001; Katzmann *et al.* 2001). Como ya hemos comentado, el dominio UEV posee además un sitio de unión para secuencias peptídicas cortas que contienen la secuencia PTAP. El motivo P(T/S)AP está presente, además de en los dominios L víricos (Freed 2002; Owen Pornillos *et al.* 2002b), en algunas proteínas implicadas en la estabilización del genoma, mitosis, transcripción genética, {Freed, 2002 #421; Pornillos, 2002 #334; Strack, 2000 #339; Garrus, 2001 #304}, e incluso algunas proteínas de

## Antecedentes



**Figura 1.15.** Representación esquemática de la estructura de la proteína Tsg101, donde se representa el dominio UEV que reconoce la secuencia PTAP, el dominio rico en Pro (PRD), el dominio no estructurado o tipo coil, y la caja o zona que dota de estabilidad a la proteína (SBOX) (Feng et al. 2000). También se muestra la secuencia PTAP, que a su vez contiene Tsg101, localizada entre la zona tipo coil y la de estabilidad.

membrana (Berthoux et al. 1997; Garrus et al. 2001). La secuencia PSAP también aparece en la proteína Hrs/Hgs de drosófila, requerida para la interacción con tirosina-quinasas y la formación de vesículas para su destrucción (Komada & Kitamura 2001; Lloyd et al. 2002). Incluso la misma proteína Tsg101 contiene una secuencia PTAP, localizada entre las zonas “coil-coiled” y SBOX (Figura 1.15). Existe, de este modo, la posibilidad de que Tsg101 se organice espacialmente para que su dominio UEV se pueda unir a su propia secuencia PTAP, originando una conformación que se auto inhibiría *per se*. Por lo tanto, es posible que la regulación en la unión a secuencias que contienen PTAP sea importante para la función de Tsg101, y algunos virus como el VIH y ébola simplemente han intentado imitar este modo de unión.

El papel de la interacción entre los dominios L de tipo PTAP y el dominio Tsg101-UEV en las etapas de gemación de los virus está bien establecida, ya que mutaciones en restos claves del dominio Tsg101-UEV así como en la secuencia PTAP del virus (Gottlinger et al. 1991; Huang et al. 1995) bloquean la liberación de éste. Así, se ha descrito que las mutaciones por Ala de Tyr63 y Met95 en el dominio Tsg101-UEV producen disminuciones en su afinidad por ligandos peptídicos correspondientes a las secuencias del dominio L del VIH en torno al 14% y al 52% respectivamente, mientras que la mutación por Ala de Val89, situado en el fondo del bolsillo hidrofóbico, no provoca cambio alguno en la afinidad de unión. De forma paralela, los ensayos realizados *in vivo* demuestran que los mutantes de Tyr63 y Met95 por Ala provocan un bloqueo de las partículas virales del VIH e impiden su liberación de la célula infectada, mientras que la mutación de Val89 no produce

ningún efecto (resultados no publicados, J. Garrus). La buena concordancia entre los fenotipos de las partículas virales que se liberan de la célula y los experimentos realizados *in vitro* apoyan la teoría de que las interacciones establecidas entre Tsg101-UEV y la secuencia PTAP son muy importantes para la correcta liberación de las partículas virales desde la célula infectada. Por este motivo, sería de gran interés llegar a conocer en detalle el modo en que el dominio Tsg101-UEV establece las interacciones con los dominios víricos tipo L, lo que es fundamental para sentar las bases moleculares necesarias para el desarrollo de nuevos ligandos con propiedades antivirales.

En el caso del dominio UEV, como hemos comentado previamente, no existe ningún trabajo o estudio de tipo termodinámico de las interacciones que establece con estos virus. Sólo existe un trabajo realizado mediante la técnica BIAcore con un ligando del dominio L del VIH (O. Pornillos *et al.* 2002; Owen Pornillos *et al.* 2002a), que revela un valor de la constante de disociación en torno a 3  $\mu\text{M}$ , lo que indica que las interacciones establecidas por el dominio UEV son del mismo orden que las obtenidas para otros dominios modulares como WW y SH3 (Arold *et al.* 1998; Ferreon & Hilser 2004; Meiyappan *et al.* 2007; Morales *et al.* 2007; Palencia *et al.* 2004; Renzoni *et al.* 1996; Wang *et al.* 2001).

Para los dominios víricos tipo PPxY, presentes en virus como el de Mason-Pfizer en monos, el de la leucemia bovina, o el del sarcoma de Rous, la interacción se establece con la ligasa de ubiquitina Nedd4 (Martin-Serrano *et al.* 2001; Yasuda & Hunter 1998; Yasuda *et al.* 2003). Aunque todavía no se dispone de suficiente información acerca de esta proteína, hoy día parece estar bien establecido que las interacciones entre el dominio L tipo PPxY de algunos virus y Nedd4, concretamente con su dominio WW3, son fundamentales para la correcta liberación de estos virus de la célula infectada (Demirov & Freed 2004; Freed 2002). La proteína Nedd4 pertenece a la familia de las ubiquitina-ligasas que regulan un gran número de funciones celulares (Harvey & Kumar 1999). La proteína, ilustrada esquemáticamente en la Figura 1.16, posee tres tipos de dominios, uno en el extremo N-terminal que tiene la función de reconocer determinados lípidos, denominado dominio C2, otro los dominios WW, numerados del 1 al 4, y el dominio Hect (*Homologous E-6-associated protein C terminus*) que posee actividad ligasa

## Antecedentes



**Figura 1.16.** Representación esquemática de la estructura de la proteína Nedd4, donde se representa el dominio en el extremo C-terminal, denominado C2, que reconoce proteínas lipídicas, los dominios WW1, WW2, WW3 y WW4, encargados del reconocimiento de secuencias ricas en prolina, y el dominio en el extremo N-terminal Hect, que posee actividad ubiquitina ligasa.

para ubiquitina. Una de las funciones principales de Nedd4 es la regulación de la expresión del canal epitelial de sodio (EnaC), proceso que es dependiente de ubiquitina (Harvey *et al.* 1999; Staub *et al.* 1996). Además, el homólogo de Nedd4 en levadura, denominado Rsp5, controla el anclaje de al menos tres proteínas de membrana mediante la regulación de ubiquitina. Otra función importante para Nedd4 parece estar relacionada con el ciclo de vida del virus del sarcoma (RSV), para el que se ha descrito que la sobreexpresión de los dominios WW impide la liberación de los viriones de RSV de la célula infectada. Igualmente, también se ha observado que la sobreexpresión de Nedd4 disminuye la cantidad de partículas virales que se liberan para algunos virus, como es el caso del de la leucemia humana tipo-1 (Sakurai *et al.* 2004). Este virus, que contiene solapados un dominio L tipo PPxY y otro tipo PTAP, parece beneficiarse de las interacciones establecidas por el dominio PPxY para la liberación de sus viriones, mientras que el papel del dominio tipo PTAP no parece estar todavía lo suficientemente claro (Bouamr *et al.* 2003; Sakurai *et al.* 2004). En el caso del virus del ébola, el cual también contiene dos dominios L, uno tipo PPxY y otro PTAP, éste parece haber desarrollado la habilidad de utilizar ambos dominios L para escapar de la célula, bien a través de la interacción entre el dominio WW3 de Nedd4 y PPxY o bien entre el dominio UEV de Tsg101 y PTAP (Harty *et al.* 2000; Licata *et al.* 2003).

Cabe decir que por el momento no se dispone de información estructural del dominio WW3 de la ubiquitina ligasa humana de Nedd4, ni, al igual que Tsg101-UEV, amén de estudios de tipo termodinámico de las interacciones que establece con dominios virales tipo L.

## **MATERIALES Y MÉTODOS**



## **2.1. MATERIALES**

Los ligandos peptídicos para el estudio con la proteína Abl-SH3 fueron proporcionados por el servicio de síntesis de péptidos del EMBL (Laboratorio Europeo de Biología Molecular, Heidelberg, Alemania), a excepción de p40 y p41, los cuales fueron adquiridos de las empresas DiverDrugs (Barcelona, España) y SIGMA-Genosys. Todos los ligandos para el estudio de los dominios L víricos fueron adquiridos de SIGMA-Genosys y SYNBIOSCI. En todos los casos, los péptidos fueron acetilados y amidados en los extremos carbono y nitrógeno terminal, respectivamente. Los péptidos fueron sintetizados en fase sólida en columna MPS y su pureza (> 95%) fue confirmada por HPLC y espectrometría de masas (EM).

Para la preparación de todas las disoluciones se utilizó agua destilada y desionizada mediante una resina de intercambio iónico, y posteriormente desionizada y filtrada de nuevo en un equipo Milli-Q plus (Millipore). Los siguientes productos, empleados para la preparación de las diferentes disoluciones, se obtuvieron de las siguientes compañías:

de MERCK: glicina, fosfato sódico dibásico y monobásico, sulfato amónico, ácido acético, ácido clorhídrico, hidróxido sódico con pureza mínima del 99.5 %, cloruro sódico, IPTG y  $\beta$ -mercaptoetanol;

de SIGMA: Bicina, TRIS, imidazol, citrato sódico, MES, MOPS y HEPES con pureza mínima del 99.5 %, hidróxido sódico, ácido clorhídrico y aceite mineral.

El material de laboratorio para la preparación de disoluciones se adquirió de las siguientes compañías:

de SARSTED, los filtros de tamaño de poro de 0.45  $\mu$ m y 0.2  $\mu$ m, y tubos de plástico de diferentes tamaños;

de SPECTRA/POR las membranas de diálisis de tamaño de poro de 3.5 kDa y de 1 kDa de peso molecular excluido.

Los reactivos y enzimas de biología molecular necesarios para el clonado, expresión y purificación de las proteínas se adquirieron de las siguientes compañías:

de Bonsái Technologies, oligonucleótidos para el clonado de dominios WW3 y WW4 de Nedd4, WW de Yap65 y para la mutagénesis dirigida del dominio Abl-SH3,

## ***Materiales y métodos***

de ROCHE, ditiotretitol (DTT), mezcla de nucleótidos (dNTP mix), Ampicilina, Cloramfenicol, los enzimas de restricción NcoI y HindIII, los reactivos para la ligación de ADN "*Rapid DNA Ligation Kit*", y los reactivos para purificación de plásmidos "*HighPure Isolation plasmid*";

de STRATAGENE se adquirieron las células supercompetentes XL-1 Blue, la Taq y Pfu polimerasas y los reactivos de mutagénesis dirigida "*QuickChange site-directed mutagénesis*";

de MBI FERMENTAS, se adquirió la fosfatasa alcalina SAP;

de NOVAGEN, se adquirieron las células para expresión BL21-DE3 y BL21-DE3 Codon+;

de DIFCO, triptona y extracto de levadura para la preparación de medios de cultivo y placas de sembrado;

de MediTest, tiras para medición de glucosa;

de QUIAGEN, se adquirieron la resina para cromatografía de afinidad de Níquel y los reactivos de purificación de productos de PCR "*Purification Kit PCR Quick*".

Para la cristalización de todas las proteínas y complejos se utilizaron algunos reactivos especiales y material de las siguientes compañías:

de Fluka, el PEG200, PEG400 y PEG4000;

de Hampton Research (<http://hamptonresearch.com>) los reactivos HR2-110, HR2-112, bandejas de 24 y de 48 reservorios, cera para sellado de capilares y herramientas para montaje de cristales para difracción;

de Triana S&T ( <http://www.trianatech.com> ), los reactivos para escaneado de pH's con sulfato amónico y formiato sódico, las cajas para 6 capilares GCB y capilares de 0.1, 0.2 y 0.3 mm.

## **2.2. INSTRUMENTACIÓN**

Los experimentos de ITC se han realizado en un microcalorímetro MicroCal MCS-ITC (MicroCal Inc., Northampton, MA, EEUU), conectado a un baño externo de NESLAB modelo RTE 100, y los de DSC en un calorímetro diferencial de barrido VP-DSC, también de MicroCal. El análisis de los datos de ITC se ha realizado con las

versiones 5.0, 6.1 y 7.0 de Origin (MicroCal Inc.) y las simulaciones para optimizar las concentraciones de proteína y ligando y volúmenes de inyección en los experimentos de ITC con MatLab® 6.5 de MathWorks Inc. El análisis de datos de DSC se ha realizado con la versión 6.1 de Origin.

Los experimentos de fluorescencia se han realizado con un espectrofluorímetro PERKIN ELMER modelo LS55. La termostatación de las muestras y de la cubeta situada en el portacubetas se llevó a cabo recirculando agua desde un baño termostaticado a 25° C de la marca SELECTA modelo Ultraterm 6000383. Como foco frío se dispuso de un dedo frío inmerso en el baño, también de SELECTA.

Los espectros de dicroísmo circular se han realizado en un espectropolarímetro JASCO, conectado a un Peltier para el control de temperatura PTC-348WI. El análisis de los datos de fluorescencia y dicroísmo circular se ha realizado con el programa Origin 6.1.

La difracción de cristales de proteínas se ha realizado en un difractor de rayos-X Bruker Smart6000 con detector tipo CCD y configuración Kappa (X8 Proteum) utilizando radiación  $K\alpha$  de cobre desde un microfoco Bruker MicroStar (Montel optics) y originada por un ánodo rotatorio a 45 kV y 60 mA. Esta instrumentación existe en el Laboratorio de Estudios Cristalográficos (CSIC), del Campus de la Salud de Granada.

Las simulaciones de dinámica molecular se han realizado con el programa AMBER versión 8.0; para el modelado, el programa Insight-II, y para la visualización de trayectorias y estructuras tridimensionales, VMD, versión 1.8.5. Las trayectorias de dinámica molecular se han calculado en el centro de Biotecnología BioTec en Dresden (Alemania). Concretamente, se ha utilizado una estación de trabajo 10 Dual XEON DELL (3 GHz y 4GB de RAM), conectada a una red de 16 ordenadores (4 GB de RAM por cada ordenador), que dispone una capacidad de almacenamiento de 2.5 TB y copia de seguridad de datos diaria. Para el modelado de estructuras se ha utilizado una estación de trabajo SGI FUEL con 2GB de memoria RAM y capacidad de visualización en estéreo.

## ***Materiales y métodos***

Los espectros UV-visible de las distintas muestras de proteína se han obtenido con un espectrofotómetro PERKIN ELMER modelo Lambda 25 de doble haz. Las medidas de índice de refracción para los cálculos de los coeficientes de extinción molar de las proteínas se realizaron con un refractómetro de mano ATAGO modelo R5000.

Las medidas de pH se han realizado con electrodos de vidrio conectados a un pH-metro digital CRISON GLP-21, de sensibilidad 0.01 unidades. La calibración del pH-metro se realizó sistemáticamente con disoluciones estándar de pH 4.01 y 7.00 (a 20°C), también de la marca CRISON.

La clonación de los dominios WW y mutantes de Abl-SH3 se ha realizado en un termociclador TECHNE PROGENE mediante la reacción en cadena de la Polimerasa. La esterilización de medios de cultivo se ha realizado en un autoclave "Autester-D", la incubación de los cultivos celulares en un incubador de Ing. CLIMAS con agitación orbital, y su centrifugación en las centrifugas KOKUSAN H-251, con rotor M1 y Hettich Roto Super 40, con rotor A.69. El lisado de los cultivos celulares se ha realizado en una prensa *French*, de la casa AMINCO y su ultracentrifugación en una centrífuga Optima™ LE-80K de BECKMAN COULTIER, con rotor 45Ti.

Para la purificación de la proteína Abl-SH3 se utilizaron una columna de exclusión molecular Hi-load SuperDex75™, un inyector INV-150, un colector de fracciones Frac-950 y un detector UV UPC-900, todos ellos de la casa de AMERSHAM PHARMACIA BIOTECH AB. Para la purificación de las proteínas Tsg101-UEV, Nedd4-WW3, Nedd4-WW4 y Yap65-WW se utilizó una columna estándar de 5 cm de largo x 10 mm de diámetro, sobre la que se empaquetó la resina de Ni-NTA de Quiagen.

Las electroforesis en geles de poliacrilamida se han realizado con una cubeta vertical Biorad PowerPac Basic, y en geles de agarosa con una cubeta SCIE-PLAS. La concentración de disoluciones de proteína se ha realizado en las células de ultrafiltración de 3 y 50 mL con membranas de 1 KDa para los dominios WW, y de 3.5 KDa para Abl-SH3 y Tsg101-UEV, todo de Millipore.

Para mantener la temperatura constante a 15 °C en los experimentos de cristalización hemos utilizado un enfriador modelo *wine cooler* de BOSCH.

## **2.3. MÉTODOS**

### **2.3.1. Clonado, expresión y purificación del dominio SH3 de Abl (*virus Leucémico de Abelson*)**

El clon del dominio Abl-SH3 fue ofrecido generosamente por el Dr. Luís Serrano (EMBL, Heidelberg, Alemania). Dicho dominio fue reclonado en el plásmido pBAT4 (Peranen *et al.* 1996) por la Dra. Eva Sánchez-Cobos, miembro de nuestro laboratorio, con el propósito de optimizar su expresión. Para la expresión y purificación del dominio Abl-SH3 se siguieron los protocolos descritos previamente en (Viguera *et al.* 1994). El gen del dominio Abl-SH3 que codifica por la secuencia MENDPNLFVA LYDFVASGDN TLSITKGEKL RVLGYNHNGE WCEAQTKNQ GWVPSNYITP VNS fue expresado en células *E.Coli* BL21-DE3 usando IPTG como agente inductor de expresión hasta una concentración final de 0.2 mM. Los cultivos celulares fueron inducidos con IPTG cuando su densidad óptica (Absorbancia a 600 nm) alcanzó el valor de 0.6.

Las células fueron recogidas por centrifugación 2000 x g durante 15 minutos, resuspendidas en Tris 100 mM pH 9.0, lisadas mediante una prensa *French* a una presión de 1000 psi y ultracentrifugadas a 12000 x g durante 30 minutos. El dominio Abl-SH3 fue precipitado del sobrenadante resultante añadiendo sulfato amónico muy lentamente y con agitación hasta una saturación del 75%, y posteriormente redisoluelto en tampón fosfato sódico 10 mM, cloruro sódico 500 mM pH 6.5.

La disolución obtenida, previamente pasada por un filtro de 0.4 micras de diámetro, se sometió a una etapa de cromatografía de exclusión molecular utilizando una resina "Hi-Load Superdex-75", previamente equilibrada en el mismo tampón. La proteína se obtuvo con una pureza superior al 99%, comprobada por electroforesis en gel de poliacrilamida (Laemmli 1970) y espectrometría de masas (Servicio de espectrometría de masas del CIC, Universidad de Granada). Para su

## ***Materiales y métodos***

almacenamiento, la proteína se concentró hasta  $11 \text{ mg}\cdot\text{mL}^{-1}$  en el mismo tampón, se congeló rápidamente en nitrógeno líquido y se mantuvo a  $-20 \text{ }^\circ\text{C}$  hasta su utilización. El rendimiento final aproximado es de 100 mg de proteína por litro de cultivo celular.

### **2.3.2. Obtención de variantes del dominio SH3 de Abl por mutagénesis dirigida**

Para la obtención de variantes del dominio SH3 de Abl hemos utilizado el método de mutagénesis dirigida *Quikchange*<sup>™</sup> de STRATAGENE, que permite realizar mutaciones (puntuales y múltiples) en un plásmido de cadena doble de forma sencilla.

Los oligonucleótidos cebadores se diseñaron de acuerdo con la mutación deseada de forma que fueran complementarios entre sí y con la cadena doble del plásmido, y con longitudes aproximadas de 30 a 40 bases de modo que su temperatura de fusión ( $T_f$ ) oscilará entre  $73$  y  $78 \text{ }^\circ\text{C}$ , ya que elevadas  $T_f$  favorecen altas temperaturas de anillamiento ( $T_a$ ). La  $T_f$  de los oligonucleótidos en este caso se estimó de acuerdo con la siguiente ecuación empírica:

$$T_f = 81.5 + 0.41 \cdot (\%G/C) - 675/N - (\% \text{ bases mutadas en el oligo}) \quad (2.1)$$

donde N es el número total de bases del oligonucleótido y %G/C es el tanto por ciento de Guaninas y Citosinas. Además, se ha intentado que el triplete que codifica la mutación este siempre lo más centrado posible en la secuencia del oligonucleótido y que corresponda con el triplete con mayor propensión en *E. Coli*. La reacción en cadena de la Polimerasa se ha diseñado según las recomendaciones del método *Quikchange*<sup>™</sup> de STRATAGENE, que se ilustra en la Tabla 2.1.

**Tabla 2.1:** Programa de la PCR seleccionado en el termociclador para la clonación de mutantes de *Abl-SH3*.

Programa	Número de ciclos	Temperatura	Duración
1	1	95°C	2 min.
		95°C	30 seg.
2	20	55°C	30 seg.
		72°C	2min./kb plásmido
3	1	4°C	Fin

### **2.3.3. Expresión y purificación del dominio UEV de Tsg101 (Tumor Susceptibility Gene 101)**

El dominio Tsg101-UEV fue clonado en el plásmido pRSETA (Schoepfer 1993) como una proteína de fusión a una cola de histidina en el laboratorio del profesor Winfried Weisserhorn (EMBL, Grenoble, Francia) y ofrecido generosamente a nuestro grupo de investigación. El gen del dominio Tsg101-UEV, que codifica por la secuencia *MRGSHHHHHGMAS MAVSESQLKK MVSKYKYRDL TVRETVNVIT LYKDLKPVLD SYVFNDGSSR ELMNLTGTIP VPYRGNTNIP ICLWLLDTYP YNPPCFVKPT SSMTIKTGKH VDANGKIYLP YLHEWKHPQS DLLGLIQVMI VVFGDEPPVF SRP*, fue expresado en células *E.Coli* BL21-DE3 Codon+ (Stratagene), diseñadas especialmente para secuencias con un alto contenido en tripletes que codifican Arg y Lys. El principio de la secuencia (en cursiva) corresponde con la zona del vector pRSETA que contiene la cola de 6 histidinas, utilizada para la purificación del dominio por afinidad a Níquel.

Tras la transformación celular y una vez que aparecieron las colonias en las placas de sembrado (entre 12 y 16 horas a 37 °C) se seleccionaron dos colonias para la preparación de un preinoculo de 100 mL de medio de cultivo enriquecido RB que se incubó durante 12 horas con agitación a 37 °C. Tanto en las placas de sembrado como en el preinoculo se utilizó una concentración de 50 mg·L<sup>-1</sup> de los antibióticos Ampicilina y Cloranfenicol. Seguidamente, se inocularon 10 mL de preinoculo por cada litro de medio de cultivo (generalmente se utilizaron 10 litros

## ***Materiales y métodos***

repartidos en 12 matraces de litro) RB sin antibióticos que se incubaron a 37 °C con agitación. Una vez que la densidad óptica a 600 nm llegó a 1, se indujo la expresión del dominio Tsg101-UEV mediante la adición de IPTG hasta una concentración final de 1 mM, incubando a 37 °C durante 4 horas. Durante la expresión se controló el nivel de glucosa, que debe ser superior a 1.5 g·L<sup>-1</sup>.

Transcurridas 4 horas desde la inducción, las células fueron recogidas por centrifugación a 7000 x g y 4 °C durante 15 minutos, resuspendidas en tampón fosfato sódico 50 mM, cloruro sódico 300 mM, β-mercaptoetanol (BME) 5 mM pH 8.0 (Tampón de Columna: TC), lisadas mediante una prensa *French* con una presión de 1000 psi y ultracentrifugadas a 12000 x g a 4 °C durante 30 minutos.

El dominio Tsg101-UEV, que se encuentra en el sobrenadante, fue purificado por cromatografía de afinidad de Níquel, utilizando una resina Ni-NTA de Quiagen previamente equilibrada en el tampón TC. Una vez agregado el sobrenadante a la columna por gravedad se llevaron a cabo tres etapas de lavado una con 100 mL del TC, otra con 50 mL de TC más 20 mM de imidazol, y una tercera con 50 mL de TC más 50 mM de imidazol, para ir eliminando progresivamente impurezas unidas a la columna de Níquel mediante interacciones inespecíficas. La proteína se eluyó con 50 mL de TC más 500 mM de imidazol en fracciones de 5 mL. Las fracciones que contenían la proteína pura se identificaron mediante electroforesis en gel de acrilamida, se dializaron extensivamente en tampón glicina 50 mM pH 3.0 y se concentraron hasta 5 mg·mL<sup>-1</sup>. Finalmente, se congelaron de inmediato en nitrógeno líquido y se almacenaron a -80 °C. La pureza, identidad e integridad del dominio se comprobó mediante electroforesis en gel de acrilamida y espectrometría de masas. El rendimiento final aproximado es de 10 mg de proteína por litro de cultivo celular.

### **2.3.4. Clonado, expresión y purificación de dominios WW (WW3 y WW4 de Nedd4 humana y WW de Yap65 humana)**

#### **2.3.4.1. Clonado *de novo* de dominios WW**

Para el clonado de todos los dominios WW utilizados en este trabajo se ha utilizado la misma estrategia. En el caso del dominio Yap65-WW, hemos utilizado la

secuencia correspondiente al mutante L30K, de mayor estabilidad, utilizado para la determinación de la estructura tridimensional en disolución del complejo con el péptido GTPPPYTVG (Pdb: 1jmq) (Macias *et al.* 1996). Las secuencias correspondientes a los dominios WW3 y WW4 de la ubiquitina ligasa humana Nedd4 se obtuvieron a partir de la secuencia aminoacídica de la proteína completa (código P46934) obtenida de la base de datos SwissProt/TrEMBL (<http://www.expasy.org/sprot/>). Los límites de ambos dominios se establecieron utilizando la secuencia del dominio Yap65-WW como patrón. Tras los procesos de clonado, las secuencias finales obtenidas son las siguientes:

**Yap65-WW**

GAMGFEIPDD VPLPAGWEMA KTSSGQRYFK NHIDQTTTWQ DPRKAMLSQ

**Nedd4-WW4**

GAMGDTSNDL GPLPPGWEER THTDGRIFYI NHNIKRTQWE DPRLNVAIT

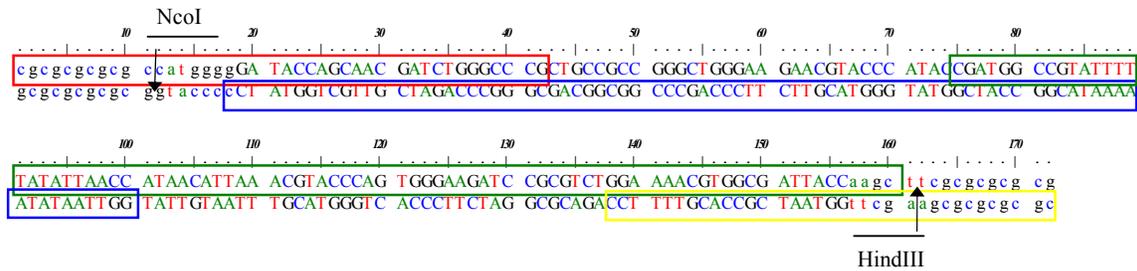
**Nedd4-WW3**

GAMGPSEIEQ GFLPKGWEVR HAPNGRFFI DHNTKTTTW EDPRLKIPAH

Las secuencias de aminoácidos de los tres dominios se tradujeron de acuerdo con los codones de nucleótidos de uso más frecuente en el organismo de expresión *E. coli*, para lo que utilizamos el programa *BioEdit* (Hall 1999). A las secuencias de nucleótidos resultantes (de aproximadamente 140 bases) se le añadieron las secuencias diana correspondientes a los enzimas de restricción NcoI y HindIII en los extremos 5' y 3', respectivamente. Los oligonucleótidos utilizados para el dominio Nedd4-WW3 se muestran como ejemplo en la Figura 2.1. Las secuencias iniciales CGCGCGCGCG son necesarias para que los enzimas de restricción NcoI y HindIII puedan anclar al fragmento de ADN y desarrollar su actividad restrictasa adecuadamente.

Puesto que la metionina inicial (ATG) está incluida en la diana de restricción de NcoI y deja una guanina (G) que desplaza el marco de lectura de los tripletes de pares de bases del dominio WW, se decidió añadir otras dos guaninas a la secuencia de la diana de restricción, de forma que después del triplete de la metionina inicial se leyera el triplete GGG, que *E. coli* traduce como una glicina, aminoácido con el

## Materiales y métodos



**Figura 2.1.** Secuencia de nucleótidos completa de WW4-NEDD4. Los oligonucleótidos 1, 2, 3 y 4, seleccionados para las dos reacciones de la polimerasa en cadena, se muestran en recuadros en color rojo, azul, verde y amarillo, respectivamente. Los sitios de corte para los enzimas de restricción NcoI y HindIII se marcan con flechas.

menor impacto posible en las características del dominio WW que se quiere producir. De este modo, tras el proceso de clonado, expresión y corte con la proteína TEV, el extremo N-terminal posee los aminoácidos GAMG en los tres dominios.

La estrategia utilizada para el clonado de los dominios WW consiste en el diseño de cuatro oligonucleótidos que solapan entre sí y que permitirán obtener la secuencia completa del dominio mediante dos reacciones de PCR consecutivas (Figura 2.1). En todos los casos los oligonucleótidos se diseñaron de acuerdo a los siguientes criterios: 1) longitud total menor de 80 bases; 2) la posición y longitud de las regiones de hibridación entre oligonucleótidos se determinó de forma que su temperatura de desnaturalización fuera similar en todos los casos y cercana a la temperatura óptima de hibridación (72 °C). La temperatura de fusión teórica se calculó por medio de la relación empírica que tiene en cuenta los puentes de hidrógeno de las bases que solapan ( $T_f = [2 \cdot (A+T)] + [4 \cdot (G+C)]$  (° C)); 3) Siempre que fue posible los extremos de los oligonucleótidos se acotaron con Guaninas y Citosinas, por el motivo descrito anteriormente.

En la primera PCR, detallada en la Tabla 2.2, de 30 ciclos, se amplificó la secuencia del dominio codificada por los dos oligonucleótidos centrales, denominados 2 y 3. El producto de esta primera PCR se purificó y se empleó como ADN molde en la segunda PCR, con la misma secuencia de programas que la anterior, en la que se extendió la secuencia central del dominio haciendo uso de los oligonucleótidos 1 y 4.

Tras 30 ciclos de extensión se purificaron los productos de reacción y se comprobó mediante secuenciación que correspondían a las secuencias de ADN esperadas. Una vez obtenido el ADN correspondiente a cada uno de los dominios, se clonó en el vector de expresión para *E. coli* pETM30 (G. Stier, EMBL, Heidelberg, Alemania), que confiere a las bacterias resistencia al antibiótico kanamicina e incluye dianas de restricción para las enzimas NcoI y HindIII. Este vector permite expresar el dominio WW fusionado por su extremo N-terminal a la secuencia diana de siete aminoácidos para la proteasa TEV, a la Glutathion S-Transferasa (GST) y a una cola de histidinas. Para ello, se llevó a cabo una etapa de digestión del ADN con las enzimas NcoI y Hind III, tanto del dominio WW como del vector pETM30, utilizando los tampones recomendados por ROCHE, y que consistió en una etapa de incubación durante una hora a 37 °C y posteriormente durante 15 minutos a 65 °C para lograr la desactivación de los enzimas. Los productos digeridos se purificaron con los reactivos de purificación de productos de PCR, y los extremos del vector pETM30 que se generaron después de la digestión se desfosforilaron, usando la fosfatasa alcalina SAP, para evitar su recircularización.

Posteriormente, se realizó la ligación o unión del ADN del dominio al vector mediante los reactivos "*Rapid DNA Ligation Kit*" de ROCHE y finalmente se purificó el vector conteniendo el ADN de interés siguiendo el protocolo descrito para los reactivos "*High Pure Plasmid Isolation Kit*" de QUIAGEN.

Todas las etapas de manipulación de ADN se siguieron mediante electroforesis en gel de agarosa al 2%. Finalmente, las secuencias del vector conteniendo el ADN de cada dominio WW fueron secuenciadas por el Servicio de Secuenciación del Instituto de Parasitología y Biomedicina "López-Neyra" (CSIC, Granada, España).

El ADN plasmídico purificado se introdujo en células competentes comerciales XL-1 Blue (Stratagene). Las células bacterianas conteniendo el plásmido se hicieron crecer en placas Petri de LB en presencia de kanamicina a 37 °C durante unas 12 horas, hasta observar la aparición de colonias resistentes al antibiótico empleado, entre las que se escogieron 4-5 que se hicieron crecer en medio de cultivo LB. El plásmido de ADN recombinante se purificó a continuación y se introdujo en células competentes BL21-DE3.

## **Materiales y métodos**

*Tabla 2.2. Programa de la PCR seleccionado en el termociclador para la clonación de dominios WW.*

<b>Programa</b>	<b>Número de ciclos</b>	<b>Temperatura</b>	<b>Duración</b>
<b>1</b>	1	95°C	1 min.
		95°C	1.min
<b>2</b>	30	55°C	1 min
		72°C	1 min
<b>3</b>	1	72°C	10 min
<b>4</b>	1	4°C	FIN

En este caso, con cada una de las colonias seleccionadas de la placa se llevaron a cabo ensayos de expresión, mediante la preparación de minicultivos de 10 mL en medio LB en presencia de kanamicina, en los que se varió el momento de adición de agente inductor IPTG, hasta una concentración 1mM. La expresión de cada colonia se controló mediante la intensidad de la banda correspondiente a la proteína de interés en un gel de acrilamida. El ADN plasmídico contenido en las colonias de mayor expresión fue purificado finalmente mediante los reactivos “*High Pure Isolation Plasmid*” de QUIAGEN y secuenciado nuevamente, para comprobar que la secuencia de bases que codifica el gen de la proteína era correcta.

### **2.3.4.2. Expresión y purificación de dominios WW**

La expresión de los dominios WW se llevó a cabo en células BL21-DE3 (Novagen) que se incubaron 12 horas a 37 °C en 5 L de medio LB en presencia de IPTG a 1mM. Las células se recogieron mediante centrifugación a 7000 x g y 4 °C durante 15 minutos y el precipitado celular se resuspendió en el tampón de columna TC (50 mM fosfato sódico, 300 mM cloruro sódico pH 8.0) conteniendo un cóctel de inhibidores de proteasas. Posteriormente se lisaron las células mediante una prensa *French* a 1000 psi. y, después de ultracentrifugar a 12000 x g durante 30 min a 4 °C,

se llevó a cabo una primera etapa de cromatografía de afinidad a níquel con la fracción soluble, que contiene la proteína de fusión GST-WW, siguiendo un procedimiento similar al descrito en la purificación del dominio Tsg101-UEV. La proteína de fusión eluída, se dializó utilizando membranas de tamaño de poro 3.5 kDa frente al tampón recomendado para la hidrólisis con la proteasa TEV (50 mM Tris-HCl, 0.5 mM EDTA, 1 mM DTT pH 8.0) y se incubó durante 1 hora a temperatura ambiente con 1  $\mu$ L de proteasa TEV por mL de disolución de proteína para hidrolizar. Los productos de la digestión se dializaron de nuevo frente al tampón TC esta vez utilizando membranas de tamaño de poro de 1 KDa y se sometieron a una segunda etapa de cromatografía de afinidad a níquel en la que la proteína de fusión que haya quedado sin hidrolizar, la GST y la proteasa TEV que contienen colas de histidinas quedan retenidas por su afinidad al níquel. El dominio WW se eluye puro generalmente en el primer lavado con el tampón TC o, en el caso del dominio Nedd4-WW3 en el segundo lavado con tampón TC más 20 mM imidazol. Las fracciones con proteína pura se dializaron en tampón TC para quitar el imidazol y se concentraron por ultrafiltración (con membranas de tamaño de poro de 1 KDa) hasta una concentración de 2 mg·mL<sup>-1</sup>, se congelaron rápidamente en nitrógeno líquido y se guardaron a -80° C. En estas condiciones la proteína es estable durante varios meses.

Las diferentes etapas de la purificación se controlaron mediante electroforesis en geles de acrilamida. La pureza de cada proteína se confirmó mediante espectrometría de masas (CIC, Universidad de Granada) y fue superior al 99 %. Los pesos moleculares obtenidos, coincidentes con los teóricos, fueron de 5568.2, 5720.2 y 5539.2 Da para Nedd4-WW3, Nedd4-WW4 y Yap65-WW, respectivamente. El rendimiento final aproximado fue de 10 mg por litro de cultivo.

### **2.2.5. Preparación de disoluciones tampón**

Las disoluciones tampón se prepararon pesando las formas ácida y básica del sistema tamponante, calculadas mediante la ecuación de Henderson Hasselbach:

$$pH = pK_a - \text{Log} \left[ \frac{\text{forma}\cdot\text{ácida}}{\text{forma}\cdot\text{básica}} \right] \quad (2.2)$$

## ***Materiales y métodos***

A continuación, se disolvieron en el 80 % del volumen final de disolución, se ajustó el pH añadiendo hidróxido sódico o ácido clorhídrico hasta el valor requerido, se enrasó al volumen final de disolución y finalmente se filtró a través de un filtro de 0.45  $\mu\text{m}$  de tamaño de poro. Alternativamente, el tampón puede prepararse 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, HEPES, MOPS e Imidazol.

### **2.2.6. Diálisis de proteínas**

Tras descongelar las alícuotas de proteína, se introdujeron en una membrana de diálisis de tamaño de poro adecuado (3.5 kDa para Abl-SH3 y Tsg101-UEV, y 1 kDa para los dominios WW). La membrana se cerró por ambos extremos y se introdujo entonces en un vaso de precipitado que contenía un volumen de tampón unas 500 o 600 veces mayor que el de la muestra; se agitó suavemente durante al menos el tiempo mínimo necesario para alcanzar el equilibrio de la diálisis (unas 8 horas en el caso de las membranas utilizadas) a 4 °C. Transcurrido ese tiempo se realizó un cambio del tampón que contiene la muestra por tampón fresco, dejándose entre 8 y 12 horas aproximadamente en las mismas condiciones de agitación y temperatura. A continuación se extrajo la disolución dializada de proteína de la bolsa de diálisis y se clarificó en una centrifuga de mesa a máxima velocidad.

### **2.2.7. Determinación de las concentraciones de proteína y ligando**

La determinación de la concentración se llevó a cabo midiendo la absorbancia de la disolución de proteína a 280 nm y aplicando la ley de Lambert-Beer. En primer lugar se midió la absorbancia del tampón en el que se encuentra la proteína y después se sustrae de la medida de la concentración de proteína o ligando. El error estándar asociado a la concentración de proteína es del 1 al 3% si el coeficiente de extinción está determinado correctamente. Su determinación se ha descrito en algunos trabajos de distintas formas (Martinez 1995), aunque el método más exacto y el que hemos usado en nuestro trabajo, es el de Gill & von Hippel (Gill & von

Hippel 1989). El método consiste en calcular en primer lugar el coeficiente de extinción de la proteína desnaturalizada en guanidina 6 M a partir del número de restos triptófano, tirosina y cisteína que se encuentran en cada molécula. Estos restos mencionados son los únicos que contribuyen de manera significativa a la absorbancia medida para la proteína, que presentará un máximo en el intervalo comprendido entre 276 y 282 nm (Edelhoch 1967).

Para la proteína Abl-SH3 se obtuvo un coeficiente de extinción molar a 280 nm de  $\epsilon = 16894 \text{ M}^{-1} \text{ cm}^{-1}$ , para la proteína Tsg101-UEV  $24180 \text{ M}^{-1} \text{ cm}^{-1}$ , y para Nedd4-WW3, Nedd4-WW4 y Yap65-WW de 11380, 12660 y  $12600 \text{ M}^{-1} \text{ cm}^{-1}$ , respectivamente. La concentración de los péptidos que no contienen restos triptófano se determinó mediante la absorbancia a 276 nm para los péptidos que contienen tirosinas, o bien a 259.5 nm para los que contienen sólo fenilalanina considerando coeficientes de extinción de  $1450 \cdot N_{\text{Tyr}}^{\circ} \text{ M}^{-1} \text{ cm}^{-1}$  y  $195 \cdot N_{\text{Phe}}^{\circ} \text{ M}^{-1} \cdot \text{cm}^{-1}$  respectivamente. Para el péptido 3BP1, que no contiene ninguno de los cromóforos anteriores, se preparó una disolución madre cuya concentración se determinó mediante un experimento de titulación a Abl-SH3.

## **2.2.8. Espectroscopía de Fluorescencia**

### **2.2.8.1. Introducción**

Las proteínas son capaces de absorber y emitir radiación en la zona ultravioleta del espectro electromagnético. Dicha absorbancia es debida a la presencia de los enlaces peptídicos, aminoácidos aromáticos y, en menor extensión, a la existencia de puentes disulfuro. Los tres restos aromáticos Phe, Tyr y Trp son los que contribuyen a la emisión de fluorescencia en la región ultravioleta-visible del espectro (Konev 1967) (Weinryb 1971) (Creighton 1989; Demchenko 1981; Permyakov 1993). Estos aminoácidos son relativamente poco abundantes en proteínas, lo que facilita la interpretación de los datos espectrales desde un punto de vista estructural.

## ***Materiales y métodos***

La emisión de fluorescencia de las proteínas está dominada por el triptófano, que es capaz de absorber a mayor longitud de onda que los otros dos fluoróforos y presenta un coeficiente de extinción mayor, por lo que su fluorescencia es mucho más intensa. Este hecho provoca que frecuentemente la energía absorbida por la fenilalanina y tirosina se transfiera a los restos triptófano de la misma proteína. La fenilalanina presenta las longitudes de onda de emisión y absorción más bajas, tiene una emisión bastante homogénea con un máximo cercano a 282 nm. La emisión de la tirosina en agua tiene su máximo a 303 nm y es relativamente poco sensible a la polaridad del disolvente.

El máximo de emisión del triptófano en agua se encuentra próximo a 350 nm y es muy dependiente de la polaridad y del entorno local. La alta sensibilidad del triptófano a los cambios en su entorno permite observar cambios en su espectro de emisión debido a transiciones conformacionales de la proteína, fenómenos de asociación, procesos de unión de ligandos y a la propia desnaturalización de la proteína.

### **2.2.8.2. Obtención de espectros de fluorescencia**

Como ya hemos indicado antes la fluorescencia de proteínas se puede aplicar al estudio de procesos de unión de ligandos a las proteínas. En algunas ocasiones dichos ligandos poseen fluorescencia propia, y puede cambiar cuando se unen a una proteína {Daniel, 1966 #110} {Anderson, 1969 #111} {Condie, 1983 #112}. La unión de ligandos a proteínas también puede inducir 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 la constante de asociación de un determinado proceso.

Basándonos en este fenómeno hemos realizado una serie de experimentos de titulación por adición de volúmenes variables de disoluciones de los péptidos M1 (Pires *et al.* 2001), leucemia humana tipo I, ébola y PPPY que se describirán en la sección 3.3, sobre disoluciones de las proteínas Nedd4-WW3 y Yap65-WW. Nuestro

objetivo es comprobar que el ligando interacciona con estas proteínas y estimar un valor de la constante de asociación para, en el caso de ser posible, diseñar adecuadamente los experimentos correspondientes de calorimetría isotérmica de titulación. En todos los casos, se partió de una disolución madre de proteína, dializada frente al tampón fosfato sódico 40 mM pH 7.0, de una concentración en torno a 50  $\mu\text{M}$ . Por otro lado se preparó una disolución madre de péptido a una concentración de 5  $\text{mg}\cdot\text{mL}^{-1}$  en agua. A partir de estas disoluciones se prepararon, añadiendo las cantidades necesarias de agua y tampón de diálisis, dos disoluciones con concentraciones finales de 20  $\mu\text{M}$  de proteína y 1.5-3 mM de ligando, ambos en fosfato sódico 20 mM pH 7.0.

El espectro de la proteína libre se obtuvo utilizando una cubeta de cuarzo para fluorescencia de 1 cm de paso de luz, de pequeño volumen de muestra y termostalizada a 25 °C, en la que se introdujeron 120  $\mu\text{L}$  de la disolución de proteína. A este volumen se adicionaron progresivamente volúmenes crecientes (entre 0.5 y 10  $\mu\text{L}$ ) de la disolución de ligando y se registraron los espectros de emisión de fluorescencia tras cada adición. Puesto que los ligandos contienen restos de Tirosina en su secuencia se recogieron los espectros de fluorescencia para las mismas adiciones de ligando sobre una disolución de tampón que se utilizaron como blanco para el análisis. Todos los espectros se registraron a 25 °C entre 300 y 500 nm utilizando una longitud de onda de excitación de 280 nm, y rendijas de excitación y emisión de 7 nm.

### **2.2.8.3. Análisis de los espectros de fluorescencia**

Una vez sustraída la fluorescencia intrínseca del ligando se normalizó cada espectro por la concentración de proteína total en la cubeta, teniendo en cuenta la dilución de ésta tras cada adición de ligando. Para el análisis de datos hemos utilizado tanto la intensidad a 340 nm que corresponde con el máximo del espectro de emisión de la proteína libre como el área bajo la curva del mismo espectro. Si representamos cualquiera de éstas frente a la concentración de ligando total en célula obtenemos la curva característica de unión. El ajuste por regresión no lineal

## **Materiales y métodos**

de mínimos cuadrados de los puntos que definen la curva al modelo de unión sencillo para un ligando que une a una proteína con estequiometría 1:1, a partir del que obtenemos la ecuación general 2.3, permite obtener el valor de la constante disociación:

$$F_{i,n} = \frac{F_i}{M_T} = F_M + (F_{ML} - F_M) \cdot \theta_i = F_M + \Delta F \cdot \theta_i \quad (2.3)$$

donde  $F_{i,n}$  se refiere a la intensidad de fluorescencia corregida por la fluorescencia debida la ligando y normalizada por la concentración total de proteína (o al área, según el análisis que empleemos);  $F_M$  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:

$$\theta_i = \frac{(K_d + M_T + L_T) \pm \sqrt{(K_d + M_T + L_T)^2 - 4 \cdot M_T \cdot L_T}}{2 \cdot M_T} \quad (2.4)$$

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

### **2.2.9. Calorimetría isotérmica de titulación**

#### **2.2.9.1. Introducción**

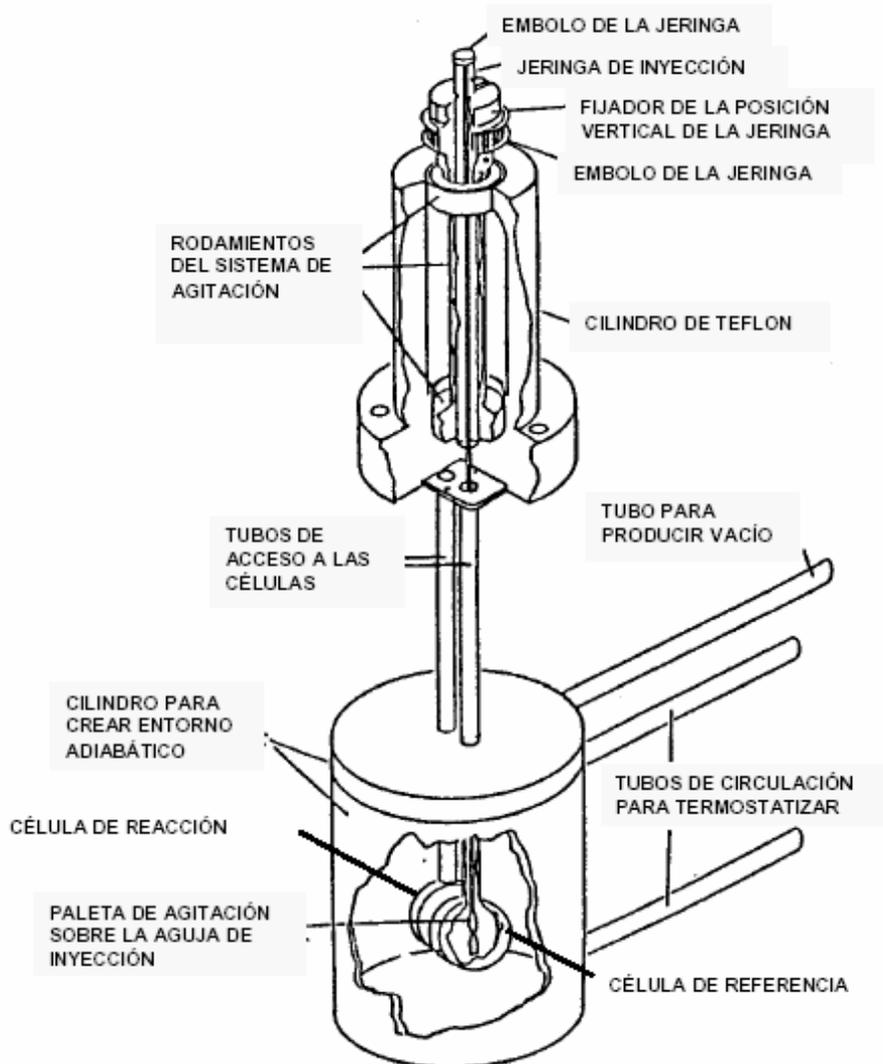
La calorimetría isotérmica de titulación (ITC, según la abreviatura en inglés, que utilizaremos a lo largo del texto) es una técnica que mide directamente el calor asociado a un determinado evento, el cual, a presión constante, equivale a la entalpía de este proceso. Debido a la sensibilidad y precisión requeridas para caracterizar las interacciones proteína-proteína o proteína-ligando, los calorímetros más utilizados pertenecen a la categoría de calorímetros de titulación con sistema de compensación dinámica y que trabajan a presión constante. Una descripción detallada del instrumento puede encontrarse en numerosos trabajos encontrados en la literatura (Breslauer *et al.* 1992; McKinnon *et al.* 1984; Velazquez-Campoy *et al.*

2004; Wiseman *et al.* 1989). Históricamente, esta técnica entró en una época dorada en los años 90 debido a los importantes avances tecnológicos y metodológicos en el campo de la microelectrónica, lo que propició que se diseñaran calorímetros con mucha más sensibilidad (con límite de detección en torno a 1  $\mu\text{cal}$ ) (McKinnon *et al.* 1984; Wiseman *et al.* 1989).

Para nuestro estudio hemos utilizado un calorímetro de la marca MicroCal, modelo MCS-ITC (MicroCal Inc., Northampton, MA, EEUU). En la Figura 2.2 se muestra un esquema de la configuración más común para este calorímetro, donde aparecen, entre otros elementos, las células de reacción y de referencia, la jeringa, el sistema de agitación y el cilindro para crear el entorno adiabático. Una de las disoluciones, bien de macromolécula o de ligando (o en su caso otra macromolécula) se ha de colocar en la célula de reacción, y la otra en la jeringa de inyección, mientras que la célula de referencia suele rellenarse con agua. Tanto la célula de reacción como la de referencia se encuentran suspendidas mediante largos tubos de acero con el fin de crear un entorno lo más adiabático posible y así minimizar, en lo posible, el intercambio de calor con el exterior.

Además, el cilindro adiabático está conectado mediante tubos de circulación de agua con un baño térmico que se utiliza para ayudar a controlar la temperatura. Para eliminar la condensación y mejorar la adiabaticidad se utiliza también un tubo para realizar vacío sobre el cilindro adiabático. El sistema de agitación-inyección es una parte muy importante del calorímetro, ya que es el responsable de la precisión en los volúmenes de inyección así como una homogenización eficaz de la disolución. Así, las jeringas de inyección vienen provistas de largas agujas de acero con su extremo terminal en forma de paleta para facilitar la agitación. El embolo de la jeringa está conectado a un motor de alta precisión, necesario para obtener volúmenes de inyección muy pequeños y precisos (1  $\mu\text{L}$ ). Todas las variables experimentales que dependen del instrumento, como la temperatura, número de inyecciones, volumen por inyección, velocidad de agitación, etc., se controlan automáticamente en un ordenador con el programa MicroCal Origin for ITC versión 4.1.

Durante el experimento, un sistema de retroalimentación genera una pequeña



**Figura 2.2.** Esquema del calorímetro isotérmico de titulación modelo MCS-ITC de MicroCal, donde se representan, entre otros, las células calorimétricas, cilindro adiabático y el conjunto agitador/inyector del calorímetro.

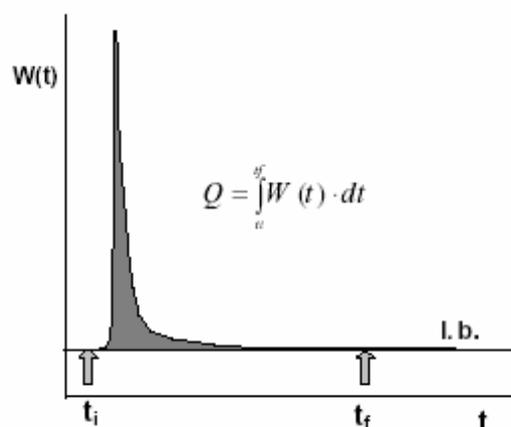
potencia térmica (indicada como trabajo,  $\mu\text{J}\cdot\text{s}^{-1}$ , en el instrumento), necesaria para mantener la misma temperatura en las células de referencia y de reacción. Cualquier proceso que tenga lugar en la célula de reacción y que genere calor, cambiará la temperatura de dicha célula y el sistema de retroalimentación generará la potencia necesaria para igualar la temperatura de las dos células. Esta potencia suministrada por el equipo tiene una forma de pico muy característica, que se representa en la Figura 2.3. En el caso de un proceso de asociación, el calor

asociado con cada inyección es proporcional al aumento de la concentración del complejo formado (avance de la reacción) y se calcula integrando el área bajo la curva de la señal medida (cantidad de calor proporcionada para mantener igual la temperatura de ambas células por unidad de tiempo). Una vez terminado el experimento podemos construir, como veremos más detalladamente en la sección 2.2.9.3, la curva isoterma de unión. El ajuste de esta curva al modelo de unión adecuado permite obtener los valores de las constantes de equilibrio de asociación, amén del resto de parámetros termodinámicos que caracterizan el proceso. Otras técnicas como las espectroscópicas, en las cuales la señal medida también es proporcional al avance de la reacción (por ejemplo, absorbancia, intensidad de fluorescencia, elipticidad molar o desplazamiento químico), permiten obtener sólo el valor de la constante de asociación, y es necesario llevar a cabo varios experimentos a diferentes temperaturas para estimar la entalpía de unión, la cual es determinada indirectamente a través de expresiones de las ecuaciones de van't Hoff (Horn *et al.* 2001):

$$\frac{\partial \ln K_a}{\partial \left(\frac{1}{T}\right)} = -\frac{\Delta H_{vH}}{R} \quad (2.5)$$

$$\frac{\partial^2 \ln K_a}{\partial \left(\frac{1}{T}\right)^2} = -\frac{T^2 \Delta C_p}{R} \quad (2.6)$$

Esto supone dos inconvenientes, primero, debido a razones prácticas, no siempre es posible cubrir un intervalo amplio de temperatura para estimar la constante de asociación, y segundo, aunque entropía y entalpía varían mucho con la temperatura, la energía libre de Gibbs muestra una dependencia más débil debido a la compensación entálpica/entrópica (Velazquez Campoy & Freire 2005). Estas dificultades hacen que haya discrepancias entre la entalpía medida calorimétricamente y la medida espectroscópicamente, denominada entalpía de van't Hoff ( $\Delta H_{vH}$ ), las cuales se han explicado en base a errores e imprecisiones inherentes a los métodos no calorimétricos (Chaires 1997; Horn *et al.* 2001; Liu & Sturtevant 1995; Naghibi *et al.* 1995).



**Figura 2.3.** Pico característico de respuesta a una inyección de un reactivo (ligando) sobre otro que llena la célula calorimétrica (proteína). El calor generado en la reacción se obtiene integrando la función  $W(t)$  sobre la línea base, l. b.

La calorimetría de titulación es una técnica única en el sentido de que es posible obtener una caracterización completa del proceso de unión, es decir información acerca de los componentes de la energía libre de Gibbs (entalpía y entropía) mediante expresiones termodinámicas sencillas:

$$\begin{aligned} \Delta G &= -RT \ln K \\ &= \Delta H - T\Delta S \end{aligned} \quad (2.7)$$

donde  $\Delta G$  es la energía libre de Gibbs,  $\Delta S$  es el cambio de entropía,  $R$  es la constante de gases y  $T$  es la temperatura. Pero la entalpía y entropía son magnitudes que dependen de la temperatura:

$$\Delta C_p = \left( \frac{\partial \Delta H}{\partial T} \right)_p = T \left( \frac{\partial \Delta S}{\partial T} \right)_p \quad (2.8)$$

donde  $\Delta C_p$  es el cambio de capacidad calorífica asociado al proceso de unión, que se puede obtener realizando experimentos a diferentes temperaturas.

La ventaja de que el calor sea la señal que medimos en un experimento de ITC reside en que no es necesario el marcaje de las moléculas para seguir el avance

de la reacción, lo que hace posible el uso de ligandos naturales sin modificar. Sin embargo, hay que ser muy cuidadoso en la preparación de muestras, desarrollo del experimento, análisis de datos y discusión de los resultados, ya que cualquier proceso adicional que genere calor contribuirá al valor de los parámetros que se obtienen del experimento (Velazquez Campoy & Freire 2005)

### **2.2.9.2. Diseño del experimento de ITC**

La estrategia que hemos utilizado para diseñar los experimentos consiste en realizar algún experimento preeliminar utilizando la información descrita en bibliografía, en caso de que exista, y a continuación utilizar una plantilla de simulación en el programa MATHCAD, disponible en la sección de docencia de la web: <http://www.ugr.es/~qmfisica/>, con objeto de optimizar las concentraciones de ligando y proteína necesarias para tener un experimento de ITC adecuado. Un parámetro muy útil es el producto  $K [M]_T$  cuyo valor debe estar entre 1 y 100, aunque el valor óptimo está entre 10 y 100, el cual es determinante para optimizar la forma de la curva isoterma de unión, que debe ser lo más sigmoideal posible (Wiseman *et al.* 1989). Dicha curva se obtiene representado el calor por mol de ligando añadido tras cada inyección frente al cociente concentración de ligando total/concentración de macromolécula total (considerando en este caso que el ligando está en la jeringa de inyección y la macromolécula en la célula de reacción). Otros parámetros utilizados para la optimización son el volumen de inyección y la concentración de ligando, los cuales, junto con la optimización del producto anterior, deben conseguir que se alcance una saturación máxima (es recomendable por encima del 80%), que indicaría que casi la totalidad de proteína que contiene la célula calorimétrica se encuentra unida al ligando (Jelesarov & Bosshard 1999; Velazquez-Campoy & Freire 2006; Velazquez-Campoy *et al.* 2004).

En la Figura 2.4 se muestran las simulaciones de un experimento convencional de ITC consistente en 20 inyecciones de 5  $\mu$ L de una disolución de ligando 5 mM sobre un volumen  $V_C = 1.347$  mL, que corresponde al volumen de la célula de reacción, de una disolución de macromolécula  $1.8 \times 10^{-4}$  M con un único sitio de unión para el ligando. Hemos considerado cuatro posibles valores de la

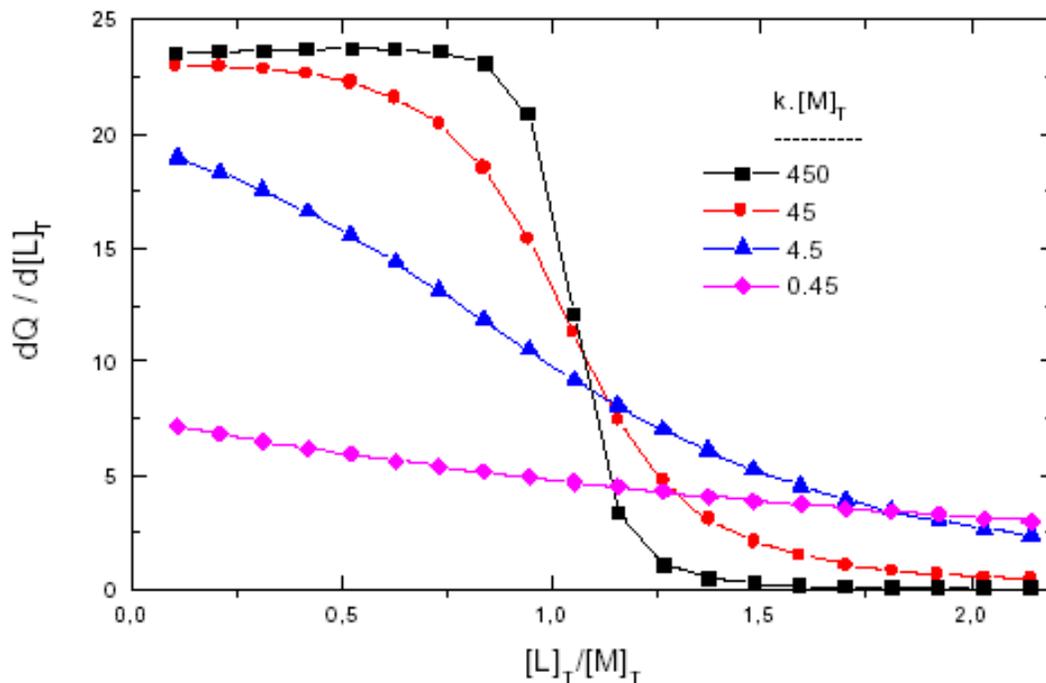
## ***Materiales y métodos***

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 relativa alta hasta una afinidad baja. Como puede observarse cuando el producto  $K [M]_T$  está dentro del intervalo adecuado (10-100) es fácil obtener una isoterma de unión de forma sigmoideal, ideal para el diseño del experimento y para que el ajuste por regresión no lineal correspondiente nos proporcione parámetros con errores estándar aceptables. Cuando el producto  $K [M]_T$  no es tan favorable, bien por tener una constante muy pequeña o muy alta, o bien porque la macromolécula sea poco soluble, no es fácil obtener la isoterma de unión completa manteniendo constante el volumen de cada inyección, proporcionando parámetros con errores significativos. Por esta razón, en nuestro grupo de investigación se ha diseñado una plantilla de simulación para el establecimiento de un perfil óptimo de volúmenes de inyección, que está disponible en la sección de docencia de la web: <http://www.ugr.es/~qmfisica/>. Otro beneficio de usar el perfil óptimo de volúmenes de inyección es la mejora de la relación señal/ruido en la parte final de la curva, donde los calores debidos a la unión del ligando a la proteína son muy pequeños.

Generalmente, en un estudio termodinámico mediante ITC se deben obtener dos termogramas; el primero corresponde a la titulación de la macromolécula con el ligando; y un segundo de titulación del ligando con el tampón de la diálisis. Este último, que representa los calores de la dilución del ligando y los posibles artefactos de inyección debe substraerse del termograma principal para obtener los efectos térmicos netos de la interacción macromolécula-ligando. Además, sería aconsejable obtener un tercer termograma correspondiente a la titulación de la proteína con el tampón de diálisis, que representaría los calores de dilución de la proteína, aunque estos calores suelen ser despreciables.

### **2.2.9.3. Formulación y análisis de un experimento directo de ITC**

En aquellos casos en que ni se muestra un efecto cooperativo ni la proteína presenta diferentes clases de sitios de unión para el ligando, el comportamiento del sistema interaccionante se describe adecuadamente a través del modelo de unión de un ligando a una macromolécula con “n” sitios idénticos e independientes. El término “idénticos” implica la misma constante de unión microscópica a cualquiera



**Figura 2.4.** Simulación del calor por mol de ligando añadido asociado a cada inyección para un experimento ITC con los siguientes parámetros experimentales:  $V_c=1.347$  mL;  $[M]$  inicial en célula de  $1.8 \cdot 10^{-4}$  M; 20 inyecciones de  $5 \mu\text{L}$ ;  $[L]$  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,  $2.5 \cdot 10^6$ ,  $2.5 \cdot 10^5$ ,  $2.5 \cdot 10^4$  y  $2.5 \cdot 10^3 \text{ M}^{-1}$ .

de los “n” sitios, y el término “independientes” que la unión a un sitio cualquiera no modifica la afinidad de los restantes. El análisis termodinámico estadístico de un proceso de unión con estas características nos demuestra que la función de partición de unión ( $Z$ ), que describe las poblaciones de los diferentes estados moleculares accesibles del sistema, es:

$$Z = (1 + K[L])^n \quad (2.9)$$

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. La razón de la concentración de ligando unido,  $[L]_b$ , a la concentración de

## **Materiales y métodos**

macromolécula total,  $[M]_T$ , conocida como parámetro de unión ( $\bar{\nu}$ ) está relacionada con la función de partición de unión por:

$$\bar{\nu} = \frac{[L]_b}{[M]_T} = \frac{\partial \ln Z}{\partial \ln [L]} \quad (2.10)$$

$$\bar{\nu} = \frac{nK[L]}{1 + K[L]} \quad (2.11)$$

donde  $[L]$  representa la concentración de ligando libre en equilibrio con las otras especies moleculares que interaccionan en dicho sistema (M, ML, ML<sub>2</sub>,..., ML<sub>n</sub>).

Un experimento convencional de ITC, está constituido por una serie de inyecciones de un determinado volumen de disolución del ligando, constante o variable para cada inyección, separadas cada una por un intervalo de tiempo suficiente para asegurar que se ha alcanzado el equilibrio y se ha transferido o compensado todo el calor liberado o absorbido. Suponiendo en principio que no existen otros procesos concomitantes, tales como ionización del tampón o cambios conformacionales (Baker & Murphy 1996; Gomez & Freire 1995; Jelesarov & Bosshard 1999), en una inyección cualquiera “i” de la serie, el calor liberado o absorbido será:

$$q_i = \Delta H_{ap} \cdot \left(\frac{kJ}{molL_b}\right) \cdot \Delta(molesL_b) \quad (2.12)$$

donde  $L_b$  representa el ligando unido y  $\Delta H_{ap}$  es el cambio de entalpía aparente por mol de ligando unido. Si  $V_C$  representa el volumen de la célula, que coincide con el volumen inicial de la muestra, la ecuación anterior puede expresarse más explícitamente como:

$$q_i = \Delta H_b \cdot V_C \cdot ([L]_{b,i} - [L]_{b,i-1}) = \Delta H_{ap} \cdot V_C \cdot (\bar{\nu}_i \cdot [M]_i - \bar{\nu}_{i-1} \cdot [M]_{i-1}) \quad (2.13)$$

donde  $[L]_b$  representa la concentración de ligando unido y  $[M]_i$ , la concentración de proteína en la célula. Si consideramos que el volumen efectivo que interviene en la reacción es aquel que llena la célula entonces las concentraciones de macromolécula y ligando actuales en la célula han de corregirse por el volumen desplazado fuera de la célula tras cada inyección. Así la concentración actual de M en la célula ( $[M]_i = [M] + [ML] + [ML_2] + \dots$ ) será:

$$[M]_i = [M]_{i-1} \cdot \frac{V_C - V_{in}}{V_C} \quad (2.14)$$

y la concentración total del ligando en la célula de reacción:

$$[L]_{T,i} = \frac{(V_C - V_{in}) \cdot [L]_{T,i-1} + V_{in} \cdot [L]_0}{V_C} \quad (2.15)$$

donde  $V_{in}$  y  $[L]_0$  son el volumen de inyección y la concentración en jeringa del ligando. El calor total acumulado después de N inyecciones será:

$$Q = \sum_{i=1}^N q_i = V_C \cdot [M]_N \cdot \Delta H_{ap} \cdot \bar{v}_N = V_C \cdot [M]_N \cdot \Delta H_{ap} \cdot \frac{nK[L]}{1 + K[L]} \quad (2.16)$$

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:

$$[L] = [L]_T - [L]_b = [L]_T - \frac{Q}{V_C \cdot \Delta H_{ap}} \quad (2.17)$$

Sustituyendo  $[L]$  en la ecuación 2.16 por la expresión 2.17 se obtiene una ecuación de segundo grado en Q cuya solución es:

$$Q = \frac{V_C \cdot \Delta H_{ap}}{2K} \cdot [1 + K[L]_T + nK[M]_T - \sqrt{(1 + K[L]_T + nK[M]_T)^2 - 4nK^2[M]_T[L]_T}] \quad (2.18)$$

## **Materiales y métodos**

Esta ecuación relaciona el calor total acumulado después de N inyecciones con las variables experimentales  $V_C$ ,  $[L]_T$  y  $[M]_T$  y los parámetros  $n$ ,  $K$  y  $\Delta H_{ap}$  a determinar en el análisis de los datos de un experimento de ITC convencional. Por otro lado la ecuación 2.17 puede reescribirse como:

$$[L] = [L]_T - [L]_b = [L]_T - [M]_T \cdot \frac{nK[L]}{1 + K[L]} \quad (2.19)$$

y despejando  $[L]$  podemos expresar esta concentración libre de ligando en términos de  $n$ ,  $K$ ,  $[L]_T$  y  $[M]_T$ :

$$[L] = \frac{1}{2K} \cdot [-(1 + nK[M]_T) - K[L]_T + \sqrt{(1 + nK[M]_T - K[L]_T)^2 + 4K[L]_T}] \quad (2.20)$$

Derivando la ecuación 2.18 respecto a  $[L]_T$  obtendremos una expresión para el calor por mol de ligando añadido asociado a cada inyección:

$$\frac{1}{V_C} \cdot \frac{dQ}{d[L]_T} \approx \frac{1}{V_C} \cdot \frac{\Delta Q}{\Delta[L]_T} = \frac{\Delta H_{ap}}{2} \cdot \left[ 1 - \frac{1 + K[M]_T \cdot \theta - nK[M]_T}{\sqrt{(1 + K[M]_T \theta - nK[M]_T)^2 - 4nK^2[M]_T \theta}} \right] \quad (2.21)$$

$$\text{donde } \theta = \frac{[L]_T}{[M]_T}.$$

De acuerdo con estas ecuaciones, existen dos formas posibles de analizar los calores experimentales, una la del calor por mol de ligando añadido asociado a cada inyección, y otra la de calor acumulado. La primera, que es la que hemos utilizado en este trabajo, tiene la ventaja de evitar la propagación de errores experimentales debido a que se puede prescindir de los puntos con error para el análisis, mientras que en el caso de los calores acumulados no es posible.

Por lo tanto, las ecuaciones 2.20 y 2.21 son las que hemos utilizado para el ajuste de los datos experimentales, mediante el módulo para calorimetría isotérmica de titulación del programa Origin 5.0. Una vez corregida la línea base, el termograma

neto se integra para obtener los calores correspondientes a la titulación, los cuales se corrigen por los calores de dilución del ligando y se normalizan por la concentración de ligando total en célula tras cada inyección, que representados frente al cociente  $[L]_T/[M]_T$  permiten obtener la isoterma de unión. Ajustando dicha curva a la ecuación 2.21 mediante regresión no lineal es posible obtener los parámetros  $n$ ,  $K$  y  $\Delta H_{ap}$  de dicha interacción.

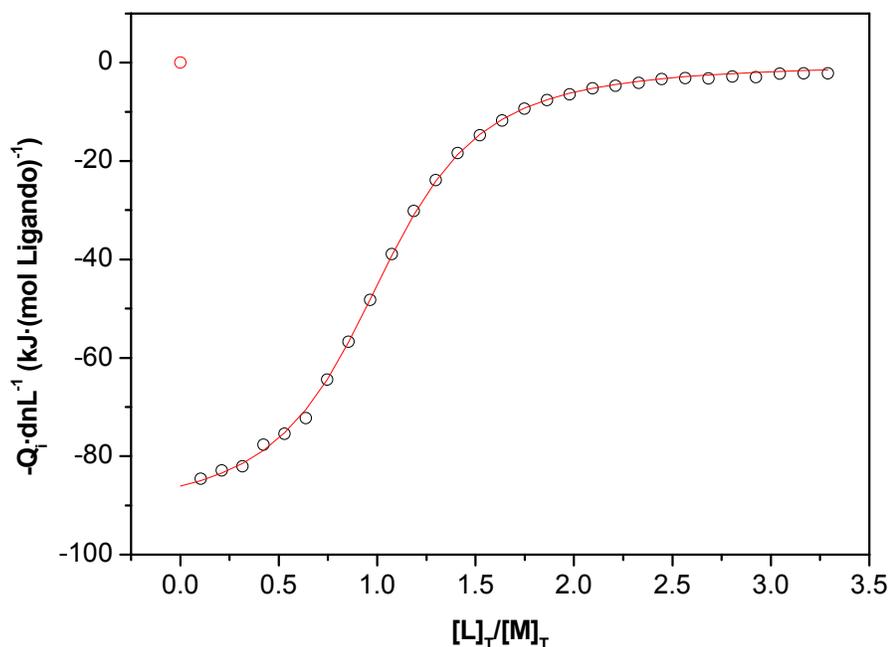
La Figura 2.5 muestra los calores por mol de ligando añadido frente al cociente  $[L]_T/[M]_T$  y su ajuste por regresión no lineal a la ecuación 2.21, de donde obtenemos el valor de los parámetros  $n$ ,  $K$  y  $\Delta H_{ap}$ , dejándolos como parámetros variables en dicho ajuste. El resto de parámetros termodinámicos que caracterizan la interacción se obtienen haciendo uso de las ecuaciones fundamentales  $\Delta G^0 = -RT \ln K$  y  $\Delta G^0 = \Delta H^0 - T\Delta S^0$  (Jelesarov & Bosshard 1999; Velazquez Campoy & Freire 2005).

#### **2.2.9.4. Formulación y análisis de un experimento de competición de ITC**

Uno de los factores limitantes de la técnica ITC es la imposibilidad de determinar los parámetros termodinámicos que caracterizan un proceso de unión con precisión cuando la afinidad es muy baja o muy alta. Como ya hemos comentado, un intervalo de medida adecuado se caracteriza por un valor del producto  $K \cdot [M]_T$  entre 1 y 100 (Wiseman *et al.* 1989), lo que limita el intervalo práctico de medida para la constante de asociación entre  $10^3$  y  $10^8$   $M^{-1}$ . Por fortuna, muchas de las interacciones establecidas entre biomoléculas están caracterizadas por afinidades dentro de estos límites, pero otras tantas tienen afinidades mayores o menores que las descritas.

La medida directa de constantes de asociación muy grandes requiere una concentración de biomolécula muy baja, lo que conlleva obtener una señal de medida pequeña y fuera del límite de detección instrumental. Excepcionalmente, cuando el valor de entalpía de unión es muy grande puede que la señal medida sea mayor, por lo que es posible determinar los parámetros termodinámicos de unión.

## Materiales y métodos

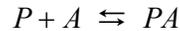


*Fig.2.5. Dependencia del calor liberado tras la inyección de L a M frente al cociente concentración de ligando total/concentración de proteína total. Los círculos representan los datos experimentales y la línea el ajuste al modelo de "n" sitios de unión idénticos e independientes (ecuación 2.21).*

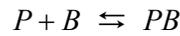
Pero para la mayoría de los casos la solución a este problema es la utilización de un experimento de competición en el cual un ligando de menor afinidad, cuyos parámetros termodinámicos de unión han sido determinados anteriormente, inhibe competitivamente al ligando de mayor afinidad (Sigurskjold 2000; Sigurskjold *et al.* 1994).

En el supuesto que tengamos una afinidad muy baja, como ocurre con algunos de los ligandos estudiados en este trabajo, necesitaríamos aumentar mucho la concentración de proteína para obtener un aumento de la señal calorimétrica, lo que no siempre es posible debido a problemas de solubilidad. La solución es llevar a cabo un experimento de competición con un ligando de mayor afinidad, cuyos parámetros de unión a la proteína son conocidos (Sigurskjold 2000; Zhang & Zhang 1998). Puesto que este ha sido el caso que se nos ha planteado en esta Memoria, desarrollaremos la formulación necesaria para el cálculo de los parámetros termodinámicos de un proceso de muy baja afinidad, no obstante, la formulación

para procesos de alta afinidad se obtiene desarrollando un procedimiento similar. Supongamos que tenemos dos ligandos, A y B, que pueden unir al mismo sitio de unión de una proteína, P:



y



por tanto, las constantes de equilibrio serían:

$$K_A = \frac{[PA]}{[P][A]} \quad (2.22)$$

y

$$K_B = \frac{[PB]}{[P][B]} \quad (2.23)$$

para  $K_A$  y  $K_B$ , respectivamente. Además, se ha supuesto que no es posible la formación de complejos ternarios como PAB, y que A y B compiten por el mismo sitio de unión. Si las concentraciones iniciales de A, B y P las llamamos,  $[A]_0$ ,  $[B]_0$  y  $[P]_0$  podemos obtener las siguientes expresiones a partir de la ley de conservación de masas:

$$[A]_0 = [A] + [PA] \quad (2.24)$$

$$[B]_0 = [B] + [PB] \quad (2.25)$$

y sustituyendo en las ecuaciones 2.22 y 2.23 llegamos a las siguientes expresiones:

$$[PA] = \frac{[P][A]_0}{1/K_A + [P]} \quad (2.26)$$

$$[PB] = \frac{[P][B]_0}{1/K_B + [P]} \quad (2.27)$$

Si A es el ligando que une moderadamente a P y B el ligando que une débilmente, entonces un experimento de titulación del ligando A sobre la

## Materiales y métodos

macromolécula P en presencia del ligando B, el ligando A desplazará parcialmente al ligando B del complejo PB en un grado que dependerá del valor de las constantes de asociación y de las concentraciones de A y B. Las concentraciones molares de especies que contengan P puede definirse como  $x_P = [P]/[P]_0$ ,  $x_{PA} = [PA]/[P]_0$ , y  $x_{PB} = [PB]/[P]_0$ , mientras que el cociente de las concentraciones estequiométricas de ligando con P se definen como  $r_A = [A]_0/[P]_0$  y  $r_B = [B]_0/[P]_0$ , y los productos de las constantes de asociación y las concentraciones estequiométricas de P como  $c_A = K_A [P]_0$  y  $c_B = K_B [P]_0$ . Sustituyendo estas expresiones en las ecuaciones anteriores llegamos a:

$$x_P + x_{PA} + x_{PB} = 1 \quad (2.28)$$

$$x_{PA} = \frac{r_A x_P}{1/c_A + x_P} \quad (2.29)$$

$$x_{PB} = \frac{r_B x_P}{1/c_B + x_P} \quad (2.30)$$

Ahora, sustituyendo las ecuaciones 2.29 y 2.30 en la ecuación 2.28 obtenemos una ecuación cúbica en función de  $x_P$ :

$$x_P^3 + ax_P^2 + bx_P + c = 0 \quad (2.31)$$

donde los coeficientes  $a$ ,  $b$  y  $c$  corresponden con:

$$a = \frac{1}{c_A} + \frac{1}{c_B} + r_A + r_B - 1, \quad (2.32)$$

$$b = \frac{r_A - 1}{c_B} + \frac{r_B - 1}{c_A} + \frac{1}{c_A c_B}, \quad (2.33)$$

y

$$c = \frac{-1}{c_A c_B} \quad (2.34)$$

La resolución matemática de la ecuación 2.31 genera tres soluciones, de las cuales la única solución física posible es:

$$x_P = \frac{2\sqrt{a^2 - 3b} \cos(\theta/3) - a}{3} \quad (2.35)$$

en la cual,

$$\theta = \arccos \frac{-2a^3 + 9ab - 27c}{2\sqrt{(a^2 - 3b)^3}} \quad (2.36)$$

Una vez que  $x_p$  es conocida, también puede conocerse  $x_{PA}$  y  $x_{PB}$  utilizando las ecuaciones 2.29 y 2.30.

En un experimento de ITC de competición de este tipo el calor liberado tras cada inyección será proporcional a los cambios en  $[PA]$  y  $[PB]$  y sus respectivas entalpías molares de unión. Por tanto, el calor que medimos será igual a:

$$\begin{aligned} \Delta Q &= V_C (\Delta H_A \Delta[PA] + \Delta H_B \Delta[PB]) \\ &= V_C [P_0] (\Delta H_A \Delta x_{PA} + \Delta H_B \Delta x_{PB}) \end{aligned} \quad (2.37)$$

donde,  $V_C$  es el volumen de la célula de reacción del calorímetro. Al igual que en un experimento directo de ITC, también tendremos que aplicar una corrección de volúmenes para calcular con precisión las concentraciones molares de las especies P, A y B en la célula del calorímetro. Así, si llamamos  $[X]$  a la concentración de cualquiera de estas especies, entonces un cambio infinitesimal de volumen afectará del siguiente modo:

$$d[X] = -\frac{dV_i}{V_0} [X], \quad (2.38)$$

y la integración de esta ecuación diferencial con  $[X]$  entre los límites  $[X]_{i-1}$  y  $[X]_i$  y  $V_i$  entre 0 y  $V_i$  resulta en la siguiente expresión:

$$[X]_i = [X]_{i-1} \exp\left(-\frac{V_i}{V_0}\right), \quad (2.39)$$

Por tanto, el factor exponencial de la ecuación 2.39 es el factor de dilución para una determinada inyección:

$$f_i = \exp\left(-\frac{V_i}{V_0}\right) \quad (2.40)$$

## Materiales y métodos

Cuando el ligando de mayor afinidad A es el que se inyecta y el ligando B es el que contiene la célula de reacción, entonces las concentraciones estequiométricas tras cada inyección serán:

$$[A]_i = [A]_{i-1}(1 - f_i), \quad (2.41)$$

,

$$[P]_i = f_i[P]_{i-1} \quad (2.42)$$

y

$$[B]_i = f_i[B]_{i-1} \quad (2.43)$$

La ecuación 2.37 puede expresarse ahora como:

$$\Delta Q_i = V_C [P_0] (\Delta H_A \{ \Delta x_{PA,i} - f_i x_{PA,i-1} \} + \Delta H_B \{ \Delta x_{PB,i} - f_i x_{PB,i-1} \}) \quad (2.44)$$

La ecuación 2.44 es la utilizada en el ajuste por regresión no lineal de mínimos cuadrados a partir de las curvas isotermas de unión de los datos de ITC, donde las concentraciones y fracciones molares son calculadas a partir de las ecuaciones 2.28-2.30, 2.32-2.36, y 2.40-2.43, implementadas en la plantilla ofrecida generosamente por el Dr. Sigurskjold (Universidad de Copenhagen, Dinamarca) (Sigurskjold 2000). El factor  $x_{PA,i} - x_{PA,i-1}$  será siempre positivo, de modo que la contribución de la unión del ligando A al calor tendrá siempre el mismo signo que  $\Delta H_A$ . La expresión  $x_{PB,i} - x_{PB,i-1}$  será siempre negativa, de modo que la disociación del ligando B del complejo PB contribuirá al calor total con el signo opuesto a  $\Delta H_B$ .

### 2.2.9.5. Interpretación de los parámetros termodinámicos obtenidos mediante ITC

Desde un punto de vista termodinámico más riguroso, es conveniente expresar la ecuación sencilla 2.7 del siguiente modo:

$$\Delta G = \Delta H - T\Delta S_{sol} - T\Delta S_{conf} - T\Delta S_{r,tr} \quad (2.45)$$

donde  $\Delta S_{sol}$  corresponde a la entropía de solvatación (cambio entrópico asociado con el cambio en solvatación de las moléculas que se unen),  $\Delta S_{conf}$  es la entropía conformacional (cambio entrópico asociado con el cambio conformacional de las moléculas que se unen) y  $\Delta S_{r,tr}$  es la entropía roto-translacional (cambio entrópico asociado con el cambio en la configuración espacial de las moléculas que se unen).

El término entálpico refleja la contribución energética de muchas interacciones individuales (enlaces de hidrógeno, van der Waals, electrostáticas, dipolo-dipolo, etc.) entre las dos moléculas, teniendo en cuenta las interacciones con el disolvente y, por tanto, refleja también la entalpía de solvatación. Este término es negativo (favorable) si las interacciones entre las moléculas que se unen compensan las interacciones que establecen con el disolvente cuando se encuentran libres, en otro caso, sería desfavorable (Velazquez-Campoy *et al.* 2004).

La entropía de solvatación representa la ganancia en grados de libertad de las moléculas de agua que previamente se encuentran en la superficie de las moléculas interaccionantes y que son liberadas tras la unión. Esta es normalmente positiva (favorable) y grande, ya que los ligandos poseen generalmente mayor porcentaje de área apolar accesible al disolvente enterrada tras la unión (Jelesarov & Bosshard 1999; Velazquez-Campoy & Freire 2006; Velazquez-Campoy *et al.* 2004). Este enterramiento de la superficie molecular expuesta al disolvente tras la unión es la mayor contribución al cambio de capacidad calorífica, aunque los cambios en los modos vibracionales de las moléculas interaccionantes también contribuyen pero en menor proporción (Gomez & Freire 1995; Makhatadze & Privalov 1995; Murphy & Freire 1992; Spolar *et al.* 1992; Sturtevant 1977; Velazquez-Campoy *et al.* 2004). Este término suele ser negativo si el complejo es tomado como estado de referencia e indica que la capacidad calorífica del complejo es menor que las de sus componentes por separado.

Finalmente, la entropía conformacional y roto-translacional son normalmente negativas (desfavorable), debido a la pérdida de grados de libertad para ambas moléculas cuando se unen. En el caso de la entropía conformacional es posible optimizar su valor modificando estructuralmente las moléculas que se unen, mientras

## ***Materiales y métodos***

que la entropía roto-translacional es esencialmente la misma en cualquier proceso de unión y no puede ser modificada (Velazquez Campoy & Freire 2005).

Aunque parece evidente que la entalpía y entropía reflejan interacciones de distinta naturaleza, es bastante frecuente la situación en que la modificación de alguna de las moléculas que interaccionan provoca un aumento de afinidad de unión en una dirección (por ejemplo, entálpicamente, mediante la formación de un nuevo puente de hidrógeno), lo que provoca un efecto opuesto en la otra dirección (entrópicamente, a través de la pérdida de grados de libertad conformacionales o una menor desolvatación tras la unión), mostrando la misma afinidad de unión, fenómeno que se conoce como compensación entálpica/entrópica.

Tradicionalmente, en la mayoría de los casos se ha conseguido un aumento de la afinidad de los ligandos manipulando la contribución entrópica, haciendo los ligandos mas hidrofóbicos y restringidos estructuralmente. De este modo, la naturaleza entrópica del ligando sería la fuerza que dirige la interacción con la molécula que se une, ya que tendría la forma y geometría más idónea para adaptarse al sitio de unión. Esta es una de las razones por las que la mayoría de librerías orientadas hacia la identificación de ligandos para aumentar la afinidad contienen compuestos generalmente hidrofóbicos (Velazquez Campoy & Freire 2005).

El termino entálpico es mas difícil de modular y requiere un conocimiento mas preciso de la interacciones intermoleculares de tipo van der Waals, puentes de hidrógeno, electrostáticas, etc., y el delicado, pero importante, balance entre la red de interacciones establecidas en la forma complejada y la red de interacciones establecidas por las moléculas libres con el disolvente. Por lo tanto, considerando que un aumento ideal en afinidad es aquel que proviene de la optimización del término entálpico y entrópico, los ligandos dirigidos entálpicamente tienen mayor potencial para optimizar, ya que puede ser más fácil mejorar la entropía que la entalpía (Velazquez-Campoy *et al.* 2000b). Por un lado, una contribución entálpica favorable es una indicación de interacciones específicas entre las moléculas que se unen y es la dirección adecuada para asegurar especificidad, selectividad y adaptabilidad. Al contrario, una entalpía desfavorable indica interacciones

inespecíficas, lo que hace bastante difícil conseguir especificidad, selectividad y adaptabilidad (Freire 2002; Ohtaka *et al.* 2004; Todd *et al.* 2000; Vega *et al.* 2004; Velazquez-Campoy *et al.* 2000a; Velazquez-Campoy *et al.* 2000b).

Como ya hemos comentado, ITC es la única técnica capaz de proporcionar información termodinámica de la unión de dos moléculas, aunque no genera información detallada (a nivel atómico o de grupo funcional) sobre el origen o la localización de interacciones individuales. Sin embargo, es posible modificar las condiciones experimentales (pH, fuerza iónica, etc.) o los grupos funcionales de las moléculas que se unen (mediante el diseño de proteínas mutantes o ligandos modificados) para obtener información acerca las fuerzas y determinantes estructurales de la interacción.

#### **2.2.10. Calorimetría diferencial de barrido**

La calorimetría diferencial de barrido (DSC, es la abreviatura del inglés) es actualmente la técnica más completa para caracterizar la energética del plegamiento-desplegamiento de las proteínas. En general, con esta técnica se estudian los cambios conformacionales inducidos por la temperatura en proteínas, ácidos nucleicos, membranas lipídicas, etc. Numerosos han sido los trabajos donde se ha estudiado la desnaturalización térmica de proteínas mediante DSC {Privalov, 1974 #89} (Privalov & Potekhin 1986) {Privalov, 1989 #83} {Freire, 1990 #92} (Sanchez-Ruiz 1995) (Filimonov *et al.* 1999), (Martinez *et al.* 1999) (Conejero-Lara & Mateo 1996; Ruiz-Sanz *et al.* 1999).

Un experimento de calorimetría diferencial de barrido consiste en registrar la capacidad calorífica aparente de una disolución de macromolécula en función de la temperatura. El posterior análisis basándonos en la termodinámica de equilibrio puede proporcionar la caracterización completa del proceso de desplegamiento térmico de una proteína. Esto permite obtener la capacidad calorífica parcial absoluta del sistema y los parámetros termodinámicos asociados, tales como el cambio de entalpía,  $\Delta H_D$ , de entropía,  $\Delta S_D$ , de energía de Gibbs,  $\Delta G_D$ , y de capacidad calorífica,  $\Delta C_p$ , correspondientes a la transición inducida por la

## ***Materiales y métodos***

temperatura, así como la función de partición y la población de los estados predominantes en los que se encuentra el sistema y sus parámetros termodinámicos característicos. El modelo más simple que existe es el de dos estados, el cual supone que las proteínas sólo pueden encontrarse en estado nativo o en estado desplegado. Otros modelos suponen la existencia de intermedios de plegamiento significativamente poblados y/o estados de asociación.

### **2.2.10.1. Diseño del experimento de calorimetría diferencial de barrido**

Para la realización de todos los experimentos de DSC de esta Memoria se ha empleado el microcalorímetro VP-DSC, de Microcal Inc. {Plotnikov, 1997 #96}, que nos permite medir la capacidad calorífica de forma continua, calentando o enfriando a velocidad constante, y diferencial, midiendo la diferencia de capacidad calorífica entre dos células que contienen la disolución de la muestra objeto del estudio y la disolución de referencia.

El experimento calorimétrico se inicia llenando ambas células con el mismo tampón que el de la proteína, y se registran varios barridos de temperatura que corresponden a lo que se denomina línea base instrumental. Si las dos células del instrumento fuesen idénticas, sólo se necesitaría un barrido para determinar la diferencia de capacidad calorífica existente entre la célula de muestra y la de referencia, pero como no son iguales es necesario restar la contribución a la señal experimental debida a estos factores instrumentales. Así, se registran barridos de temperatura sucesivos (sin rellenar las células al finalizar cada registro) para asegurarnos de que la línea base es reproducible. Una vez adquirida la línea base instrumental, se llena la célula de muestra con la disolución de proteína. Después de haber realizado el primer barrido con la muestra es necesario realizar un segundo barrido, ya que para poder aplicar la termodinámica de equilibrio el proceso tiene que ser reversible térmicamente. Los termogramas obtenidos representan la capacidad calorífica aparente de la disolución de proteína,  $\Delta C_p(T)$ , en función de la temperatura. A partir de ésta podemos obtener los parámetros termodinámicos que caracterizan el proceso de desnaturalización térmica, como se describe a continuación.

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

Los ficheros generados por el programa de adquisición de datos contienen el valor de una señal proporcional a la variación de la capacidad calorífica de la disolución de proteína (en mV) con la temperatura. Los ficheros se importan con el programa MicroCal Origin for DSC v 4.1 en el que se lleva a cabo tanto una conversión de forma automática de unidades de mV a unidades típicas de capacidad calorífica como  $\text{cal}\cdot\text{C}^{-1}$  como la corrección dinámica del termograma, basándose en la última calibración realizada y en el tiempo de respuesta preseleccionado. Una vez aquí se resta la línea base instrumental y se normalizan los datos por la concentración de proteína y el volumen efectivo de la célula, expresando la capacidad calorífica en términos de capacidad calorífica molar parcial  $\text{kJ}\cdot(\text{K}\cdot\text{mol}^{-1})$ .

La diferencia de capacidad calorífica entre la disolución de macromolécula y la del tampón que la contiene es normalmente negativa ya que la capacidad calorífica de la célula de referencia es mayor que la de muestra {Privalov, 1974 #89} debido a que la proteína desplaza una cantidad de agua cuya capacidad calorífica es mayor que la de la proteína. Esta diferencia se puede escribir como:

$$\Delta C_p(T) = C_p(T) \cdot m_p - C_{p,s}(T) \cdot \Delta m_s \quad (2.46)$$

donde  $C_p(T)$  y  $C_{p,s}(T)$  son las capacidades caloríficas parciales de la proteína y del disolvente respectivamente,  $m_p$  es la cantidad de proteína en la célula calorimétrica y  $\Delta m_s$  la cantidad de disolvente desplazado por la proteína en disolución. Teniendo en cuenta además que:

$$\Delta m_s = m_p \cdot \frac{V_p(T)}{V_s(T)} \quad (2.47)$$

donde  $V_p(T)$  y  $V_s(T)$  son los correspondientes volúmenes específicos parciales de la proteína en disolución y del disolvente respectivamente, así:

$$C_p = C_{p,s}(T) \cdot \frac{V_p(T)}{V_s(T)} + \frac{\Delta C_p(T)}{m_p} \quad (2.48)$$

## **Materiales y métodos**

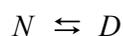
Teniendo en cuenta que generalmente se trabaja con disoluciones acuosas, que la capacidad calorífica y el volumen específico del agua son conocidos y aceptando como válido que para la proteína el volumen específico parcial es  $V_p=0.73 \text{ mL}\cdot\text{g}^{-1}$ , valor que se usa como promedio para todas las proteínas globulares (Makhatadze *et al.* 1990), se puede calcular la capacidad calorífica molar parcial de la proteína.

### **2.2.10.3. Análisis de las trazas calorimétricas según la termodinámica de equilibrio**

Del análisis de la capacidad calorífica molar parcial de la proteína como una función de la temperatura se pueden obtener los parámetros termodinámicos del proceso de desnaturalización térmica, para lo que aplicaremos un modelo de equilibrio una vez comprobado que el proceso es reversible. A continuación se describe el modelo utilizado para el análisis de los termogramas que se presentan en esta Memoria.

#### **2.2.10.3.1. Modelo de equilibrio de dos estados**

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). Así, para el proceso de equilibrio tenemos:



La constante de equilibrio aparente, para la desnaturalización, viene dada por:

$$K_D = \frac{[D]}{[N]} \quad (2.49)$$

y la función de partición del sistema, en función de la constante de equilibrio, se define como:

$$q = 1 + K_D \quad (2.50)$$

Las poblaciones para los estados N y D a partir de las ecuaciones anteriores son:

$$f_N = \frac{1}{q} \quad (2.51)$$

$$f_D = \frac{K_D}{q} \quad (2.52)$$

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

$$\langle H \rangle = f_N \cdot H_N + f_D \cdot H_D = H_N + f_D \cdot \Delta H_D = H_N + \frac{K_D}{q} \cdot \Delta H_D \quad (2.53)$$

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

$$C_p = \left[ \frac{\partial \langle H \rangle}{\partial T} \right]_p = C_{p,N} + \frac{\Delta H_D}{(1 + K_D)^2} \left[ \frac{\partial K_D}{\partial T} \right]_p + \frac{K_D}{(1 + K_D)} \left[ \frac{\partial \Delta H_D}{\partial T} \right]_p \quad (2.54)$$

donde  $C_{p,N}$  es la capacidad calorífica molar parcial del estado nativo. Si ahora expresamos las dependencias de  $K_D$  y  $\Delta H_D$  con la temperatura, nos queda:

$$\frac{\partial \Delta H_D}{\partial T} = \Delta C_p \quad (2.55)$$

$$\frac{\partial \Delta S_D}{\partial T} = \frac{\Delta C_p}{T} \quad (2.56)$$

$$K_D = \exp\left(-\frac{\Delta G_D}{T}\right) \quad (2.57)$$

## Materiales y métodos

y así,  $\Delta S_D$ ,  $\Delta G_D$  y  $\Delta C_p$  son las diferencias de entropía, de energía de Gibbs y de capacidad calorífica entre los estados D y N respectivamente. Según los trabajos de Privalov y colaboradores, la función que describe la 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 se puede estimar a partir de su secuencia aminoacídica (Makhatadze *et al.* 1990) {Privalov, 1990 #68}. Así, la función  $C_{p,D}$ , que es la capacidad calorífica molar parcial del estado desnaturalizado, puede describirse como un polinomio de segundo grado, aunque en algunos casos se han descrito modelos de equilibrio suponiendo una dependencia lineal (Filimonov *et al.* 1999). La función que describe a la capacidad calorífica molar parcial del estado nativo,  $C_{p,N}$ , es prácticamente lineal en el intervalo antes de la transición {Privalov, 1989 #83}. Así, se obtienen las siguientes ecuaciones para  $C_{p,D}$ ,  $C_{p,N}$  y  $\Delta C_p$ :

$$C_{p,D} = a + b \cdot T + c \cdot T^2 \quad (2.58)$$

$$C_{p,N} = d + e \cdot T \quad (2.59)$$

$$\Delta C_p = (a - d) + (b - e) \cdot T + c \cdot T^2 \quad (2.60)$$

Para obtener las dependencias de  $\Delta H_D$  y  $\Delta S_D$  con la temperatura integramos las ecuaciones 2.55 y 2.56:

$$\Delta H_D(T) = \Delta H_m + \int_{T_m}^T \Delta C_p \cdot dT \quad (2.61)$$

$$\Delta S_D(T) = \Delta S_m + \int_{T_m}^T \frac{\Delta C_p}{T} \cdot dT \quad (2.62)$$

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

$$\Delta S_m = \frac{\Delta H_m}{T_m} \quad (2.63)$$

Si integramos las ecuaciones de  $\Delta H_D$  y de  $\Delta S_D$  con respecto a la temperatura, nos queda que:

$$\Delta H_D = \Delta H_m + (a-d) \cdot (T - T_m) + \frac{(b-e)}{2} \cdot (T^2 - T_m^2) + \frac{1}{3} \cdot c \cdot (T^3 - T_m^3) \quad (2.64)$$

$$\Delta S_D = \frac{\Delta H_m}{T_m} + (a-d) \cdot \ln\left(\frac{T}{T_m}\right) + (b-e) \cdot (T - T_m) + \frac{1}{2} \cdot c \cdot (T^2 - T_m^2) \quad (2.65)$$

$$\Delta G_D = \Delta H_D - T \cdot \Delta S_D \quad (2.66)$$

Y así, podemos expresar la capacidad calorífica molar parcial de la siguiente manera:

$$C_p = C_{p,N} + \frac{(\Delta H_D)^2}{RT^2} \cdot \frac{K_D}{(1+K)^2} + f_D \cdot \Delta C_p \quad (2.67)$$

En el caso de que la proteína muestre una dependencia lineal frente a la temperatura para la capacidad calorífica molar parcial del estado desnaturalizado, entonces aplicando un procedimiento similar al descrito anteriormente se llegaría a las siguientes expresiones:

$$\Delta C_p = (a-d) + (b-e) \cdot T \quad (2.68)$$

$$\Delta H_D = \Delta H_m + (a-d) \cdot (T - T_m) + \frac{(b-e)}{2} \cdot (T^2 - T_m^2) \quad (2.69)$$

$$\Delta S_D = \frac{\Delta H_m}{T_m} + (a-d) \cdot \ln\left(\frac{T}{T_m}\right) + (b-e) \cdot (T - T_m) \quad (2.70)$$

De acuerdo con estas expresiones, las ecuaciones 2.66 y 2.67 se mantendrían iguales, por lo que para el ajuste no lineal por mínimos cuadrados de

## ***Materiales y métodos***

las curvas de  $C_p$  hemos utilizado las ecuaciones 2.66-2.70, implementadas en el programa Origin 6.1 (Microcal Software Inc.). Para ello, antes es preciso definir todas las magnitudes de las que depende  $C_p$  en dicha ecuación, en función de la temperatura. De esta manera,  $C_p$  queda expresada en función de la temperatura y una serie de parámetros, de cuyo ajuste obtenemos los valores de  $T_m$ ,  $\Delta H_m$  y  $\Delta C_p$  (Apéndice I)

### **2.2.11. Dicroísmo circular**

El dicroísmo circular (CD) se basa en que las moléculas quirales, entre ellas las proteínas, en ausencia de campos magnéticos no interaccionan igual con la luz polarizada circularmente hacia la derecha y hacia la izquierda (Woody *et al.* 1996). Concretamente, el CD es una técnica espectroscópica con la que se puede determinar la diferencia entre la absorción de la luz polarizada circularmente hacia la derecha y hacia la izquierda. La luz polarizada circularmente se consigue con dos haces de luz ortogonales que están polarizados en el plano y que se encuentran fuera de fase en 90°. El vector eléctrico de la luz polarizada circularmente se puede expresar matemáticamente (Kliger *et al.* 1990; Michel & Thulstrup 1986) como:

$$\mathbf{E}_{\pm} = E_0(\mathbf{i} \pm i\mathbf{j}) \exp[2\pi i(\nu t - z/\lambda)] \quad (2.71)$$

donde  $E_0$  es la amplitud de la onda,  $i$  y  $j$  son los vectores unidad en las coordenadas  $x$  e  $y$  respectivamente en un sistema de coordenadas cartesiano hacia la derecha en el que  $+z$  es la dirección de propagación,  $i$  es  $(-1)^{1/2}$  y  $\nu$  y  $\lambda$  son la frecuencia y la longitud de onda de la luz. La absorbancia de luz polarizada hacia la izquierda se puede expresar como:

$$A_i = \log_{10} \left( \frac{I_i^0}{I_i} \right) = \varepsilon_i c l \quad (2.72)$$

y hacia la derecha:

$$A_d = \log_{10} \left( \frac{I_d^0}{I_d} \right) = \varepsilon_d cl \quad (2.73)$$

donde  $I^0$  es la intensidad de luz polarizada que incide sobre la muestra,  $I_i^0$  e  $I_d^0$  son las intensidades de luz polarizada hacia la izquierda y hacia la derecha, respectivamente, que salen de la muestra,  $l$  es la distancia que recorre la luz a través de la muestra,  $c$  es la concentración de soluto y  $\varepsilon$  es el coeficiente de extinción molar del soluto para la luz polarizada hacia el lado correspondiente. De acuerdo con los principios teóricos del CD y las ecuaciones anteriores 2.72 y 2.73 podemos obtener la siguiente expresión:

$$\Delta A = A_i - A_d = \varepsilon_i \cdot cl - \varepsilon_d \cdot cl = (\varepsilon_i - \varepsilon_d) \cdot cl = \Delta \varepsilon \cdot cl \quad (2.74)$$

donde las unidades de  $\varepsilon$  son  $M^{-1} \text{ cm}^{-1}$ ,  $c$  es en molar y  $l$  en cm. Además de expresar el CD como en la ecuación 2.74, en el campo de la biofísica también se utiliza el método de medida original propuesto por Lowry (Lowry 1935). Este se basa en que cuando una luz polarizada circularmente pasa a través de un soluto quiral, la absorción diferencial de las dos componentes circulares hace que a la salida de la muestra cada componente siga estando circularmente polarizada, pero los radios de los círculos trazados por los vectores eléctricos son ahora diferentes, y cuando estas dos ondas luminosas opuestas, polarizadas circularmente, se combinan, el resultado sea luz polarizada elípticamente. Cuando los vectores eléctricos de las dos componentes circulares van en el mismo sentido, la suma de sus magnitudes es el eje mayor de la elipse, y cuando van en sentidos opuestos, la resta de sus magnitudes nos da el valor del eje menor de la elipse. De este modo el CD se puede caracterizar por la relación entre el eje mayor y menor, que es la tangente de un ángulo  $\theta$ , al que se denomina elipticidad. Como  $\theta$  es normalmente muy pequeño, puede expresarse como:

$$\theta(rad) = \tan \theta = \frac{|\mathbf{E}_i| - |\mathbf{E}_d|}{|\mathbf{E}_i| + |\mathbf{E}_d|} = \frac{\exp\left(\frac{-A_i}{2}\right) - \exp\left(\frac{-A_d}{2}\right)}{\exp\left(\frac{-A_i}{2}\right) + \exp\left(\frac{-A_d}{2}\right)} \quad (2.75)$$

## **Materiales y métodos**

Expandiendo las exponenciales, teniendo en cuenta el valor despreciable de  $\Delta A$  frente a la unidad y convirtiendo a grados, obtenemos:

$$\theta(\text{grad}) = 180 \cdot \ln 10 \cdot \Delta A / 4\pi = 32.98 \Delta A \quad (2.76)$$

donde se puede observar que la elipticidad es proporcional al dicroísmo circular. La elipticidad molar se define como:

$$[\theta] = \frac{\theta}{10 \cdot c \cdot l} \quad (2.77)$$

o bien la elipticidad molar media por resto:

$$[\theta]_{MMR} = \frac{\theta}{10 \cdot c \cdot N_A \cdot l} \quad (2.78)$$

donde  $\theta$  se introduce en miligrados (datos de salida del espectropolarímetro),  $c$  en molar,  $N_A$  es el número de aminoácidos, y  $l$  en cm;  $[\theta]$  y  $[\theta]_{MMR}$  se expresan en grados·dmol<sup>-1</sup>·cm<sup>2</sup>.

### **2.2.11.1. Espectros de dicroísmo circular de proteínas**

Mediante espectroscopía de dicroísmo circular se puede obtener información acerca de las propiedades conformacionales de proteínas, concretamente es posible:

- Determinar si una proteína se encuentra plegada y en tal caso, podemos caracterizar su estructura secundaria (ultravioleta lejano), estructura terciaria (ultravioleta cercano) y clasificarla dentro de una familia estructural.
- Comparar las estructuras de proteínas obtenidas mediante otras técnicas (rayos-X o RMN) o las estructuras de diferentes mutantes de una misma proteína.

- Estudiar la estabilidad conformacional de una proteína (estabilidad térmica, frente a cambios de pH o en presencia de agentes desnaturizantes) y cómo esta estabilidad se altera dependiendo de las condiciones utilizadas.
- Determinar los cambios conformacionales asociados a interacciones entre moléculas asimétricas (interacciones proteína-proteína, proteína-DNA, proteína-ligando y DNA-ligando).

La estructura secundaria de las proteínas se puede determinar mediante CD en el ultravioleta lejano (190-250 nm). Esta región es muy sensible a los cambios estructurales, fundamentalmente a los de estructura secundaria, ya que la contribución de la estructura terciaria, debida al enlace peptídico fundamentalmente, es despreciable. Aunque el enlace peptídico por si sólo no es una estructura quiral, la interacción entre grupos peptídicos adyacentes dispuestos en determinadas orientaciones determina que las proteínas presenten actividad óptica en la región del UV lejano, donde absorbe el enlace peptídico.

Cada elemento principal de estructura secundaria de las proteínas (hélices alfa, láminas beta) presenta un espectro característico, tal y como se ilustra en la Figura 2.6. Para la estructura no regular ocurre lo mismo, aunque lógicamente el espectro presenta mayor variabilidad. En concreto las hélices alfa presentan un espectro de muy característico en la región del UV lejano, constituido por un doble mínimo a 222 y a 208-210 nm y un máximo más intenso a 191-193 nm. La intensidad de las tres bandas (dependiente de los ángulos diedros de la proteína) da una idea de la cantidad de contenido helicoidal que tiene la proteína.

Las estructuras tipo beta suelen presentar un único pico negativo y un único pico positivo cuyas intensidades son menores que las de las hélices. El mínimo se encuentra normalmente entre 210 y 225 nm y el pico positivo, de mayor intensidad que el negativo, aparece entre 190 y 200 nm (Figura 2.6). Se han descrito algunos ejemplos (Manavalan & Johnson 1983; Wu *et al.* 1992) de proteínas que sólo poseen estructura  $\beta$  que muestran una banda negativa intensa alrededor de 200 nm, como los péptidos con estructura no regular. La explicación que se ha dado a estos casos, según los datos de difracción de rayos X, es que estas proteínas tienen hojas  $\beta$  cortas y bastante irregulares.

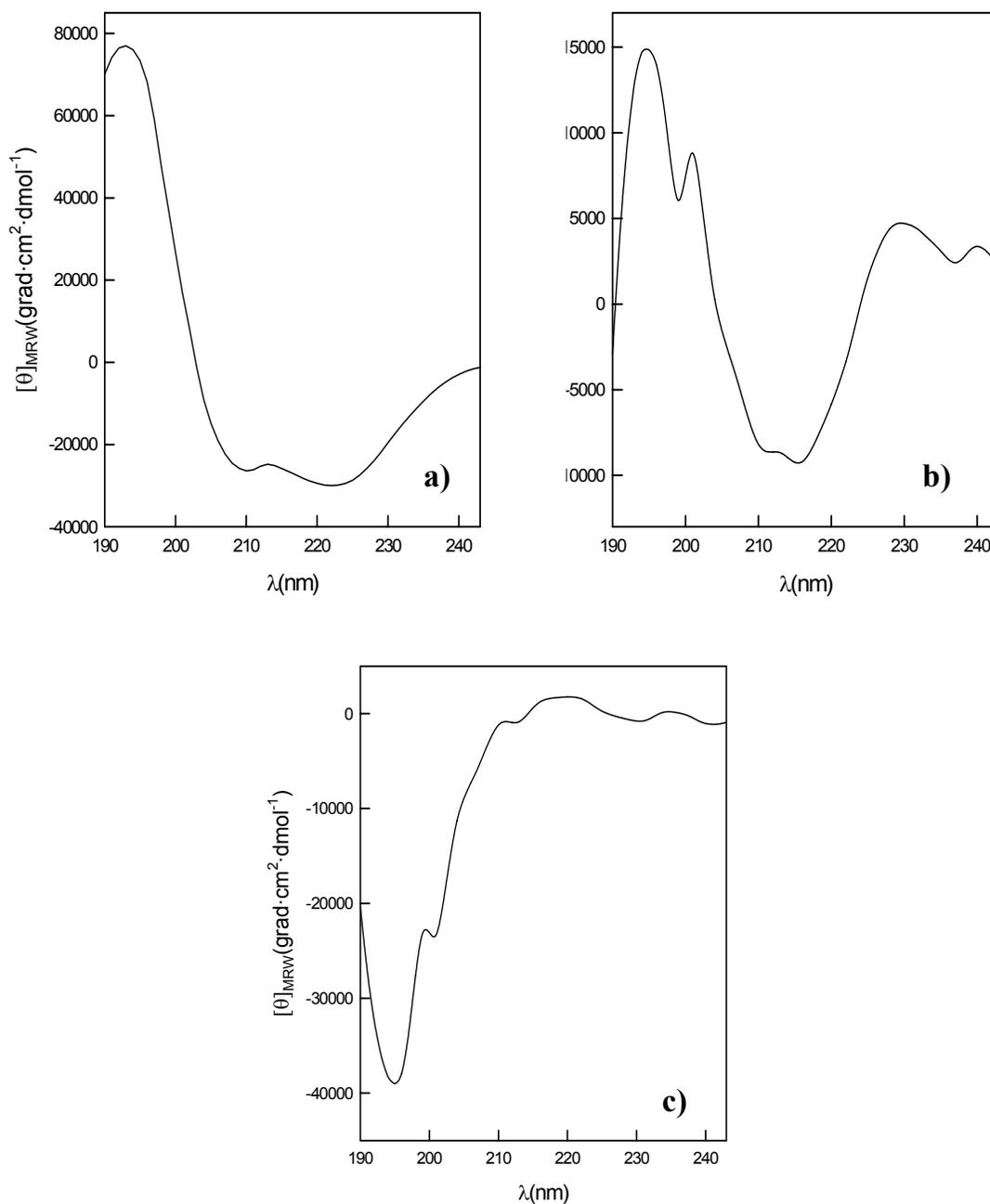
## ***Materiales y métodos***

Los péptidos y proteínas con estructura no regular (los bucles de las conformaciones nativas y, en general, las conformaciones desplegadas) presentan habitualmente una banda negativa fuerte cerca de 200 nm (Figura 2.6.) y algunas bandas débiles entre 220 y 230 nm que pueden ser tanto positivas como negativas. Hay que recordar, sin embargo, que en el UV lejano también absorben los aminoácidos aromáticos y que, aunque sus cadenas laterales no son quirales, se comportan como tales dependiendo de su entorno proteico.

Por esta razón el espectro de algunas proteínas aparece dominado por contribuciones peculiares debidas a estos restos. Con esta salvedad, la actividad óptica de los enlaces peptídicos es de gran utilidad para conocer el contenido de estructura secundaria de las proteínas. Así, cada proteína tiene un espectro de dicroísmo circular en el ultravioleta lejano que es una combinación lineal de los espectros característicos de las estructuras secundarias y no periódicas que contiene.

El espectro ultravioleta cercano (250-350 nm) de una proteína es sensible a cambios en la estructura terciaria. A estas longitudes de onda los cromóforos son los aminoácidos aromáticos y los enlaces disulfuro cuando se encuentran inmovilizados en un entorno estructural asimétrico, típicamente el estado nativo de algunas proteínas.

La contribución de la fenilalanina al espectro es pequeña, debido a su bajo coeficiente de extinción; además, es poco sensible a las alteraciones de su ambiente. El espectro se caracteriza por una banda de absorción en la región 250-270 nm. La tirosina tiene una banda de absorción más intensa a 276 nm y un hombro a 283 nm, que pueden desplazarse hacia el rojo debido a la formación de puentes de hidrógeno por el grupo fenólico. El triptófano presenta la banda de absorción más intensa, que está centrada a 282 nm y también puede desplazarse hacia el rojo debido a la formación de un puente de hidrógeno por el grupo indol. A diferencia de lo que ocurría en el lejano, las absorciones de las cadenas laterales de estos restos aromáticos no se mezclan con las del grupo peptídico (que no absorbe en esta región).



**Figura 2.6.** Espectro ultravioleta lejano de CD de una proteína con estructura de hélice alfa (a), de una proteína con estructura tipo beta (b), y de una proteína con estructura no regular (c).

Así, el espectro de dicroísmo circular de una proteína en el ultravioleta cercano es una pequeña huella dactilar de la conformación nativa cuya pérdida nos indicará que han ocurrido cambios conformacionales apreciables.

### **2.2.11.2. Diseño y ejecución de los experimentos de CD**

Los experimentos en el ultravioleta lejano se registraron desde 260 hasta 195 nm para lo que se utilizaron cubetas de 2 mm de paso de luz y concentraciones de muestra de 30  $\mu$ M. Inicialmente, se debe de registrar un espectro con la cubeta llena del tampón de diálisis, que se utilizará como blanco, con el fin de eliminar las diferentes contribuciones instrumentales y del disolvente al espectro de CD de la proteína. Antes de registrar cada espectro, se deja la muestra equilibrar a la temperatura deseada al menos 5 min.

El programa de adquisición del instrumento permite promediar  $N$  espectros consecutivos de forma que el nivel de ruido resultante es  $N^{1/2}$  veces menor que el de cada espectro individual, ganándose de esta manera mucha sensibilidad. En todos los casos se realizó una acumulación de 16 espectros, que en nuestro caso fue suficiente para eliminar gran parte del ruido, y el resto de parámetros instrumentales fueron:

- Resolución: 1 nm.
- Velocidad de toma de espectro: 100 nm/min.
- Tiempo de respuesta: 1 segundo.
- Ancho de ventana de toma de datos: 1 nm.
- Flujo de nitrógeno gas para enfriamiento de las lámparas: 10 L/min.

### **2.2.12. Cristalización de proteínas**

#### **2.2.12.1. Introducción**

La cristalización de moléculas en solución es un fenómeno de equilibrio químico cuya cinética y termodinámica están controladas por las características físicas (temperatura, gravedad, superficie, viscosidad, constante dieléctrica o vibraciones) y químicas (pH, fuerza iónica, iones metálicos o detergentes) del disolvente y del soluto. Bajo ciertas condiciones, el sistema origina una fase sólida a

partir de la disolución que disminuye la energía libre del sistema (núcleo crítico o un agregado no específico) (Weber 1991). Esta disminución de energía libre es mayor cuanto más ordenada sea la fase segregada.

El crecimiento cristalino generalmente se considera dividido en tres etapas: nucleación, crecimiento y cese del crecimiento. La nucleación es el proceso por el cual una fase sólida se segrega de la disolución para formar agregados cristalinos termodinámicamente estables, además proporciona al sistema superficie para seguir creciendo. Cuando la nucleación está libre de partículas y superficies extrañas se denomina nucleación homogénea. La presencia de partículas extrañas puede provocar la aparición de los núcleos a unos valores de sobresaturación inferiores a los requeridos para que se produzca la nucleación homogénea (nucleación heterogénea). El crecimiento de un cristal implica una serie de pasos que describen la migración de las moléculas desde el seno de la solución a la superficie del cristal y su posterior incorporación desplazando las moléculas de disolvente.

Hay tres causas que provocan que un cristal deje de crecer y sea de baja calidad: I) el consumo de las unidades de crecimiento (se ha alcanzado una situación de equilibrio termodinámico), II) el incremento de la energía de superficie por acumulación de imperfecciones, y III) la incorporación de impurezas sobre las caras del cristal (Feher & Kam 1985; Kam *et al.* 1978). Las impurezas (moléculas de proteína con cierta desnaturalización, fragmentos proteicos u otras proteínas copurificadas, partículas de polvo, etc.) no se incorporan perfectamente, ya que el número de contactos efectivos es menor que para una molécula en estado nativo. Sin embargo, conforme el cristal crece la concentración de impurezas aumenta.

### **2.2.12.2. Preparación de las muestras de cristalización**

Todos estos elementos anteriores se oponen al principio de uniformidad, y por tanto, a la cristalización. Generalmente, se ha establecido que las disoluciones de proteína para cristalización deben cumplir los siguientes requisitos (Ducruix & Giegé 1999; McPherson 1982; McPherson 1999): a) una concentración elevada; b) una

## ***Materiales y métodos***

pureza elevada; y c) homogeneidad y estabilidad en las condiciones de cristalización.

Para obtener soluciones de proteína de elevada concentración hemos utilizado células de ultrafiltración para concentrar las disoluciones hasta  $10 \text{ mg mL}^{-1}$  aproximadamente. La pureza de la solución de proteína debe de ser tan alta como sea posible, al menos del 95% (determinada mediante electroforesis de geles de poliacrilamida y espectrometría de masas, en nuestro caso). Aunque las impurezas pueden impedir la nucleación o limitar el tamaño de los cristales, en algunos casos la presencia aleatoria de pequeñas impurezas podría facilitar la nucleación, y por tanto, la cristalización de una proteína, aunque esto es algo muy difícil de controlar. Por último, la homogeneidad de la proteína es algo muy importante. La heterogeneidad puede aparecer como resultado de la manipulación experimental (grados distintos de oxidación, proteólisis, saturación parcial con ligandos, agregación en grado variable, desnaturalización parcial, deamidación, etc.). Hay que evitar durante la preparación de la proteína someterla a tratamientos que puedan introducir heterogeneidad, como los cambios bruscos de temperatura, y si es posible utilizar una disolución madre de proteína de una misma purificación. No es recomendable usar la liofilización como modo de preservar las proteínas para cristalizar, ya que algunas veces agregan, lo que dificulta enormemente el proceso de solubilización. La congelación y descongelación sucesivas tampoco son recomendables por estas razones.

Otro requisito muy importante debe ser que, tanto las condiciones donde se mantiene la proteína tras su purificación como las de cristalización (pH, temperatura, adición de ligandos y de protectores de grupos  $-\text{SH}$ , etc.), deben preservar al máximo la estabilidad de la proteína.

### **2.2.12.3. Técnicas de cristalización: difusión de vapor y contradifusión en capilares**

La estrategia general para la cristalización de proteínas es reducir la solubilidad de la macromolécula hasta que ésta alcanza el estado de sobresaturación. Existen diversas formas de llevar una disolución de proteína a un

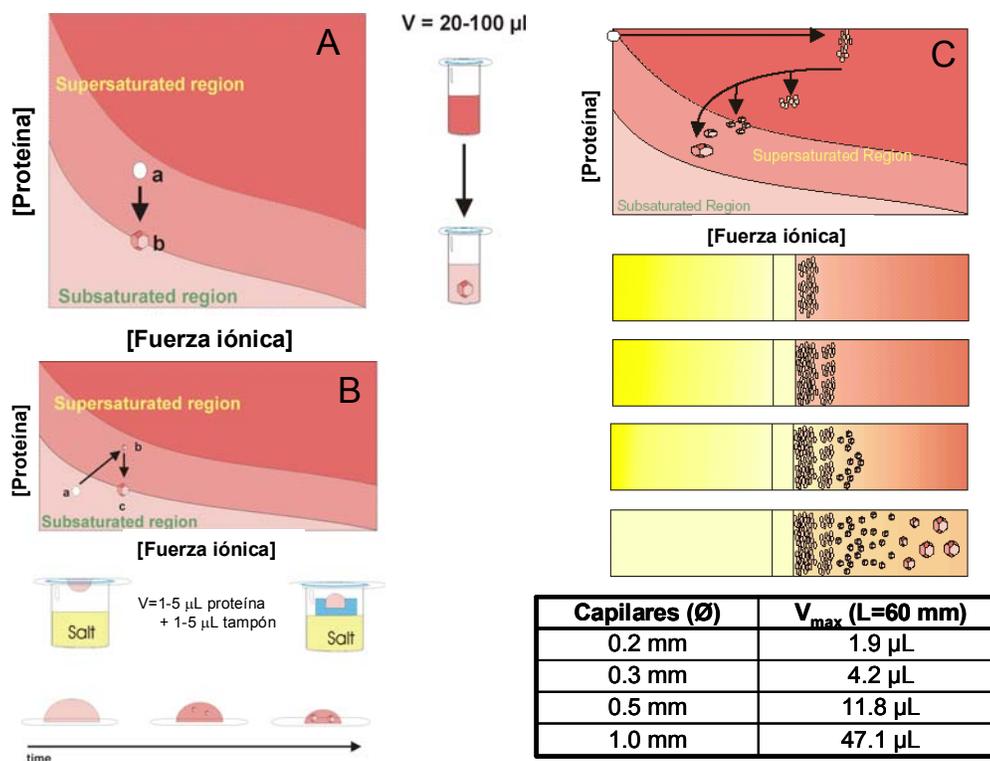
estado sobresaturado (McPherson 1999): por apantallamiento de la carga, por eliminación de moléculas de agua, por disminución de la constante dieléctrica del medio o por descenso de la temperatura. Se han desarrollado diferentes técnicas para llevar una disolución subsaturada de proteína a la región sobresaturada donde la nucleación y el crecimiento cristalino son posibles. Las técnicas más empleadas son la de mezcla directa (*batch*) o la de difusión en fase de vapor.

La cristalización por el método de mezcla directa se consigue al poner en contacto la solución de proteína con un agente precipitante. Como resultado de la adición del agente precipitante, se consigue la sobresaturación en la disolución de proteína (Figura 2.7-A). En la técnica de difusión en fase vapor, la disolución de proteína se mezcla con la disolución de agente precipitante (en una gota sentada o colgante sobre una superficie) para que la mezcla se sitúe en la región sobresaturada del diagrama de solubilidad. En un reservorio cerrado, la gota se enfrenta a una disolución que contiene el agente precipitante (reservorio) a mayor concentración. De acuerdo con la Ley de Raoult, la presión parcial del disolvente es directamente proporcional a su fracción molar en una disolución, de modo que a medida que aumenta la fracción molar del soluto la presión parcial del agua será menor.

Por consiguiente, puesto que la presión parcial del agua en la gota es mayor que en el reservorio ocurrirá una migración de agua en fase vapor hacia el reservorio hasta llegar al equilibrio, lo que provocará un aumento de la concentración de agente precipitante y de proteína en la gota. Este aumento de la concentración, representado como una trayectoria en el diagrama de solubilidad, consigue que la disolución entre en la región sobresaturada en la que pueden tener lugar la nucleación y el crecimiento de los cristales (Figura 2.7-B).

Una nueva técnica que ha venido tomando auge en los últimos años es la contradifusión, basada en la difusión líquido-líquido. Para desarrollarla, la disolución de agente precipitante y la de proteína se ponen en contacto en un medio en el que el transporte de masa está controlado por la difusión. Este requerimiento se satisface en condiciones de gravedad reducida, en medios gelificados o capilares, o en una combinación de éstos (Figura 2.7-C) (García-Ruiz 1991).

## Materiales y métodos



**Figura 2.7.** Representación esquemática de las técnicas de cristalización por difusión en fase vapor (paneles A y B) cristalización por contradifusión en capilares (panel C). Los diagramas de fases muestran las zonas sobresaturadas en un rango de concentraciones de proteína crecientes como una función de la fuerza iónica en la disolución de proteína. Las regiones lábiles y metaestables se representan en rojo oscuro y rojo claro, respectivamente (A-C). La región subsaturada en capilares se representa en amarillo. Durante el proceso de cristalización, la disolución de proteína viaja hacia la región sobresaturada disminuyendo la solubilidad de la proteína. En la región lábil ocurre la nucleación, acto seguido la solubilidad de la disolución de proteína disminuye hasta la región metaestable donde el crecimiento del cristal es más favorable (a**→**b**→**c). En la parte inferior de la figura, se muestran los distintos volúmenes de disolución de proteína necesario para ambas técnicas y las dimensiones de los capilares para contradifusión (Ng et al. 2003).

Otros dos requerimientos necesarios para que esta técnica desarrolle todo su potencial son unas condiciones iniciales muy alejadas del equilibrio (elevada concentración de agente precipitante o de proteína) y una única dimensión de desarrollo. Generalmente, se utiliza un capilar en el que se pone la disolución de proteína y una cámara con el precipitante, en un tampón con unas condiciones de pH y fuerza iónica determinadas. La difusión del agente precipitante (mayor coeficiente de difusión que la proteína) en la cámara de proteína originará un gradiente continuo de sobresaturación hasta alcanzar condiciones de equilibrio dependientes del espacio y del tiempo a lo largo de todo el capilar (Garcia-Ruiz 2003; Ng et al. 2003). La mejor estrategia para reducir la velocidad de

sobresaturación es minimizar la cinética del proceso de transporte de masa, lo que a su vez se consigue minimizando los procesos de convección y permitiendo que el transporte de masa ocurra según las leyes de la difusión. Se ha demostrado que los geles son un medio muy bueno para crecer cristales de alta calidad (en el caso de geles de agarosa, una concentración del 0.12 %p/v es suficiente para eliminar la convección). Además, aunque existen diferentes formas de implementar la contradifusión (Garcia-Ruiz 2003), el empleo de capilares de 0.1 mm de diámetro asegura la minimización de los fenómenos convectivos y reduce significativamente la cantidad de proteína necesaria para llevar a cabo los experimentos (300 nL por cada capilar de 50 mm de longitud).

Para la cristalización de proteínas estudiadas en esta Memoria hemos utilizado técnicas tanto de difusión en fase vapor como de contradifusión. Para las primeras hemos utilizado los reactivos “Cristal Screen HR2-110” y “HR2-112” de Hampton Research para explorar condiciones de cristalización preliminares, o bien hemos partido de condiciones ya descritas en bibliografía. Para las proteínas cristalizadas por contradifusión hemos utilizado unas cajas de cristalización, modelo GCB proporcionadas por Triana Science & Tech, que permiten explorar condiciones de pH desde 3.0 hasta 9.0 en los precipitantes sulfato amónico y formiato sódico. Estos experimentos se han llevado a cabo en capilares de borosilicato de 0.1, 0.2 y 0.3 mm de diámetro y 50 mm de longitud. En el caso de los capilares de 0.2 y 0.3 mm de diámetro hemos gelificado la solución de proteína utilizando una concentración final de 0.05 % de agarosa para evitar fenómenos convectivos en el capilar. Los capilares se han sellado utilizando cera líquida o plastilina.

#### **2.2.12.4. Seguimiento del crecimiento cristalino**

Es muy importante etiquetar las bandejas de cristalización de gota sentada con toda la información posible acerca de las condiciones físico químicas experimentales, identificación de la proteína a cristalizar, y fecha de preparación. Es conveniente preparar un documento para cada bandeja de cristalización que debe contener todas las condiciones experimentales, volúmenes de reservorio, de gota inicial, composición de la gota, temperatura, fecha, etc. En nuestro caso hemos

## ***Materiales y métodos***

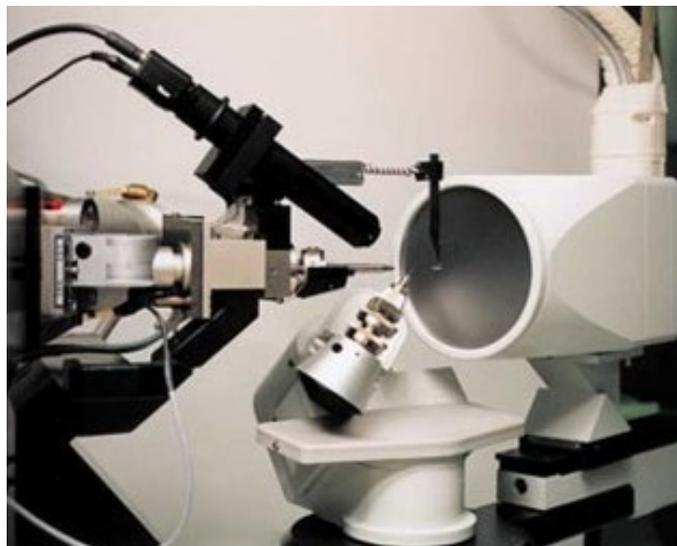
utilizado un documento modelo de seguimiento de crecimiento cristalino para bandejas de 24 pocillos proporcionado por Hampton Research. En el caso de los reactivos "Cristal Screen I" y "II" de Hampton research, existe un documento modelo para controlar la evolución de cada gota en la bandejas de 48 pocillos. Ambos modelos están disponibles en la página web [www.hamptonresearch.com](http://www.hamptonresearch.com) .

Para el control de los capilares se anota la información detallada en un documento que se identifica con cada caja que contiene los capilares (conteniendo hasta un máximo de 6 en nuestro caso) y se colocan en un orden determinado para poder identificarlos, ya que no es posible escribir en ellos.

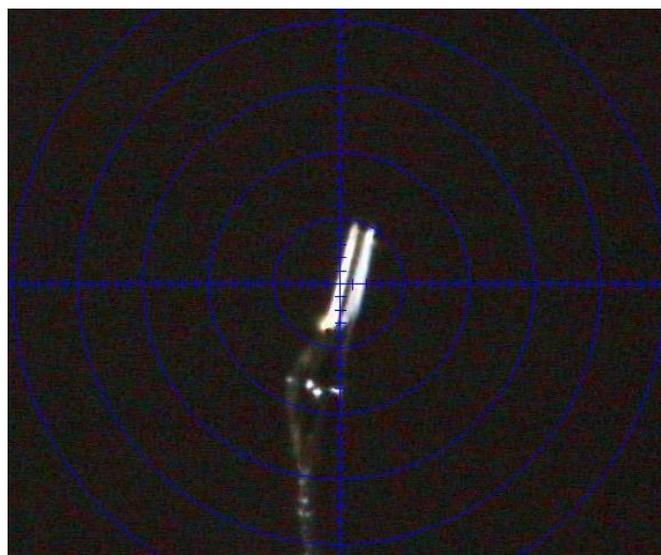
El seguimiento de la evolución espacial y temporal del crecimiento cristalino se ha realizado con un equipo de microscopía óptica de transmisión de Leyca equipado con una cámara digital conectada a un ordenador personal y una lupa binocular de Olympus. De esta forma, podemos conseguir imágenes de cualquier etapa del crecimiento cristalino.

### **2.2.12.5. Metodología utilizada para la difracción de rayos X**

En primer lugar es necesario preparar el cristal que vamos a difractar. En el caso de los cristales obtenidos utilizando la técnica de gota sentada hemos utilizado glicerol a una concentración final del 5% como agente crioprotector, lo que evita la formación de hielo y la consiguiente rotura del cristal. Los cristales se han montado utilizando lazos de Hampton research desde 0.05 hasta 0.6 micras y se han congelado bajo un flujo laminar de nitrógeno gas (Krioflex) a una temperatura de 100 K. Esto evita la generación de calor y radicales libres durante el tiempo de medida que ocasionarían daños irreversibles en el cristal. Para los cristales obtenidos por contradifusión en capilares hemos difractado directamente en los propios capilares a temperatura ambiente. Para ello se corta la parte del capilar que contiene el cristal a difractar hasta dejar una longitud de 10-15 mm y se sellan los extremos con pegamento de secado rápido. Finalmente, se acopla el capilar a una cabeza goniométrica que posee varios grados de libertad (Figura 2.8).



**Figura 2.8.** Difractómetro de rayos X de ánodo rotatorio Microstar de BRUKER Microfocus (Montel Optics) junto con un detector CCD (Charged Coupled Device) Proteum X8.



**Figura 2.9.** Cristal del mutante N114A de Abl-SH3 montado en un lazo para su difracción en el equipo mostrado en la Figura 2.8. El cristal se centra en el haz de rayos X con el programa Proteum2Server instalado en dicho equipo.

Una vez montado el lazo o el capilar, éste se centra con una cámara digital que se encuentra alineada en el plano del haz de rayos X, tal y como se ilustra en la

## ***Materiales y métodos***

Figura 2.9, para lo que se usa un ordenador con el programa Proteum 2Server. Antes de recoger los datos se hace una prueba de difracción en la que se expone el cristal durante 15-120 segundos y se gira a través de uno de los grados de libertad en intervalos de  $0.5^\circ$ . A la imagen recogida se le asignan los correspondientes índices de Miller y se define la celdilla unidad y grupo de Laue (McPherson 1999). El cálculo de la estrategia del experimento de difracción se hace para recoger el máximo de reflexiones posibles en el menor tiempo de exposición. Una vez calculada la estrategia, se importa y se comienza la recogida de reflexiones.

Los cristales se han difractado utilizando un generador de rayos X de ánodo rotatorio Microstar de *BRUKER* Microfocus (Montel Optics), junto con un detector *CCD (Charged Coupled Device) Proteum X8*. El haz de rayos X se genera con una lámpara de ánodo rotatorio de cobre con filtro de níquel que permite seleccionar la radiación  $K\alpha$  del cobre ( $\lambda=1.54 \text{ \AA}$ ). El generador de alimentación del tubo funcionó a 45 kV y 60 mA.

La resolución estructural de las proteínas estudiadas en esta Memoria ha sido realizada por la Dra. Ana Cámara Artigas (Universidad de Almería).

### **2.2.13. Dinámica molecular de proteínas**

En el campo de la química, la dinámica molecular es una técnica de simulación en la que se permite que átomos y moléculas interactúen por un periodo de tiempo bajo las leyes de la física clásica. En realidad, la dinámica molecular es un campo interdisciplinario ya que sus leyes y teorías provienen de las matemáticas, la física y la química, y emplea algoritmos de las ciencias de la computación y teoría de la información. Esta metodología permite estudiar los materiales y las moléculas no como entidades rígidas, sino como cuerpos en movimiento y su campo de aplicación va desde superficies catalíticas hasta sistemas biológicos (McCammon *et al.* 1977; McCammon & Harvey 1997).

Al menos desde la época de Boltzmann (siglo XIX) sabemos que la materia está constituida por partículas en movimiento que interactúan, ya que previamente

se había considerado a las moléculas como los típicos modelos estáticos de museo. Richard Feynman, en 1963, dijo que "todo lo que hacen los seres vivos puede ser entendido a través de movimientos de los átomos." (Feynman 1963). Una de las contribuciones más importantes de la dinámica molecular ha sido crear conciencia de que los sistemas biológicos y sus elementos (el ADN y las proteínas) se componen de partículas en movimiento, de ahí que sea utilizada para explorar la relación entre estructura, movimiento y función. También se la ha llamado "mecánica numérico-estadística" o "la visión de Laplace de la mecánica Newtoniana", en el sentido de predecir el futuro al animar las fuerzas de la naturaleza. Resulta tentador pensar que la dinámica molecular es como un microscopio virtual, sin embargo, las simulaciones largas en tiempo pueden estar limitadas desde un punto de vista matemático, lo cual genera errores numéricos acumulativos. Esto quiere decir que debemos usar con cautela la idea de que estamos siguiendo el comportamiento real de una molécula en el tiempo. De cualquier forma, la dinámica molecular nos permite explorar su comportamiento representativo en tiempo detallado y resolución espacial.

El resultado de una simulación de dinámica molecular son las posiciones  $X$  y velocidades  $V$  de cada átomo de la molécula para cada instante en el tiempo, lo que se llama comúnmente trayectoria. La principal justificación del método de dinámica molecular se basa en que promedios estadísticos de un conjunto son iguales a promedios de tiempo del sistema, lo que se conoce como "hipótesis ergódica". La forma más simple de dinámica molecular se basa en el conjunto microcanónico, es decir, un sistema aislado de volumen fijo y que no intercambia masa ( $M$ ) ni energía ( $E$ ) con el entorno. En este contexto, para un sistema de  $N$  partículas con coordenadas  $X$  y velocidades  $V$ , se pueden plantear las siguientes ecuaciones diferenciales de primer orden:

$$F(X) = -\nabla E(X) = MV(t) \quad (2.79)$$

$$X(t) = V(t) \quad (2.80)$$

La función de energía potencial  $E(X)$  tiene en cuenta las atracciones y repulsiones que sienten los átomos entre sí debido a los enlaces químicos, interacciones electrostáticas, de van der Waals, etc., de las moléculas. A  $E(X)$

## ***Materiales y métodos***

también se le conoce como “campo de fuerza” y es una función de las coordenadas de las partículas  $X$  que normalmente proviene de cálculos basados en mecánica estadística.

La trayectoria de las partículas debe ser discreta en el tiempo lo que significa que, si no queremos acumular error en el cálculo de la posición y la velocidad de la partícula, debemos de elegir un incremento de tiempo entre cada cálculo lo suficientemente pequeño para que no sea mayor que el movimiento vibracional más rápido del sistema (generalmente 1-2 femtosegundos es suficiente). Además, estos errores se pueden minimizar considerablemente eligiendo los parámetros y algoritmos adecuados. Por ejemplo, para el incremento de tiempo, se integra la posición  $X$  y velocidad  $V$  con el método de integración de Verlet (Verlet 1967), por el cual, dadas las posiciones iniciales (p.ej. la estructura de rayos X de una proteína) y las velocidades iniciales (p.ej., las generadas por un campo de fuerzas), es posible calcular todas las posiciones y velocidades en el futuro. El método se basa en la expansión de polinomios de Taylor (Kline 1990), que utilizan la velocidad y coordenadas de una etapa de integración anterior, junto con las de la actual (tiempo= $t$ ), para calcular su velocidad y coordenadas en la siguiente etapa futura (tiempo= $t+\Delta t$ ). Las ecuaciones finales a las que se llega se describen a continuación:

$$x(t + \Delta t) = x(t) + v(t) \cdot \Delta t + \frac{1}{2} a(t) \cdot (\Delta t)^2 \quad (2.81)$$

$$v(t + \Delta t) = v(t) + \frac{a(t) + a(t + \Delta t)}{2} \cdot \Delta t \quad (2.82)$$

donde  $x$ ,  $v$  y  $a$  son la posición, velocidad y aceleración, respectivamente, que dependen del tiempo, y  $\Delta t$  es el incremento de tiempo entre una etapa pasada o futura respecto al tiempo  $t$ .

### **2.2.13.1 Formulación de las simulaciones de dinámica molecular**

El diseño de una simulación de dinámica molecular debe de tener siempre en cuenta la capacidad computacional disponible. El tamaño del sistema (número total de átomos del sistema) y la duración total de la simulación son los factores que

tenemos que controlar para tener un tiempo de simulación razonable. En la literatura científica más reciente encontramos simulaciones con duración total que van desde nano- hasta microsegundos, lo que requiere desde varios días hasta años de tiempo computacional. Durante una simulación de dinámica molecular la evaluación del potencial en función de las coordenadas internas de las partículas que componen el sistema de estudio es la tarea que consume más tiempo computacional. Para la evaluación del potencial, éste suele dividirse como la suma de un término covalente y otro no covalente:

$$E_{total} = E_{covalente} + E_{no\ covalente} \quad (2.83)$$

donde el término covalente puede ser separado en sus contribuciones debidas a la energía entre los dos átomos que forman el enlace covalente ( $E_{enlace}$ ), a la geometría (considerando sus ángulos) de los orbitales electrónicos implicados en el enlace ( $E_{ángulo}$ ), y a la energía debida a la torsión del enlace ( $E_{diedro}$ ), que depende del orden del enlace (por ejemplo, doble enlace) y de los pares electrónicos vecinos a dicho enlace:

$$E_{covalente} = E_{enlace} + E_{ángulo} + E_{diedro} \quad (2.84)$$

El término no covalente puede, a su vez, desdoblarse en sus contribuciones electrostática y de van der Waals:

$$E_{no\ covalente} = E_{electrostático} + E_{vanderWaals} \quad (2.85)$$

Este último término se estudia como un potencial par-par atómico, y es la parte que consume más tiempo computacional. La contribución electrostática, por su parte, es integrada según la ley de Coulomb:

$$E_{ij}(r_{ij}) = \sum \frac{z_i \cdot z_j}{4\pi\epsilon_0} \frac{1}{r_{ij}} + \sum A_l \exp \frac{-r_{ij}}{p_l} \quad (2.86)$$

donde el primer término corresponde con la ecuación de Coulomb para un par de iones  $i$  y  $j$ , siendo  $z$  las cargas de cada ion,  $\epsilon_0$  la constante dieléctrica del medio, y  $r_{ij}$  la distancia entre ellos; y el segundo término, el cual muchas veces se obvia, es la

## **Materiales y métodos**

repulsión de corto alcance justificada por el principio de exclusión de Pauli; siendo  $A_l$  el potencial electromagnético y  $p_l$  el momento.

La contribución al potencial debida a las fuerzas de van der Waals se calcula normalmente mediante el potencial de Lennard-Jones:

$$E(r) = 4\varepsilon' \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^6 \right] \quad (2.87)$$

donde  $\varepsilon'$  es el mínimo del potencial,  $\sigma$  es la distancia a la cual la fuerza entre las partículas interaccionantes es cero, y  $r$  es la distancia entre dichas partículas. Para reducir el coste computacional del cálculo de estos potenciales suele utilizarse el método de Ewald o "Particle mesh Ewald" (PME) (DeLeeuw *et al.* 1980). Este método considera las interacciones de tipo no covalente mediante un potencial para las de corto alcance ( $E_{ca}$ ) y otro para las de largo alcance ( $E_{la}$ ):

$$E_{total} = \sum_{i,j} \Phi(r_j - r_i) = E_{ca} + E_{la} \quad (2.88)$$

donde,  $\Phi$  es el potencial de interacción general, y  $E_{ca}$  y  $E_{la}$  pueden expresarse como:

$$E_{ca} = \sum_{i,j} \Phi_{ca}(r_j - r_i) \quad (2.89)$$

$$E_{la} = \sum_{i,j} \Phi_{la}'(k) |\rho(k)|^2 \quad (2.90)$$

donde  $\Phi_{la}'(k)$  es la transformada de Fourier para el potencial de largo alcance y  $\rho(k)$  es la densidad de carga. Esta nueva definición para el potencial de interacciones no covalentes es especialmente útil, ya que las interacciones de corto alcance son evaluadas en espacio real mientras que las de largo alcance en espacio de Fourier, lo que acelera enormemente su cálculo.

Otra de las aproximaciones utilizadas frecuentemente en simulaciones de dinámica molecular es la aplicación del algoritmo SHAKE (Ryckaert *et al.* 1977), mediante el cual es posible fijar las longitudes de determinados enlaces (por ejemplo, los enlaces de hidrógeno en las moléculas de agua) lo que acelera

muchísimo el cálculo computacional de la simulación. El algoritmo se basa en la aplicación de las siguientes etapas, nombradas I-IV:

I. Se permite el movimiento de todos los átomos del sistema aplicando el método de integración de Verlet (Verlet 1967), asumiendo que no existen enlaces rígidos.

II. Una vez que se ha movido cada átomo, la desviación de cada enlace se usa para calcular la correspondiente fuerza para hacer volver cada átomo y enlace a su posición original.

III. Una vez que se ha aplicado esta corrección a todos los enlaces, se chequea de nuevo cada longitud de enlace. Si la desviación mayor encontrada supera los límites de tolerancia impuestos (generalmente entre  $10^{-4}$  y  $10^{-8}$  Å), entonces se vuelve a realizar otro cálculo para su corrección.

IV. Las etapas II y III son repetidas hasta que la longitud de cada enlace satisface el criterio de convergencia elegido.

Otro factor importante a la hora de diseñar una simulación es la elección del medio en el que se va a llevar a cabo, es decir, se puede hacer el cálculo en vacío o en solución. Además, en el caso de hacerlo en solución, se puede elegir un modelo “explícito” o “implícito” para el disolvente. Un modelo de tipo implícito considera al disolvente como un medio continuo caracterizado por propiedades promedio de lo que sería el disolvente real (Tomasi & Persico 1994). El modelo continuo más riguroso es el que utiliza la ecuación de Poisson-Boltzmann (PB) para calcular el potencial electrostático resultante,  $\Phi_{\text{elect}}(r)$ , a una distancia  $r$  de la densidad de carga  $\rho(r)$ , usando la constante dieléctrica del sistema  $\epsilon'(r)$ :

$$\nabla \cdot \epsilon'(r) \nabla \phi(r) = -4\pi\rho(r) \quad (2.91)$$

Aunque este modelo es más rápido en las simulaciones que un modelo de tipo explícito, sigue siendo muy lento para situaciones donde necesitamos realizar un número de cálculos de dinámica molecular muy elevado, como por ejemplo para

## ***Materiales y métodos***

estudios de cribado de extensas librerías de bases de datos de ligandos para una proteína diana. Por esta razón son necesarios métodos más rápidos.

La siguiente aproximación al método de PB la constituyen el modelo de tipo Born Generalizado (GB) (Still *et al.* 1990) que utiliza el radio efectivo de Born para cada átomo de la molécula soluto y se basa en la teoría electrostática (Tamm 1979); y el modelo de Born Generalizado Superficial (SGB) (Ghosh *et al.* 1998), que utiliza aproximaciones empíricas para introducir efectos de polarización a través de todo el volumen del sistema (Zauhar & Morgan 1985).

Los modelos de tipo explícito ofrecen una descripción mas detallada del disolvente, ya que tienen en cuenta de una forma mas precisa la granularidad y viscosidad del disolvente (Jorgensen & Gao 1988; Kollman 1993; Lybrand *et al.* 1985). Estos modelos emplean cientos o miles de moléculas de disolvente para realizar las simulaciones, por lo que el tiempo de computación es significativamente mayor.

Hay que tener en cuenta que el volumen donde se hacen las simulaciones debe ser suficientemente grande para evitar la aparición de artefactos relacionados con la salida de soluto y de disolvente, lo que se puede evitar mediante la utilización de condiciones periódicas enlazantes (Figura 2.10) (Cheatham *et al.* 1995). Un modelo realista requiere un número bastante grande moléculas de disolvente que deben ser incluidas junto con el soluto. Tanto las partículas de soluto como las de disolvente son encerradas en una caja central la cual se replica de un modo periódico tridimensionalmente. De este modo, las partículas que abandonan la caja por una sus de caras vuelven a entrar por la cara opuesta como si fueran su imagen (Figura 2.10), de manera que el número total de partículas en la caja central permanece constante.

### **2.2.13.2 Simulaciones de dinámica molecular para los complejos del dominio SH3 de Abl, y los mutantes N114A y N114Q con el ligando p41**

Para llevar a cabo las simulaciones de dinámica molecular con los complejos formados por el dominio Abl-SH3 y sus mutantes con el ligando p41 hemos contado



## ***Materiales y métodos***

Para llevar a cabo todas las simulaciones hemos utilizado el paquete de programas de AMBER versión 8.0 (Case *et al.* 2004), siguiendo una serie de etapas en las que, en primer lugar se preparan los ficheros de topología y coordenadas iniciales, y en segundo lugar se llevan a cabo etapas de minimización energética y la propia simulación. Dichas etapas se describen a continuación:

**Preparación de los ficheros de entrada y creación de ficheros de topología y coordenadas iniciales:** hemos utilizado las coordenadas de las estructuras cristalográficas del complejo formado por Abl-SH3, el mutante N114A y el N114Q con p41, que corresponden con los ficheros Pdb 1bbz, 2f0r, para Abl-SH3 y N114A, respectivamente, y pendiente de ser depositado en el banco de datos de proteínas en el caso de N114Q. En el caso de la proteína silvestre hemos dejado enterradas las cinco aguas de la interfase de unión de la estructura a/b del complejo, que se corresponden con las aguas 1015, 1064, 2103, 1082, 1105 del fichero Pdb 1bbz. Para el mutante N114A, hemos dejado las aguas enterradas en la interfase de unión de la estructura a/c del fichero Pdb 2o88, que se corresponden con las aguas 29, 55, 4 y 9, y para el mutante N114Q las aguas enterradas en la interfase de unión de la estructura a/c de su fichero Pdb, que se corresponden con las aguas 48, 14, 1 y 7. En primer lugar se cargan los ficheros Pdb en el módulo xLeap de Amber, donde generaremos los ficheros de topología y de coordenadas iniciales. El fichero de topología contiene información estructural teórica acerca de los átomos y enlaces presentes en cada estructura, los cuales son utilizados por el campo de fuerzas para calcular las coordenadas iniciales de cada átomo en la dinámica molecular. El campo de fuerzas que se aplica para generar el fichero de coordenadas iniciales es el ff99 (Cornell *et al.* 1995), implantado en el módulo xLeap de AMBER, y que está desarrollado para realizar simulaciones de proteínas y ácidos nucleicos en un modelo explícito de disolvente.

**Comprobación de la estructura inicial generada:** es necesario comprobar que los átomos de hidrógeno han sido añadidos correctamente y que no hay ningún problema de conectividad entre los restos que componen la proteína. Para este propósito hemos utilizado el programa Visual Molecular Dynamics (VMD) (Humphrey *et al.* 1996).

**Comprobación de la carga total del sistema:** es necesario comprobar la carga total del sistema con el que vamos a trabajar. Puesto que utilizaremos el campo de fuerzas ff99, que evalúa la energía debida a interacciones electrostáticas, y el método de Ewald para la evaluación de las interacciones electrostáticas de largo y corto alcance, es necesario minimizar la energía electrostática global del sistema. En nuestro caso la carga total del sistema es de -2, por lo que para neutralizarla utilizamos de nuevo el módulo de xLeap añadiendo dos átomos de sodio. Hay que comprobar que los iones han sido añadidos correctamente en la estructura, y que no están cerca del sitio de unión para los ligandos.

**Solvatación del sistema:** hemos llevado a cabo las simulaciones utilizando un modelo de disolvente explícito como es el TIP3P (agua de tipo triangular, en la que el ángulo entre los hidrógenos se mantiene fijo) ya que reproduce adecuadamente las propiedades dinámicas del disolvente y permite estudiar moléculas de agua determinadas de un modo individual (Jorgensen *et al.* 1983), que es uno de nuestros principales propósitos. Hemos utilizado una caja con forma de octaedro truncado de 8 Å para solvatar nuestro sistema, de modo que así se asegura que desde cualquier átomo de soluto a una cara del octaedro hay una distancia mínima de 8 Å. Además, este tipo de caja es la idónea para llevar a cabo simulaciones de dinámica molecular de proteínas globulares en disolución debido a su forma particular, la cual permite ahorrar en número de moléculas de disolvente, y por tanto en tiempo de computación. Posteriormente, hemos comprobado que una caja de 4 Å es suficiente para el propósito de nuestras simulaciones y además reduce el tiempo computacional significativamente.

Una vez que tenemos nuestro sistema neutralizado en carga total, y solvatoado, crearemos los ficheros finales de topología y coordenadas iniciales con el módulo xLeap de AMBER, con los cuales llevaremos a cabo una serie de minimizaciones energéticas antes de la propia simulación.

**Minimización energética inmovilizando el movimiento del complejo:** vamos a utilizar condiciones periódicas enlazantes basándonos en el método de Ewald para la evaluación de las interacciones electrostáticas de largo alcance por debajo de 12 Å (un límite de 10 Å se considera razonable para simulaciones en

## **Materiales y métodos**

modelos explícitos de disolvente). Se realizaron 1000 ciclos de minimización energética. Esta etapa la vamos a llevar a cabo a presión constante, por lo que vamos a permitir variar el volumen de la caja de solvatación para asegurar que las moléculas de agua se muevan más fácilmente y se equilibren alrededor de nuestro soluto, sobre el que aplicamos un pequeño potencial armónico para evitar que ocurran cambios conformacionales inesperados tanto en la proteína como en el ligando (constante armónica  $k = 500 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{Å}^{-2}$ ). Utilizaremos también el algoritmo SHAKE (Ryckaert *et al.* 1977) para impedir el movimiento de los enlaces mediados por hidrógeno.

**Minimización energética del sistema completo sin restricciones:** utilizaremos los mismos algoritmos y parámetros que en la primera etapa, pero ahora sin aplicar un potencial armónico sobre ningún átomo de nuestro sistema. Esta etapa la llevaremos a cabo a presión constante, de modo que tanto los átomos de soluto como de disolvente puedan moverse más fácilmente ya que no aplicamos ningún tipo de restricción sobre el volumen de la caja. En este caso la etapa constará de 2500 ciclos de minimización energética.

**Calentamiento del sistema completo de 0 a 300 K aplicando un potencial pequeño sobre los átomos del complejo:** utilizando los mismos algoritmos y parámetros que en etapas anteriores, y a volumen constante, calentaremos desde 0 a 300 K aplicando un pequeño potencial armónico con una constante armónica de  $10 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{Å}^{-2}$  sobre los átomos de soluto, para evitar que ocurran cambios conformacionales drásticos como consecuencia del calentamiento. Para evitar calentamientos bruscos de temperatura usaremos el termostato de Langevin, mediante el cual se aplica una fuerza friccional que se opone a la fuerza proporcionada por el campo de fuerzas seleccionado para controlar la energía cinética de las partículas y alcanzar la temperatura deseada (Grest & Kremer 1986). La ecuación de Langevin utiliza la siguiente expresión:

$$ma = -\xi v + f + f' \quad (2.92)$$

donde  $m$  es la masa de la partícula,  $a$  es su aceleración,  $v$  es la velocidad de la partícula,  $\xi$  es una constante friccional,  $f(r)$  es la fuerza generada por el campo de

fuerzas seleccionado, y  $f'$  es una fuerza aleatoria. Puesto que la constante friccional tiene un valor positivo, la fuerza friccional proporcionada por el producto  $-\xi v$  reducirá la aceleración de la partícula y provocará una disminución de la temperatura. La fuerza aleatoria  $f'$  es utilizada para variar la energía cinética de las partículas mediante una distribución de tipo gaussiano, que depende de la temperatura y tiempo de integración seleccionado. En nuestro caso, se han llevado a cabo 10000 simulaciones de dinámica molecular de 2 fs cada una para calentar el sistema paulatinamente desde 0 a 300 K. En realidad, esta etapa constituye una primera simulación de dinámica molecular, ya que es aquí donde se generaran las velocidades iniciales mediante una distribución de Boltzman.

**Simulaciones de dinámica molecular de nuestro sistema sin restricciones:** el objetivo es llevar a cabo la simulación a temperatura y presión constante que son las condiciones más similares a las de laboratorio, por lo que utilizaremos una presión de 1 atmósfera que se mantendrá constante en promedio durante el resto de la simulación. Hemos llevado a cabo las simulaciones de dinámica molecular sin aplicar ningún potencial armónico sobre ningún átomo del sistema. Para ello se realizarán 6,000.000 simulaciones de dinámica molecular con un tiempo de integración de 2 femtosegundos, lo que constituye una simulación total de 12 nanosegundos. Se generaran tres ficheros, uno que contiene los resultados promedio de la simulación (energía cinética total, energía potencial total, temperatura promedio, volumen promedio, etc.), otro que contiene la trayectoria propiamente dicha (posición y velocidad de cada átomo del sistema) y un último fichero de reinicio, que sirve para reiniciar la simulación en el caso de que ésta no llegue a completarse totalmente. Cada uno de estos ficheros se actualiza cada 1000 simulaciones completadas. Para completar esta etapa se requiere aproximadamente una semana de cálculo computacional de la red de ordenadores que dispone del grupo de la Dra. Pisabarro, cuyas características técnicas ya hemos descrito.

### **2.2.13.3 Análisis y visualización de trayectorias de dinámica molecular**

Mientras que la trayectoria está siendo calculada es conveniente realizar periódicamente algunas pruebas para comprobar que nuestro sistema es estable.

## ***Materiales y métodos***

Generalmente, estas comprobaciones consisten en analizar algunas de sus propiedades, como las energías cinética, potencial y total, la presión, la temperatura, la densidad, y el volumen de nuestro sistema, además de la desviación cuadrática media (RMSD) de los restos de soluto. La representación de cada una de estas propiedades frente al tiempo de simulación está caracterizada por mostrar valores que tienden a estabilizarse en el tiempo, siempre y cuando la dinámica sea estable. Así, por ejemplo, para la densidad de nuestro sistema debemos obtener un valor ligeramente superior a  $1.00 \text{ g}\cdot\text{cm}^{-3}$  (densidad del agua a 300 K y 1 atm de presión). En el caso de la presión, veremos que al principio tiene un valor constante de cero como consecuencia de la primera etapa de simulación, mientras que después toma valores que fluctúan en torno a 1 atm. Si el termostato de Langevin está funcionando correctamente veremos que la temperatura toma un valor estable de 300 K. Otra prueba importante que nos permite conocer si se mantiene la integridad del complejo y no hay cambios conformacionales significativos es el análisis de RMSD, que suele calcularse sobre los carbonos alfa del complejo, o sobre la proteína y el ligando individualmente.

También es conveniente visualizar la trayectoria de dinámica molecular y ver cómo se comportan la proteína, el péptido y las moléculas de agua, para lo que hemos utilizado el programa Visual Molecular Dynamics (VMD) (Humphrey *et al.* 1996). Previamente es necesario reprocesar las imágenes de dinámica molecular, de modo que se mantenga constante el número de moléculas de disolvente en la caja de solvatación durante la trayectoria. Este proceso es posible gracias a que estamos utilizando condiciones periódicas enlazantes, de modo que cuando una molécula abandona la caja, su molécula imagen entrará a ocupar una posición equivalente. Para esta operación se utiliza el comando “reimage” del módulo ptraj de AMBER 8.0.

Una vez hechas estas correcciones tenemos las trayectorias preparadas para analizarlas. Para ello, hemos seguido una estrategia basándonos en las etapas que se describen a continuación:

**Cálculo de la densidad de hidratación en la interfase de unión de los complejos:** para comprobar que las posiciones de las moléculas de agua

observadas en las estructuras cristalográficas son ocupadas durante la simulación e identificar nuevos sitios de hidratación hemos llevado a cabo el cálculo de la densidad de hidratación en la interfase de unión de los complejos. Hemos utilizado el comando “grid” del modulo ptraj de AMBER, que permite generar un fichero con la densidad de hidratación que se puede visualizar en VMD. El cálculo consiste en generar un cubo tridimensional (de 50 Å de lado), que se divide a su vez en cubos de menor tamaño que se diferencian entre sí 0.5 Å de lado, y que explora una zona específica de la caja de la trayectoria. Cada vez que una cara de un cubo identifique una molécula de agua, esta será contabilizada para el cálculo de la densidad de hidratación.

**Identificación de las moléculas de agua y cálculo de su tiempo de residencia en una posición de la interfase de unión:** en primer lugar hemos creado un fichero Pdb donde se identifica cada átomo del complejo y cada una de las moléculas de agua en la caja de solvatación, para lo que se utiliza el comando “ampdb” de AMBER. El criterio para considerar una molécula de agua enterrada en el sitio de unión ha sido la formación de al menos uno de los enlaces de hidrógeno que establece en las estructuras cristalográficas. Hemos permitido que este criterio sea infringido por períodos iguales o inferiores a 2 picosegundos, que es el tiempo entre cada imagen de la dinámica molecular. Visualizaremos la trayectoria de dinámica molecular con el programa VMD, seleccionando las aguas que están formando enlaces de hidrógeno con los restos deseados a una distancia igual o menor de 3.5 Å, y un ángulo mayor de 90 grados entre  $H \text{-----} X \text{-----} H$  , donde X se considera el átomo aceptor del enlace de hidrógeno. Para el cálculo del tiempo de residencia de una molécula de agua, ocupando una posición cristalográfica hemos contabilizado el número de imágenes que cumplen el criterio definido anteriormente. A su vez, el tiempo de residencia promedio,  $\langle t_r \rangle$  de una posición se obtiene aplicando un procedimiento estadístico que pondera el valor del tiempo de residencia de cada molécula de agua por el número de imágenes que permanece en la interfase de unión:

$$\langle t_r \rangle = \frac{\sum t_r \times n_i}{n_T} \quad (2.93)$$

## **Materiales y métodos**

donde  $t_r$  es el tiempo de residencia de una molécula de agua determinada,  $n_T$  es el número total de imágenes de la simulación y  $n_i$  es el número de imágenes que una molécula de agua permanece en una posición de la interfase de unión. Además, hemos calculado la ocupancia de cada una de las posiciones de las moléculas de agua, lo que corresponde con el porcentaje de tiempo de la simulación que una posición determinada esta ocupada por moléculas de agua, y es igual al cociente del sumatorio del tiempo de residencia de cada molécula de agua individual entre el tiempo total de simulación.

**Análisis de las los puentes de hidrógeno establecidos en la interfase de unión:** una vez que hemos identificado las moléculas de agua que ocupan cada una posiciones en el sitio de unión, hemos analizado las interacciones que establece. Hemos estudiado tanto la distancia, ángulo y frecuencia de formación de los puentes de hidrógeno que forman las moléculas de agua utilizando el comando “*hbond*” de AMBER. También hemos estudiado los puentes de hidrógeno, tanto inter- como intramoleculares, formados por el ligando y la proteína en la interfase de unión. Con este comando es posible definir los átomos de la proteína, del ligando y de las moléculas de agua como posibles donores o aceptores de puentes de hidrógeno durante toda la simulación. En el caso de las moléculas de agua, hemos definido como donador y aceptor los átomos de hidrógeno y de oxígeno respectivamente de cada molécula de agua ocupando una posición cristalográfica, y hemos calculado todas las interacciones que establece a una distancia menor o igual a 3.5 Å y un ángulo mayor o igual a 90 grados formado entre los átomos pesados del puente de hidrógeno. Se consideró una frecuencia de formación mayor del 5 % durante la dinámica. Para el cálculo de las interacciones mediadas por ligando y proteína, hemos seleccionado los átomos de las cadenas laterales y los carbonilo y nitrógeno del enlace peptídico para los restos que establecen enlaces de hidrógeno en las estructuras cristalográficas, así como de todos los restos del ligando p41. Para el cálculo del ángulo promedio,  $\langle \text{ángulo} \rangle$ , de los puentes de hidrógeno establecidos por las moléculas de agua que ocupan una determinada posición en la interfase de unión hemos aplicado un procedimiento estadístico que pondera su valor por el tiempo de residencia de cada molécula de agua del siguiente modo:

$$\langle \text{ángulo} \rangle = \frac{\sum (\text{ángulo} \times t_r)}{t_T} \quad (2.94)$$

donde se aplica el sumatorio del producto del valor del ángulo por el tiempo de residencia de cada molécula de agua,  $t_r$ , y se divide por el tiempo total de simulación,  $t_T$ . Para la distancia promedio y la frecuencia de formación de cada puente de hidrógeno se utiliza un procedimiento similar.

**Cálculo de la fluctuación atómica y del factor B teórico de las moléculas de agua en la interfase de unión:** para calcular el valor de fluctuación atómica hemos utilizado el comando “atomicfluct” de AMBER teniendo en cuenta sólo la parte de la trayectoria que una determinada molécula de agua se encuentra en una posición original de la estructura cristalográfica. Posteriormente, el cálculo de factor B se puede obtener a través de la siguiente ecuación:

$$factorB = \left( \frac{8 \times \pi^2}{3} \right) \times \frac{(ri - \langle ri \rangle)^2}{n} \quad (2.95)$$

donde  $ri$  es la posición de la molécula de agua en cada momento,  $\langle ri \rangle$  es la posición promedio de esa molécula de agua durante  $n$  imágenes, que son las que esa molécula permanece en la posición original en la interfase de unión. Igualmente, para el cálculo del *factor B* promedio de las moléculas de agua que ocupan una posición determinada se aplica un procedimiento estadístico similar al utilizado para el tiempo de residencia promedio.



## **RESULTADOS**



### **3.1. ESTUDIO DEL RECONOCIMIENTO DE LIGANDOS RICOS EN PROLINA POR EL DOMINIO SH3 DE ABL**

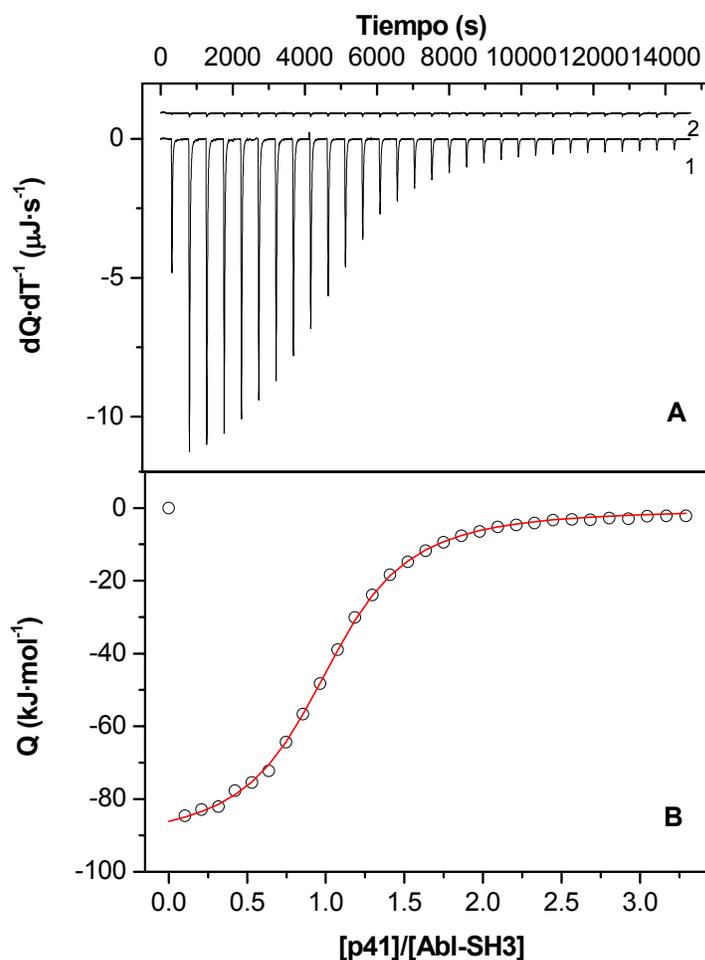
#### **3.1.1. Análisis termodinámico-estructural de la interacción de Abl-SH3 con el ligando de diseño p41**

##### **3.1.1.1. La unión Abl-SH3 con p41 presenta un patrón termodinámico incongruente con la elevada hidrofobicidad de la interfase de unión**

Hemos abordado la caracterización termodinámica mediante calorimetría isotérmica de titulación de la interacción entre el dominio SH3 de la tirosina quinasa del oncogén Abl y el ligando p41, de secuencia APSYSPPPPP. Este péptido, obtenido mediante métodos de diseño racional a partir del ligando natural 3BP1 (APTMPPLPP) (Pisabarro *et al.* 1994; Pisabarro & Serrano 1996), es el ligando de mayor afinidad descrito hasta la fecha para el dominio Abl-SH3. En la Figura 3.1 se muestra como ejemplo la titulación calorimétrica de p41 y Abl-SH3 a 25 °C en tampón fosfato 20 mM pH 7.0 (panel superior) junto con la curva isoterma de unión y el mejor ajuste al modelo de unión de “n” sitios idénticos e independientes, de acuerdo a la ecuación 2.21 (panel inferior). Según este análisis, el ligando p41 se une a un único sitio de unión en el dominio Abl-SH3 con una constante de disociación 2.3  $\mu\text{M}$ , lo que concuerda con los valores obtenidos en trabajos anteriores mediante espectroscopia de fluorescencia (Pisabarro & Serrano 1996). La interacción está dirigida por una entalpía de unión muy exotérmica ( $\Delta H_{ap} = -91.7 \text{ kJ}\cdot\text{mol}^{-1}$ ) parcialmente compensada por una contribución entrópica desfavorable ( $-T\Delta S_{ap} = 59.4 \text{ kJ}\cdot\text{mol}^{-1}$ ). El cambio de capacidad calorífica ( $\Delta C_{pap} = 0.73 \text{ kJ}\cdot(\text{K}\cdot\text{mol})^{-1}$ ) se obtuvo del análisis de la dependencia con la temperatura de la entalpía de unión bajo las mismas condiciones de tampón y pH.

La característica más sobresaliente de los parámetros termodinámicos es el carácter fuertemente exotérmico de la entalpía de unión, que contrasta notablemente con la elevada hidrofobicidad de las superficies de interacción tanto en la proteína como en el ligando. Como se ha comentado anteriormente, y se ilustró en la Figura 1.2 (página 18), desde un punto de vista estructural la interacción se basa fundamentalmente en la intercalación de restos hidrofóbicos, de modo que las

## Resultados



**Figura 3.1.** Experimento de titulación del dominio Abl-SH3 con el péptido p41 mediante ITC en fosfato 20 mM, pH 7.0, a 25 °C. **A.** (1) Termograma de titulación formado por los calores por unidad de tiempo liberados tras la inyección de p41 (30 inyecciones de 7  $\mu\text{L}$  de una disolución de p41 a 0.6 mM) sobre la proteína Abl-SH3 (32  $\mu\text{M}$  en la célula calorimétrica). (2) Termograma de dilución del ligando en las mismas condiciones que el experimento de titulación (desplazados en el eje para mayor claridad). **B.** Isotherma de unión correspondiente al experimento en A, obtenida después de normalizar los calores de titulación y corregir por los calores de dilución. Los datos experimentales se representan como círculos, mientras que la línea continua corresponde al mejor ajuste para el modelo de “ $n$ ” sitios de unión idénticos e independientes (ecuación 2.21).

únicas interacciones polares serían dos puentes de hidrógeno establecidos entre Abl-SH3 y p41: *i)* un puente de hidrógeno, altamente conservado en otros complejos SH3-péptido, establecido entre el resto Tyr115 en el dominio y el oxígeno carbonilo de P8 en el péptido; *ii)* un puente de hidrógeno, característico del ligando p41, establecido entre el residuo Y4 de p41 y los restos Ser75 y Asp77 en el lazo RT del dominio Abl-SH3.

Este mismo comportamiento termodinámico se ha descrito para otros dominios SH3 en trabajos previos, los cuales se recogen en la Tabla 3.1 (Arold *et al.* 1998; Ferreon & Hilser 2004; Renzoni *et al.* 1996; Wang *et al.* 2001). Sin embargo, aunque la afinidad de la interacción es de magnitud similar en todos los casos, las contribuciones entálpicas y entrópicas observadas para la interacción entre Abl-SH3 y p41 son considerablemente mayores que las descritas para otros dominios SH3, que oscilan entre  $-33.5$  y  $-63.8$   $\text{kJ}\cdot\text{mol}^{-1}$  para la entalpía de unión, y  $8.4$  y  $27.3$   $\text{kJ}\cdot\text{mol}^{-1}$  para las respectivas contribuciones entrópicas (Arold *et al.* 1998; Ferreon & Hilser 2004). Esto es sorprendente teniendo en cuenta la ausencia especialmente notoria de interacciones polares e iónicas en el complejo Abl-SH3/p41 en comparación con los otros ejemplos.

La unión de Abl-SH3 y p41 conlleva el enterramiento de un total de  $700 \pm 10$   $\text{Å}^2$  de área apolar accesible al disolvente y  $250 \pm 10$   $\text{Å}^2$  de área polar, de los cuales  $440 \pm 5$   $\text{Å}^2$  de área apolar y  $90 \pm 5$   $\text{Å}^2$  de área polar corresponden al ligando y  $240 \pm 10$   $\text{Å}^2$  de área apolar y  $160 \pm 5$   $\text{Å}^2$  de área polar provienen de la ocultación al disolvente de grupos del dominio SH3. Estos valores representan la media aritmética (y la correspondiente desviación estándar) obtenida de las cuatro estructuras presentes en la unidad asimétrica del complejo Abl-SH3/p41 (Pdb: 1bbz). Con estos resultados se obtiene una relación  $\Delta\text{ASA}_{\text{ap}}/\Delta\text{ASA}_{\text{pol}}$  de 2.81, característica de interacciones predominantemente hidrofóbicas para las que, como ya hemos discutido en la sección 1.1, sería de esperar un comportamiento termodinámico diferente (interacciones dirigidas por contribuciones entrópicas favorables asociadas con el efecto hidrofóbico y contribuciones entálpicas desfavorables o sólo ligeramente favorables).

Hemos aplicado la parametrización estructural de la energética de plegamiento de proteínas e interacción con ligandos, desarrollada por el Dr. Freire (Universidad Johns Hopkins, Baltimore, USA) y colaboradores (Luque & Freire 1998; Luque & Freire 2002), a las cuatro estructuras del complejo Abl-SH3/p41 presentes en la unidad asimétrica. Aunque estas ecuaciones predicen adecuadamente el carácter exotérmico de la interacción, las entalpías de unión calculadas a  $25$   $^{\circ}\text{C}$  ( $-10.1$ ,  $-9.4$ ,  $-10.8$  y  $-13.1$   $\text{kJ}\cdot\text{mol}^{-1}$  para los complejos a/b, c/d, e/f y g/h, respectivamente) son notablemente menos negativas que el valor determinado

## Resultados

*Tabla 3.1. Parámetros termodinámicos obtenidos para la unión de ligandos/proteínas con secuencias ricas en prolina a dominios tipo SH3.*

Referencia	Ligando	Proteína	T (°C)	$K_a$ ( $\times 10^4 M^{-1}$ )	$\Delta G_{ap}$ ( $kJ \cdot mol^{-1}$ )	$\Delta H_{ap}$ ( $kJ \cdot mol^{-1}$ )	$-T\Delta S_{ap}$ ( $kJ \cdot mol^{-1}$ )
(Arold <i>et al.</i> 1998)	Nef-SIDA	Hck-SH3	25	166.6	-35.5	-62.8	27.3
(Arold <i>et al.</i> 1998)	Nef-SIDA	Src-SH3	25	8.7	-28.2	-41.8	13.6
(Renzoni <i>et al.</i> 1996)	P2L	Fyn-SH3	30	6.3	-27.8	-51.4	23.6
(Ferreon & Hilser 2004)	SosY	Sem-SH3	15	3.6	-25.1	-33.5	8.4
(Wang <i>et al.</i> 2001)	RLP2	Src-SH3	25	5.3	-26.8	-42.7	15.9
(Palencia <i>et al.</i> 2004)	P41	Abl-SH3	25	4.8	-32.2	-91.7	59.5

experimentalmente ( $-92 \text{ kJ} \cdot \text{mol}^{-1}$ ). Es interesante señalar que estas ecuaciones empíricas reflejan exclusivamente las interacciones directas entre las interfases de unión del ligando y la proteína. De igual modo, debemos considerar que el error promedio con el que estas ecuaciones predicen generalmente las entalpías de unión es del orden de  $5 \text{ kJ} \cdot \text{mol}^{-1}$  (Luque & Freire 2002), lo que indicaría que existen contribuciones importantes a los parámetros termodinámicos cuyo origen no está incluido en estas ecuaciones. De igual manera, el cambio de capacidad calorífica de unión ( $1.04 \pm 0.02 \text{ kJ} \cdot (\text{K} \cdot \text{mol})^{-1}$ , calculado como promedio entre las cuatro estructuras) presenta una discrepancia con el valor experimental ( $0.73 \text{ kJ} \cdot (\text{K} \cdot \text{mol})^{-1}$ ) mayor que el error del 9-10% con el que estas ecuaciones generalmente predicen las capacidades caloríficas. En todo caso, ambos valores estarían dentro de lo esperado para el tamaño y naturaleza de nuestro sistema (Jelesarov & Bosshard 1999; Makhatadze & Privalov 1995) y son muy similares a los obtenidos para otros complejos SH3-péptido (Ferreon & Hilser 2004).

En definitiva, los resultados de este análisis indican claramente que el comportamiento termodinámico observado no puede explicarse adecuadamente en términos de interacciones directas entre interfases de unión hidrofóbicas y apuntan

hacia una mayor complejidad, de modo que deben existir factores adicionales que confieran a una interacción aparentemente muy hidrofóbica el patrón termodinámico característico de interacciones más polares. Estos factores podrían ser: *i*) la existencia de equilibrios acoplados por la ionización de grupos del dominio o del ligando (Baker & Murphy 1996; Gomez & Freire 1995; Mason & Jensen 2007; Velazquez-Campoy *et al.* 2000a); *ii*) la existencia de algún cambio conformacional en la proteína y/o el ligando (Arold *et al.* 1998; Ferreon & Hilser 2004; Luque & Freire 2002); *iii*) la desolvatación incompleta de las superficies de interacción (Bhat *et al.* 1994; Li & Lazaridis 2007; Lu *et al.* 2007; Petrone & Garcia 2004; T. M. Raschke 2006). De este modo, podríamos expresar la entalpía aparente de unión ( $\Delta H_{ap}$ ) obtenida experimentalmente como:

$$\Delta H_{ap} = \Delta H_{int} + \Delta H_{conf} + n_H \times \Delta H_{ion} \quad (3.1)$$

donde  $\Delta H_{int}$  es la entalpía intrínseca de unión, es decir, debida exclusivamente a la interacción establecida entre el ligando y la proteína, que es dependiente del estado de protonación de proteína y ligando;  $\Delta H_{conf}$  es la entalpía asociada a cambios conformacionales en proteína y/o ligando;  $\Delta H_{ion}$  es la entalpía de ionización del tampón; y  $n_H$  es el número de protones aceptados o liberados durante el proceso de unión.

### **3.1.1.2. La termodinámica de unión de Abl-SH3 con p41 no refleja la influencia de equilibrios de ionización acoplados**

Es frecuente que como consecuencia del proceso de unión se produzca un cambio de  $pK_a$  de alguno de los grupos ionizables de las moléculas que interaccionan (Baker & Murphy 1996; Gomez & Freire 1995; Velazquez-Campoy *et al.* 2000a). Como hemos comentado, la consecuencia sería que la entalpía aparente determinada experimentalmente ( $\Delta H_{ap}$ ) podría tener un valor diferente de la entalpía intrínseca de unión ( $\Delta H_{int}$ ). La existencia e influencia de estos equilibrios de ionización se puede evaluar llevando a cabo experimentos a un determinado valor de pH pero en varios tampones con distintos calores de ionización, lo que permitiría estimar el número de protones intercambiados y el valor de la entalpía intrínseca de unión (dependiente del pH pero independiente de la ionización del tampón). Con

## Resultados

este propósito hemos hecho experimentos en tres tampones con diferentes calores de ionización (fosfato, HEPES e imidazol) a pH 7.0 y a tres temperaturas (15, 20 y 30 °C), lo que además nos permitirá obtener el cambio de capacidad calorífica intrínseco de la interacción ( $\Delta C_{p,int}$ ). Los valores de los parámetros termodinámicos obtenidos de este análisis se recogen en la Tabla 3.2.

En la Figura 3.2 se representa la entalpía aparente obtenida en cada uno de los tampones frente a la entalpía de ionización propia de cada tampón para todas las temperaturas estudiadas. El análisis mediante regresión lineal de estos datos proporcionaría directamente de la ordenada en el origen la entalpía intrínseca de unión ( $\Delta H_{int}$ ) a cada temperatura si no existieran cambios conformacionales acoplados ( $\Delta H_{conf} = 0$ ; ecuación 3.1) y el número de protones intercambiados con el tampón por nuestro sistema,  $n_H$ , a partir de la pendiente (ecuación 3.1). Como se refleja en la Figura 3.2 y en la Tabla 3.2 el número de protones intercambiados por la proteína y el ligando con el tampón es nulo, dentro del error experimental y a cualquier valor de temperatura, mientras que los valores de entalpías intrínsecas obtenidas son muy cercanos a las entalpías aparentes medidas en fosfato sódico, que es el tampón con menor calor de ionización. De igual modo, si representamos las entalpías intrínsecas de unión frente a la temperatura obtenemos un valor de  $-0.724 \text{ kJ}\cdot(\text{K}\cdot\text{mol})^{-1}$  para la capacidad calorífica intrínseca de la interacción ( $\Delta C_{p,int}$ ), prácticamente idéntico al de capacidad calorífica aparente obtenido en tampón fosfato ( $-0.73 \text{ kJ}\cdot(\text{K}\cdot\text{mol})^{-1}$ ).

Por consiguiente, a partir de estos resultados podemos afirmar con seguridad que no existe un intercambio neto de protones con el tampón asociado a la formación del complejo. Sin embargo, siempre existe la posibilidad de que varios procesos de ionización de distinto signo se cancelaran entre sí de modo que, aunque no hubiera una contribución entálpica por parte de la ionización del tampón, sí podría haber efectos considerables asociados a la protonación o desprotonación de los grupos en la proteína o el ligando. Esta situación podría detectarse estudiando el efecto del pH sobre la entalpía intrínseca de unión (Baker & Murphy 1996; Luque & Freire 2002; Velazquez-Campoy *et al.* 2000a).

Tabla 3.2. Análisis termodinámico de la interacción entre p41 (APSYSPPPPP) y el dominio Abl-SH3.

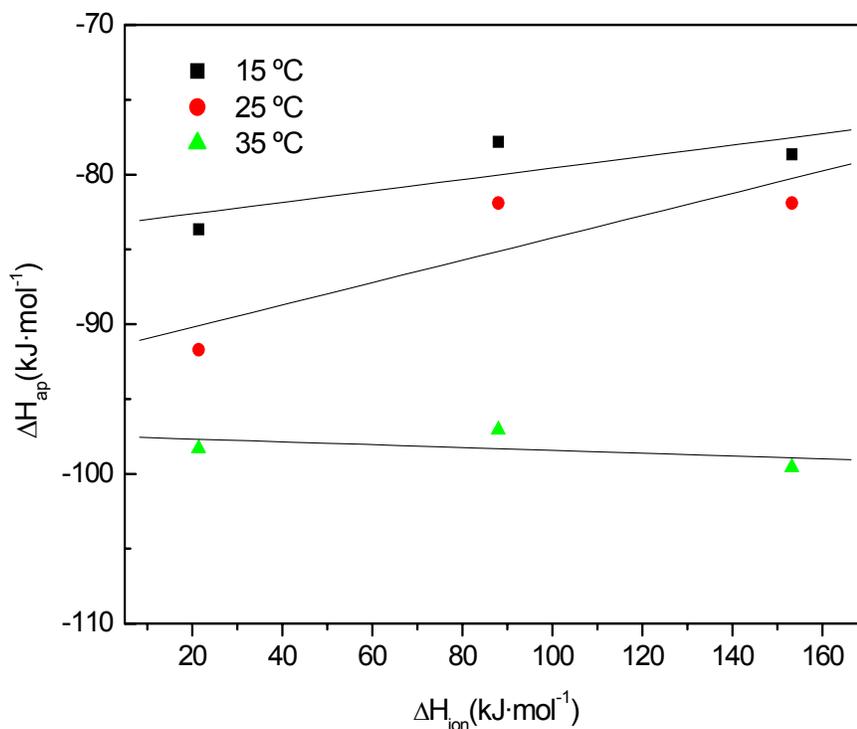
pH	T (°C)	Tampón	n*	$\Delta H_{ap}$ <sup>*</sup> (kJ·mol <sup>-1</sup> )	$K_a$ <sup>*</sup> (x10 <sup>5</sup> M <sup>-1</sup> )	$\Delta C_{p,ap}$ <sup>*</sup> J·(K·mol) <sup>-1</sup>	n <sub>H</sub> <sup>a</sup>	$\Delta H_{int}$ <sup>a</sup> (kJ·mol <sup>-1</sup> )	$\Delta C_{p,int}$ <sup>a</sup> J·(K·mol) <sup>-1</sup>
15		fosfato	0.99	-83.7	12.2				
		HEPES	1.04	-77.9	9.2		0.1	-80.8	
		imidazol	1.04	-78.7	9.9				
7.0	25	fosfato	1.05	-91.7	4.8	-729			
		HEPES	1.01	-87.9	3.9	-955	0.1	-91.7	-724
		imidazol	1.02	-87.9	4.3	-1034			
35		fosfato	1.05	-98.4	1.5				
		HEPES	0.98	-97.1	1.3		0.0	-98.0	
		imidazol	1.03	-99.7	1.1				
5.0	25	acetato	0.98	-92.1	2.1				
3.0	25	glicina	1.03	-93.4	2.9				

\*El error en los datos experimentales, obtenidos del ajuste por regresión no lineal mediante la ecuación 3.21, está en torno al 1% para el número de sitios de unión, 5% para la entalpía de unión y 10 % para la constante de afinidad. <sup>a</sup>Estos valores se obtuvieron considerando los valores de ionización publicados (Christensen 1976; Fukada & Takahashi 1998) para fosfato sódico ( $\Delta H_{ion}=5128$  J mol<sup>-1</sup>,  $\Delta C_{p,ion}=-187.2$  J·(K mol)<sup>-1</sup>), HEPES ( $\Delta H_{ion}=21044$  J mol<sup>-1</sup>,  $\Delta C_{p,ion}=49.0$  J·(K mol)<sup>-1</sup>) e imidazol ( $\Delta H_{ion}=36650$  J mol<sup>-1</sup>,  $\Delta C_{p,ion}=-16.3$  J·(K mol)<sup>-1</sup>).

Con este fin, hemos llevado a cabo titulaciones a pH 5.0 y pH 3.0, utilizando tampones (acetato sódico 20 mM y glicina 20 mM respectivamente) caracterizados por calores de ionización muy bajos (Christensen 1976), lo que asegura que las entalpías aparentes medidas en estas condiciones son muy próximas a las entalpías intrínsecas a estos valores de pH. Como puede apreciarse en la Tabla 3.2, dentro del error experimental no se observan cambios significativos en ninguno de los parámetros termodinámicos de la interacción a los distintos valores de pH, lo que descartaría definitivamente la presencia de procesos de ionización acoplados a la unión del ligando p41 a Abl-SH3.

En definitiva, este análisis descarta totalmente la posibilidad de que la naturaleza fuertemente exotérmica de esta interacción se deba a la influencia de equilibrios de ionización acoplados a la formación del complejo.

## Resultados



**Figura 3.2.** Representación de la entalpía aparente de unión obtenida para cada tampón a 15, 25 y 35 °C, frente a la entalpía de ionización de cada tampón. Si extrapolamos a calor de ionización cero obtenemos la entalpía intrínseca de unión de p41 y Abl-SH3; y de la pendiente de dicha recta obtenemos el número de protones intercambiados con el tampón. Dichos valores aparecen recogidos en la Tabla 3.1.

### 3.1.1.3. No existen cambios conformacionales significativos asociados a la unión de Abl-SH3 y p41

Según lo dicho (ecuación 3.1), también podría ocurrir que, bien la proteína, bien el ligando, o ambos, no adopten la misma conformación espacial cuando se encuentran libres que cuando están formando el complejo, de modo que al unirse tuviera lugar un cambio conformacional en sus estructuras que, lógicamente, podría tener consecuencias en la energética de unión (Luque *et al.* 1998; Velazquez-Campoy *et al.* 2004).

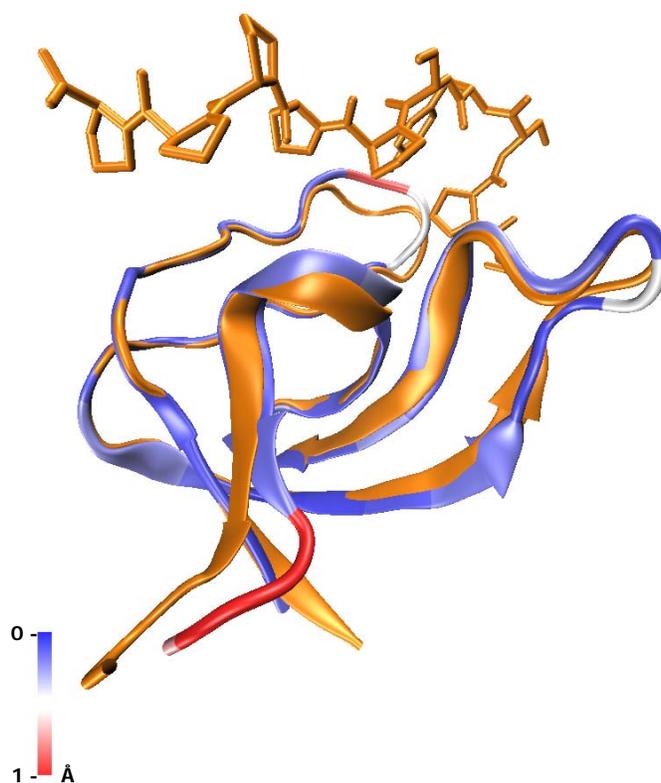
En la actualidad se dispone de estructuras cristalográficas de calidad tanto para el dominio Abl-SH3 libre (Pdb: 1abq) (Musacchio *et al.* 1994) como para el complejo Abl-SH3/p41 (Pdb: 1bbz) (Pisabarro *et al.* 1998) que contiene cuatro estructuras diferentes en la unidad asimétrica. La comparación de la estructura del

dominio libre con cada una de las 4 estructuras del complejo que existen en la unidad asimétrica parece indicar que no existen cambios conformacionales significativos asociados a la interacción (Figura 3.3). De hecho, los valores de desviación cuadrática media (RMSD) son en todos los casos muy pequeños y oscilan entre 0.65 y 0.96 Å para las distintas estructuras, con un valor promedio de 0.75 Å. Sólo se observan ligeras diferencias en los extremos de los lazos RT y n-Src, del mismo orden que las descritas para otros dominios SH3 y sus complejos (Arold *et al.* 1998; Renzoni *et al.* 1996; Wang *et al.* 2001). De hecho, estas regiones son las que muestran mayor flexibilidad en las estructuras lo que se ha propuesto como un factor importante en la determinación de la especificidad de unión en las interacciones de dominios SH3 con sus ligandos (Arold *et al.* 1998; Ladbury & Arold 2000). Aunque la rigidificación de estas zonas flexibles en el complejo pudiera tener consecuencias energéticas, no se observan diferencias significativas entre Abl y otros dominios SH3 que pudieran justificar las diferencias en las magnitudes de las contribuciones entálpicas y entrópicas observadas.

En un estudio reciente llevado a cabo con la proteína Sem-SH3 y un ligando rico en prolina se ha propuesto que el equilibrio conformacional del ligando parece ser uno de los factores determinantes del patrón termodinámico anómalo observado en las interacciones de dominios SH3 (Ferreon & Hilser 2004). Los autores estiman que la organización del ligando desde las conformaciones más desplegadas hacia aquellas con la disposición adecuada para unirse a la proteína (hélice de PPII) contribuye favorablemente al valor de entalpía de unión con unos 10 kJ·mol<sup>-1</sup>, suponiendo que el ligando está completamente desplegado en disolución. En nuestro caso, se ha comprobado experimentalmente que el equilibrio conformacional del ligando p41 libre está muy desplazado hacia la conformación PPII (Pisabarro *et al.* 1998), lo que minimizaría estas contribuciones. De acuerdo con esto, aunque existiera una hipotética redistribución conformacional, sería de esperar que los efectos sobre la entalpía de unión fueran muy pequeños comparados con el valor de entalpía de -91.7 kJ·mol<sup>-1</sup> obtenido para la interacción Abl-SH3/p41.

Por tanto, teniendo en cuenta la información estructural disponible no podemos concluir que existan cambios conformacionales significativos acoplados a la unión del ligando a la proteína que pudieran justificar los valores de los

## Resultados



**Figura 3.3.** Superposición de las estructura del complejo formado entre el péptido p41 y Abl-SH3 (en estructura de varillas y en estructura tipo cinta, respectivamente, ambos en color naranja), y la estructura de Abl-SH3 libre (estructura tipo cinta sobre la que se muestra el valor de RMSD respecto al complejo). Como puede apreciarse las dos estructuras son casi totalmente superponibles. El valor de RMSD entre las dos estructuras se muestra en escala gradual de colores desde al azul (RMSD=0 Å) hasta el rojo (RMSD=1 Å). Los regiones correspondientes a los lazos n-Src y -RT son las que muestran las mayores diferencias, aunque estas no superan el valor de 0.9 Å.

parámetros termodinámicos observados. Los ligeros cambios detectados, así como la posible contribución del equilibrio conformacional del ligando, tampoco parecen suficientes *per se* como para poder justificar la elevada exotermicidad mostrada por la interacción.

### 3.1.1.4. La interacción entre Abl-SH3 y p41 está mediada por un conjunto de moléculas de agua ocluidas en la interfase de unión

Una vez evaluados los posibles equilibrios acoplados al proceso de unión, podríamos decir que el patrón termodinámico descrito para la interacción Abl-

SH3/p41 se debería a causas intrínsecas a la propia interacción. En este sentido, la unión de un ligando a una proteína es un proceso que tiene lugar en disolución y es frecuente que cuando ambos interaccionan la desolvatación de la interfase de unión no sea completa. De hecho, se ha descrito en numerosas ocasiones la presencia de moléculas de agua atrapadas en las interfases de unión, generalmente caracterizadas por presentar valores de área accesible al disolvente muy bajos (menores de  $10 \text{ \AA}^2$ ) (Bhat *et al.* 1994; Covell & Wallqvist 1997; Wang *et al.* 1996a; Wang *et al.* 1996b) y que frecuentemente actúan como mediadores de las interacciones entre el ligando y la proteína (Damjanovic *et al.* 2005; Helms 2007; Li & Lazaridis 2007; Lu *et al.* 2007; Petrone & Garcia 2004). Es de esperar que la presencia de moléculas de agua, que actuarían como adaptadores rellenando espacios en la superficie de interacción, optimice la formación de interacciones de van der Waals y de puentes de hidrógeno y facilite la disipación de cargas, lo que contribuiría favorablemente a la entalpía de unión. Por otro lado, la desolvatación incompleta de la superficie de interacción ligando-proteína implicaría una penalización entrópica que compensaría parcialmente los efectos entálpicos favorables, de modo que la influencia de estas moléculas de agua en la energía de Gibbs suele ser generalmente modesta.

Por consiguiente, la existencia de interacciones mediadas por moléculas de agua parece una hipótesis plausible para explicar el origen del comportamiento termodinámico observado para el complejo Abl-SH3/p41. Hemos examinado las estructuras cristalográficas disponibles para este sistema y para nuestro análisis hemos seleccionado aquellas moléculas de agua que se encuentran simultáneamente a menos de  $5 \text{ \AA}$  de los átomos del ligando y de la proteína y que presentan áreas accesibles al disolvente menores de  $10 \text{ \AA}^2$ . Como se ha descrito previamente, este criterio proporciona una buena aproximación a la situación en disolución (Luque & Freire 2002).

El análisis de las cuatro estructuras del complejo Abl-SH3/p41 presentes en la unidad asimétrica (Pdb: 1bbz) revela la presencia de varias moléculas de agua en la interfase de unión que cumplen nuestros criterios. La identidad de cada una de esas moléculas, así como su accesibilidad al disolvente en cada uno de los complejos, se recogen en la Tabla 3.3.

## Resultados

**Tabla 3.3.** Moléculas de agua en la interfase de unión del complejo formado entre el dominio Abl-SH3 y el ligando p41, en el dominio libre, en la estructura de la quinasa completa de Abl, y en el complejo de Abl-SH3 con el ligando 3BP1.\*

Nº agua	Abl/p41 (a/b) <sup>1</sup>	Abl/p41 (c/d) <sup>1</sup>	Abl/p41 (e/f) <sup>1</sup>	Abl/p41 (g/h) <sup>1</sup>	Abl-SH3 <sup>2</sup>	Abl quinasa <sup>3</sup>	Abl/3BP1 (a/c) <sup>4</sup>	Abl/3BP1 (b/d) <sup>4</sup>
<b>1</b>	w 1015 (0.03)	w 1036 (1.12)	w 1019 (3.76)	w 2036 (2.85)	w 8 (36.86)		w 8 (8.22)	
<b>2</b>	w 1064 (0.00)	w 1025 (0.00)		w 1028 (0.00)		w 170 (14.09)	w 35 (9.78)	w 26 (10.55)
<b>3</b>	w 2103 (5.55)	w 2016 (16.73)		w 1003 (8.93)			w 60 (7.19)	
<b>4</b>	w 1082 (2.37)	w 1067 (4.89)	w 1097 (0.00)	w 1001 (6.25)		w 35 (9.91)	w 42 (6.51)	w 82 (6.27)
<b>5</b>	w 1105 (0.00)	w 1060 (0.00)	w 1061 (0.00)	w 1018 (0.54)		w 44 (5.06)		w 88 (4.65)
			W 1089 (0.00)					
			w 2082 (1.14)					

\*El área accesible al disolvente de cada una de las aguas se muestra entre paréntesis, en unidades Å<sup>2</sup>. El número de identificación que acompaña a cada molécula es el mismo que se había asignado a las estructuras cristalográficas resueltas en la base de datos de proteínas. Para mayor claridad hemos asignado un número del 1 al 5 (en negrita) para definir cada una de las posiciones que ocupan las moléculas de agua.

<sup>1</sup>Cuatro estructuras diferentes encontradas en la unidad asimétrica del complejo Abl-SH3 y p41 (1bbz).

<sup>2</sup>Estructura del dominio Abl-SH3 libre (1abq).

<sup>3</sup>Estructura de la quinasa completa de Abl (1opk).

<sup>4</sup>Dos estructuras diferentes encontradas en la unidad asimétrica del complejo Abl-SH3 y 3BP1 (1abo).

Para mayor claridad hemos asignado un número del 1 al 5 para las moléculas de agua 1015, 1064, 2103, 1082 y 1105 respectivamente, observadas en la estructura con los identificadores a/b del fichero Pdb 1bbz. Hemos encontrado moléculas de agua en posiciones equivalentes en las otras tres estructuras de la unidad simétrica, a excepción de la e/f en la que las aguas 2 y 3 no se observan.

También se puede apreciar en la Tabla 3.3 que algunas de estas moléculas de agua están presentes también en la estructura con el ligando natural 3BP1 (Pdb: 1abo) (Musacchio *et al.* 1994), a partir del cual se diseñó el péptido p41, así como en la estructura de la tirosina quinasa de Abl completa, donde el dominio SH3 se encuentra interaccionando con la secuencia de conexión entre el dominio SH2 y el

dominio quinasa (Pdb: 1opk) (Nagar *et al.* 2006; Nagar *et al.* 2003). Además, incluso en el dominio libre SH3 se observa la presencia de una de estas moléculas.

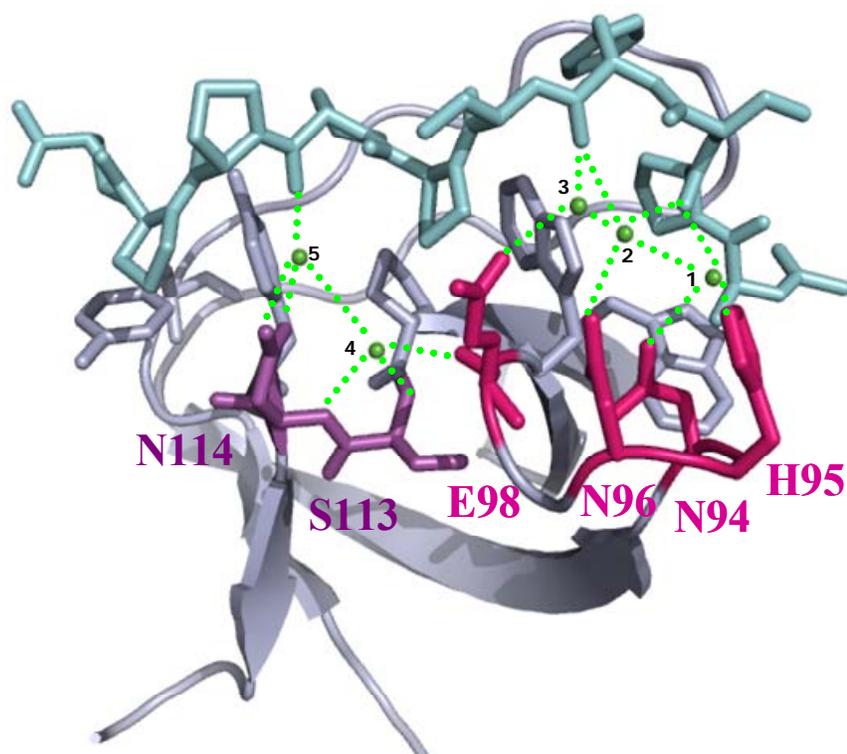
Como se muestra en la Figura 3.4, las moléculas de agua identificadas en la interfase de unión del complejo Abl-SH3/p41 ocupan posiciones idóneas para el establecimiento de una compleja red de puentes de hidrógeno que median las interacciones entre el péptido y un conjunto de restos en el dominio Abl-SH3 diferentes de los que conforman el sitio de unión característico para el reconocimiento de secuencias ricas en Prolina.

Se observan dos regiones de hidratación claramente diferenciadas en el sitio de unión: una formada por las moléculas de agua 1, 2 y 3 que median las interacciones entre los restos N94, H95, N96 y E98 en el lazo n-Src del dominio SH3 y la región de especificidad del ligando; y una segunda definida por moléculas de agua 4 y 5, que establecen interacciones entre los restos S113 y N114 en la hélice  $3_{10}$  y el motivo central del ligando en conformación PPII.

Para más claridad, en la Figura 3.5 se muestra una representación bidimensional de los puentes de hidrógeno mediados por estas moléculas de agua así como de las interacciones directas establecidas por los restos de la proteína y el ligando en la interfase de unión. En la Tabla 3.4. se resumen las distancias correspondientes a los distintos puentes de hidrógeno para el complejo Abl-SH3/p41 a los que se les ha asignado un número para facilitar su identificación. Se muestran también las distancias de los enlaces equivalentes observados en Abl-SH3/3BP1 y en el dominio Abl-SH3 libre. Como se puede observar, las moléculas de agua participarían en el establecimiento de 18 puentes de hidrógeno, 10 con restos del dominio SH3 (cuatro de ellos mediados por las aguas 1, 2 y 3, y seis por las aguas 4 y 5), 5 con grupos del ligando (cuatro de los cuales están establecidos con las aguas 1, 2 y 3 y sólo uno por la molécula 5) y tres entre moléculas de agua.

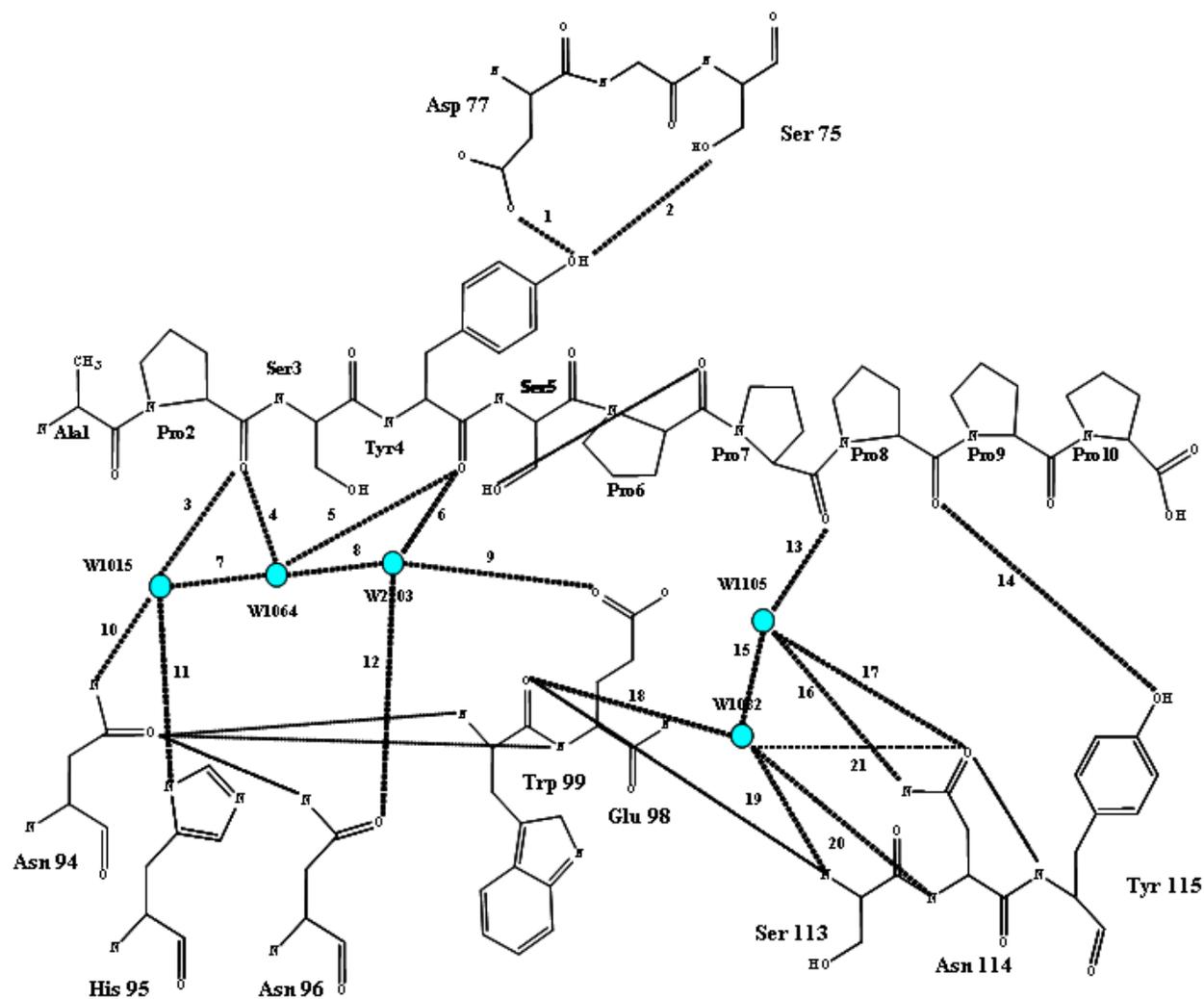
En resumen, este análisis indica que la interacción entre p41 y Abl-SH3 parece establecerse mediante un mecanismo dual que implicaría: *i*) la intercalación de restos prolina entre los restos aromáticos que delimitan los bolsillos hidrofóbicos en el sitio de unión canónico, y *ii*) el establecimiento de una compleja red de

## Resultados



**Figura 3.4.** Representación del dominio Abl-SH3 formando complejo con el péptido p41 (Pdb: 1bbz). Los residuos del lazo n-Src implicados en la formación de puentes de hidrógeno con moléculas de agua se han representado en color fucsia, y los residuos de la hélice 3<sub>10</sub> en violeta. También se han mostrado en estructura de varilla (gris azulado) los residuos aromáticos que componen el sitio de unión. El péptido p41 se muestra en color turquesa y los puentes de hidrógeno están representados por líneas discontinuas de color verde. Las moléculas de agua enterradas en la interfase de unión se indican con números del 1 al 5, también en color verde

puentes de hidrógeno entre el péptido y regiones adyacentes a estos bolsillos en el dominio SH3, mediadas por un conjunto de moléculas de agua ocluidas en la interfase de unión. Es plausible que estas interacciones, que definen una interfase de unión extendida con mayor polaridad efectiva (el cociente  $\Delta ASA_{ap}/\Delta ASA_{pol}$  disminuye de 2.88 a 1.65), tengan una influencia importante en el valor de los parámetros termodinámicos de la interacción con el ligando y contribuyan notablemente al carácter exotérmico de la interacción.



**Figura 3.5** Representación bidimensional del entramado de puentes de hidrógeno mediado por las moléculas de agua ocluidas en la interfase de unión del complejo formado entre p41 y Abl-SH3. Las cinco moléculas de agua observadas (esferas color turquesa) aparecen con el mismo número que el fichero original (Pdb: 1bbz), mientras que los aminoácidos de proteína aparecen con la numeración correspondiente a la secuencia de la quinasa completa de Abl. Hemos utilizado un número para identificar cada uno de los puentes de hidrógeno (líneas discontinuas) observados en la interfase de unión.

## Resultados

**Tabla 3.4.** Distancias de enlace de los puentes de hidrógeno mediados por moléculas de agua observados en los diferentes complejos así como en el dominio Abl-SH3 libre.

Interacción ID*	P41 (A/B) <sup>1</sup>	P41 (C/D) <sup>1</sup>	P41 (E/F) <sup>1</sup>	P41 (G/H) <sup>1</sup>	3BP1 (A/C) <sup>2</sup>	3BP1 (B/D) <sup>2</sup>	Abl-SH3 <sup>3</sup>
1	2.57	3.55	3.07	3.90	--	--	--
2	3.23	2.31	4.49	3.00	--	--	--
3	2.67	3.14	2.90	2.71	3.73	2.83	--
4	3.08	2.82	--	2.78	2.78	3.23	--
5	3.22	3.34	--	3.18	3.04	3.75	--
6	3.91	3.95	--	3.33	--	--	--
7	3.07	3.28	--	2.98	3.56	2.75	--
8	2.76	2.98	--	2.82	2.78	--	--
9	2.86	3.15	--	3.02	2.60	--	--
10	2.81	2.87	2.80	2.82	3.02	2.80	2.67
11	3.30	3.46	3.56	3.58	--	--	--
12	3.31	3.94	--	3.48	3.24	--	--
13	2.49	2.45	2.55	2.64	--	2.74	--
14	2.74	2.71	2.65	2.68	2.68	2.89	--
15	3.65	3.38	3.16	3.51	--	3.06	--
16	2.70	--	3.13	2.60	--	3.20	--
17	2.49	--	2.51	2.51	--	2.67	--
18	2.64	2.81	2.73	2.90	2.87	3.10	--
19	2.88	3.18	3.19	3.24	3.30	3.36	--
20	3.02	3.06	3.07	2.90	3.10	3.22	--
21	3.12	--	2.78	3.05	2.83	2.96	--

\*El número de interacción es el mismo que se ha utilizado para identificar los puentes de hidrógeno en la Figura 3.5.

<sup>1</sup>Cuatro estructuras diferentes encontradas en la unidad asimétrica del complejo Abl-SH3 y p41 (1bbz).

<sup>2</sup>Dos estructuras diferentes encontradas en la unidad asimétrica del complejo Abl-SH3 y 3BP1 (1abo).

<sup>3</sup> Dominio Abl-SH3 (1abq).

### **3.1.1.5. Diseño de mutaciones conservativas en Abl-SH3 para perturbar la red de interacciones mediadas por moléculas de agua**

Con el fin de investigar la función de las moléculas de agua en la interfase de unión del complejo Abl-SH3/p41 hemos generado un conjunto de mutantes conservativos del propio dominio Abl-SH3. Nuestra intención es perturbar suficientemente la red de interacciones mediadas por moléculas de agua pero causando el mínimo impacto en las propiedades estructurales y conformacionales del dominio. Es de esperar que la investigación de los efectos de estas sustituciones sobre la energética de unión, la estructura y la propia dinámica de la interacción proporcione información de utilidad para la identificación de las interacciones clave para la unión. Asimismo, esta información es de indudable interés en el desarrollo de estrategias de diseño que permitan explotar las interacciones mediadas por moléculas de agua para la optimización de la afinidad, e incluso la especificidad de unión, en este sistema.

Como dianas para las mutaciones hemos elegido las posiciones N94, N96 y N114, cuyas cadenas laterales se encuentran implicadas en la coordinación de las moléculas de agua (Figuras 3.4 y 3.5). Se descartó de antemano la posibilidad de mutar otras cadenas laterales que también participan en la coordinación de las moléculas de agua como H95, E98, W99 o S113, para evitar la introducción de sustituciones no conservativas que pudieran complicar la interpretación de los efectos termodinámicos.

Las mutaciones se han seleccionado de manera que la cadena lateral del aminoácido implicado en alguno de los puentes de hidrógeno ha sido, bien eliminada mediante sustitución por Ala, o bien modificada en longitud pero manteniendo su potencial para participar en enlaces de hidrógeno cambiándola por Thr o Gln. Es de esperar que cualquiera de estas sustituciones provoque la pérdida o el establecimiento de nuevas interacciones, o incluso la liberación de alguna molécula de agua, por lo que el análisis de sus efectos en la energética de unión proporcionará información valiosa sobre las interacciones establecidas en la interfase de unión.

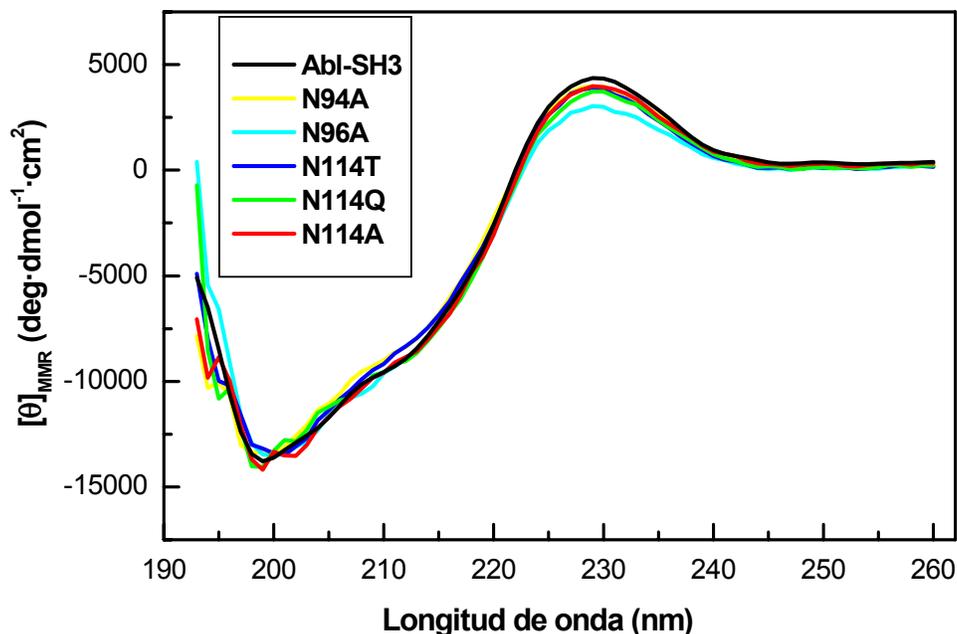
## **Resultados**

Hemos generado los mutantes a Ala en las tres posiciones y hemos seleccionado para la introducción de mutaciones por Thr y Gln un resto de cada una de las zonas de hidratación descritas en las Figuras 3.4 y 3.5. Concretamente, hemos seleccionado la posición N94 que se encuentra en la zona del lazo n-Src e interacciona directamente y con la menor distancia de enlace con la molécula de agua más enterrada en su zona (agua 1 en las Figuras 3.4 y 3.5); y también la posición N114 que está en la región de la hélice 3<sub>10</sub> y es la única cadena lateral de Abl-SH3 implicada en interacciones con moléculas de agua en esta zona. Por tanto, se han generado un total de 7 mutantes (N94A, N94T, N94Q, N96A, N114A, N114T y N114Q).

### **3.1.1.6. Las mutaciones en posición N114 no perturban las propiedades conformacionales del dominio Abl-SH3**

Con objeto de evaluar el impacto de las mutaciones diseñadas sobre la estabilidad estructural y propiedades conformacionales del dominio Abl-SH3 hemos realizado una serie de experimentos de dicroísmo circular y calorimetría diferencial de barrido. Este estudio nos permitirá asegurarnos de que los posibles efectos de las mutaciones en la energética de unión al ligando p41 sean debidos exclusivamente a la modulación de las interacciones que establece la cadena lateral de los restos mutados en la interfase de unión y no responden a alteraciones en las propiedades conformacionales y estructurales del dominio.

En la Figura 3.6 se muestran los espectros de dicroísmo circular en la región UV-lejano recogidos entre 193 y 260 nm a una temperatura de 5 °C. En todos los casos, los experimentos se han realizado en tampón glicina 50 mM pH 2.5, que corresponde a las condiciones de mayor reversibilidad utilizadas para los experimentos de calorimetría diferencial de barrido como explicaremos después. Los espectros muestran que los mutantes se encuentran correctamente plegados, observándose diferencias en general poco significativas con la forma silvestre. En todos los casos se observa una evolución espectral acorde con los elementos de estructura secundaria que componen el dominio SH3. Así, podemos apreciar el máximo a 220 nm y el mínimo a 200 nm característicos de estructura no periódica desordenada, el mínimo a 215 nm al que contribuyen las láminas beta, además del



**Figura 3.6.** Espectros de dicroísmo circular de la proteína silvestre Abl-SH3 y los mutantes diseñados tomados en glicina 50 mM, pH 2.5, y a una temperatura de 5 °C.

aumento de la señal por debajo de 200 nm característico de estructuras periódicas (láminas beta fundamentalmente en nuestro caso).

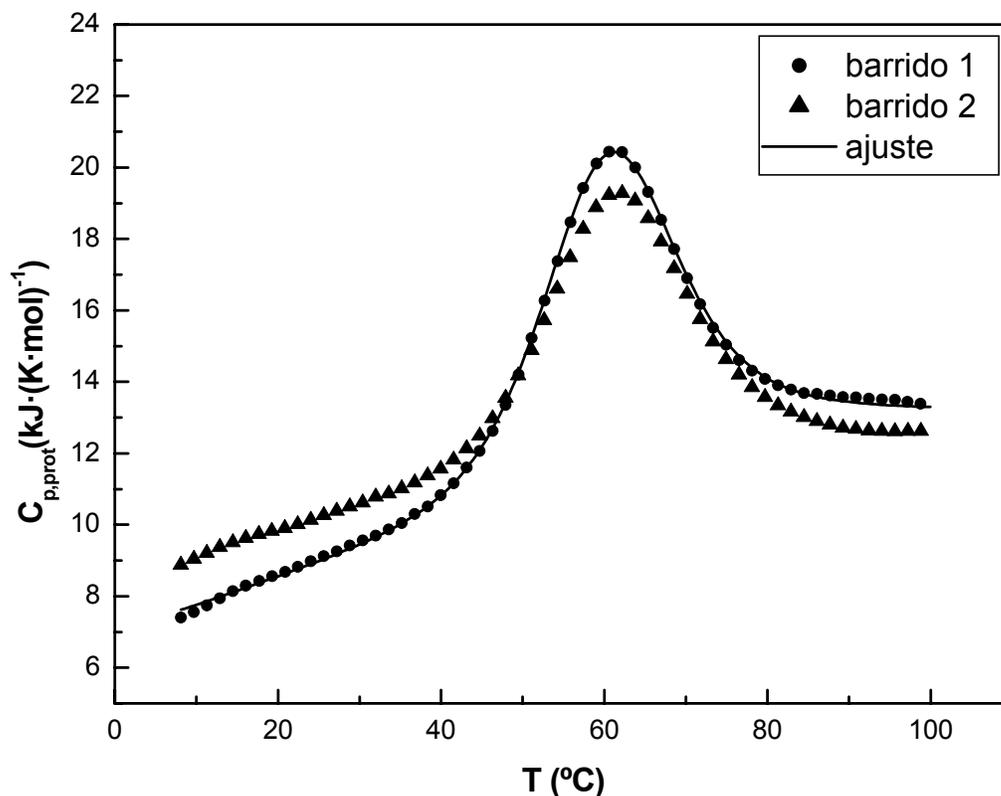
Con el fin de caracterizar el equilibrio conformacional de los mutantes en condiciones similares a las escogidas para los estudios de interacción investigamos la reversibilidad del desplegamiento de los mutantes en fosfato sódico 50 mM pH 7.0, condiciones en las que también se ha estudiado la estabilidad de la proteína silvestre (Filimonov *et al.* 1999). Sin embargo, hemos observado que en estas condiciones tanto la proteína silvestre como los mutantes presentan un cierto grado de irreversibilidad y algunos de los mutantes precipitan una vez superada la temperatura de desplegamiento, lo que impide el análisis de las transiciones según un modelo de equilibrio. Como consecuencia, hemos explorado condiciones experimentales alternativas y finalmente hemos encontrado que tanto la proteína silvestre como todos los mutantes presentan un desplegamiento altamente reversible en glicina 50 mM pH 2.5, condiciones en las que, por tanto, sería posible aplicar un modelo de equilibrio para el análisis de los termogramas obtenidos. En general, como es característico para proteínas de bajo peso molecular, las trazas son poco cooperativas lo que dificulta el cálculo de los parámetros termodinámicos.

## Resultados

Por este motivo, hemos empleado concentraciones de  $3 \text{ mg}\cdot\text{mL}^{-1}$  lo que permite obtener una intensidad de la señal adecuada para el análisis.

Para el análisis de las curvas de DSC a pH 2.5 hemos aplicado el modelo de equilibrio de dos estados, suponiendo una dependencia lineal con la temperatura para los estados nativo y desplegado (Sección 2.2.10.3; Apéndice I), como ha sido descrito previamente para la forma silvestre (Filimonov *et al.* 1999). Generalmente, para muchas proteínas de tipo globular suele utilizarse una función cuadrática para describir la dependencia de la capacidad calorífica del estado desplegado con la temperatura (Filimonov *et al.* 1982; Privalov 1979; Privalov & Potekhin 1986), aunque para esta proteína una dependencia lineal puede describirla adecuadamente (Filimonov *et al.* 1999). En la Figura 3.7 se muestran el primer y segundo barrido de temperatura obtenidos para la proteína silvestre en estas condiciones. Como podemos observar, el segundo barrido de temperatura muestra una reversibilidad en torno al 80 %, incluso completando el primer barrido hasta los 100 °C. Del ajuste del primer barrido obtenemos los parámetros termodinámicos  $T_m$  y  $\Delta H_m$ , cuyos valores fueron de 59.6 °C y  $170 \text{ kJ}\cdot(\text{K}\cdot\text{mol})^{-1}$ , respectivamente.

El grado de reversibilidad de los mutantes es similar al mostrado por la proteína silvestre. En la Figura 3.8 se muestran los termogramas obtenidos para todos ellos junto con su mejor análisis según el modelo de dos estados. Según podemos observar en la Figura 3.9, la dependencia de la entalpía de desplegamiento  $\Delta H_m$  de cada mutante con su temperatura de desplegamiento  $T_m$ , en el intervalo de temperaturas experimentalmente accesible, se ajusta bien a una línea recta en la que se incluye la propia proteína silvestre. Esto indica que todas las variantes del dominio Abl-SH3 pueden describirse de acuerdo a una única función entalpía y que, por consiguiente, no sufren cambios energéticamente significativos en su conformación como consecuencia de la mutación. Una excepción son los mutantes de la posición 94 que, en el caso de N94A y N94Q, quedan fuera de la correlación, por lo que no se han considerado para el ajuste lineal de los datos. De la pendiente de la recta obtenemos el cambio de capacidad calorífica de desplegamiento ( $\Delta C_{p,m} = 4.7 \text{ kJ}\cdot(\text{K}\cdot\text{mol})^{-1}$ ) el cual, dentro del elevado error experimental, podría compararse razonablemente bien con el obtenido previamente para la proteína silvestre ( $\Delta C_{p,m} = 3.4 \text{ kJ}\cdot(\text{K}\cdot\text{mol})^{-1}$ ) (Filimonov *et al.* 1999).

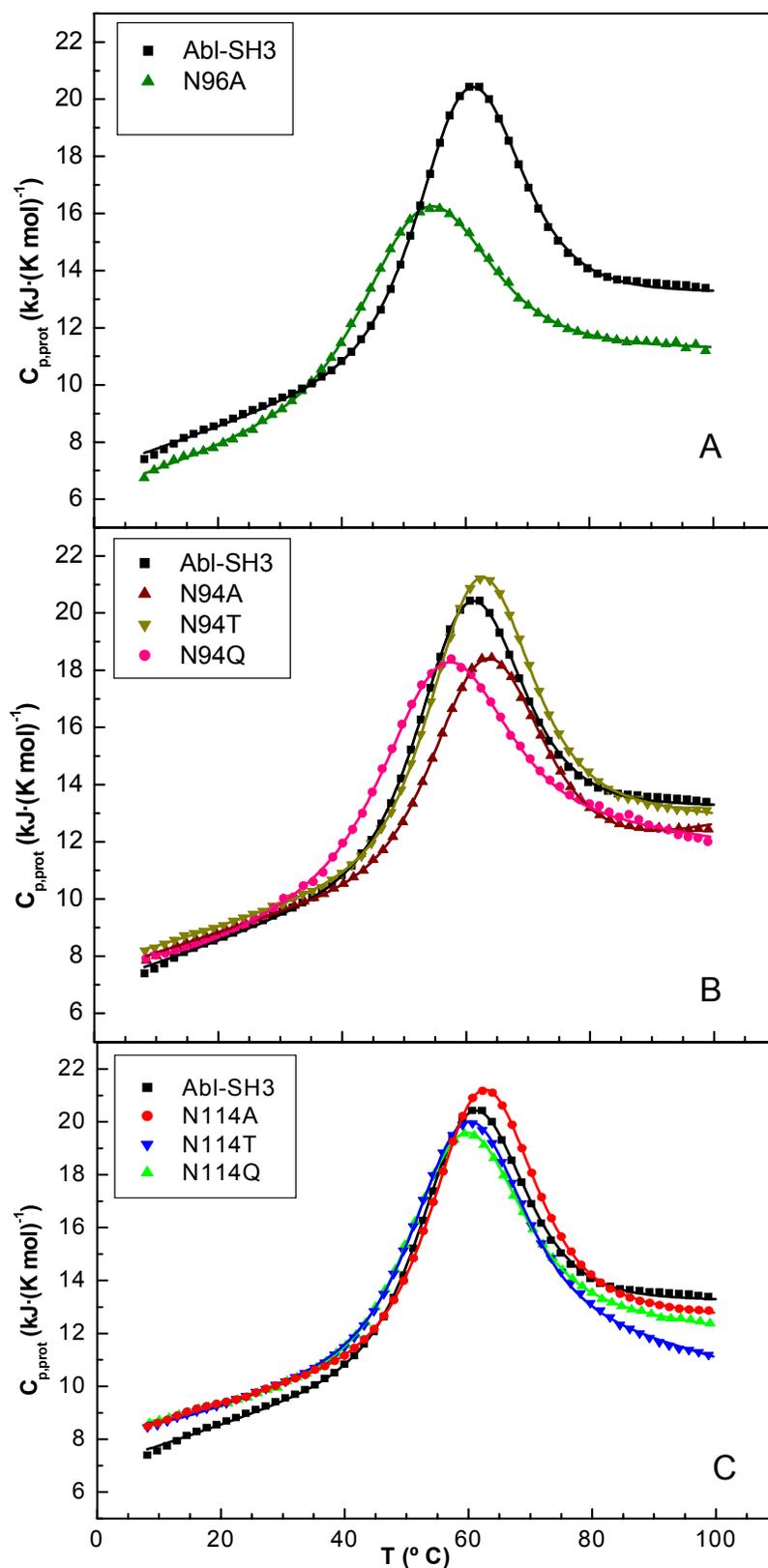


**Figura 3.7.** Dependencia de la capacidad calorífica molar parcial con la temperatura del dominio Abl-SH3, obtenida a partir del experimento realizado en glicina 50 mM, pH 2.5. Los círculos representan el primer barrido de temperatura, y los triángulos el segundo barrido. Las líneas continuas corresponden al mejor ajuste al modelo de equilibrio de desplegamiento de dos estados para el primer barrido de temperatura (sección 2.2.10.3, Apéndice I).

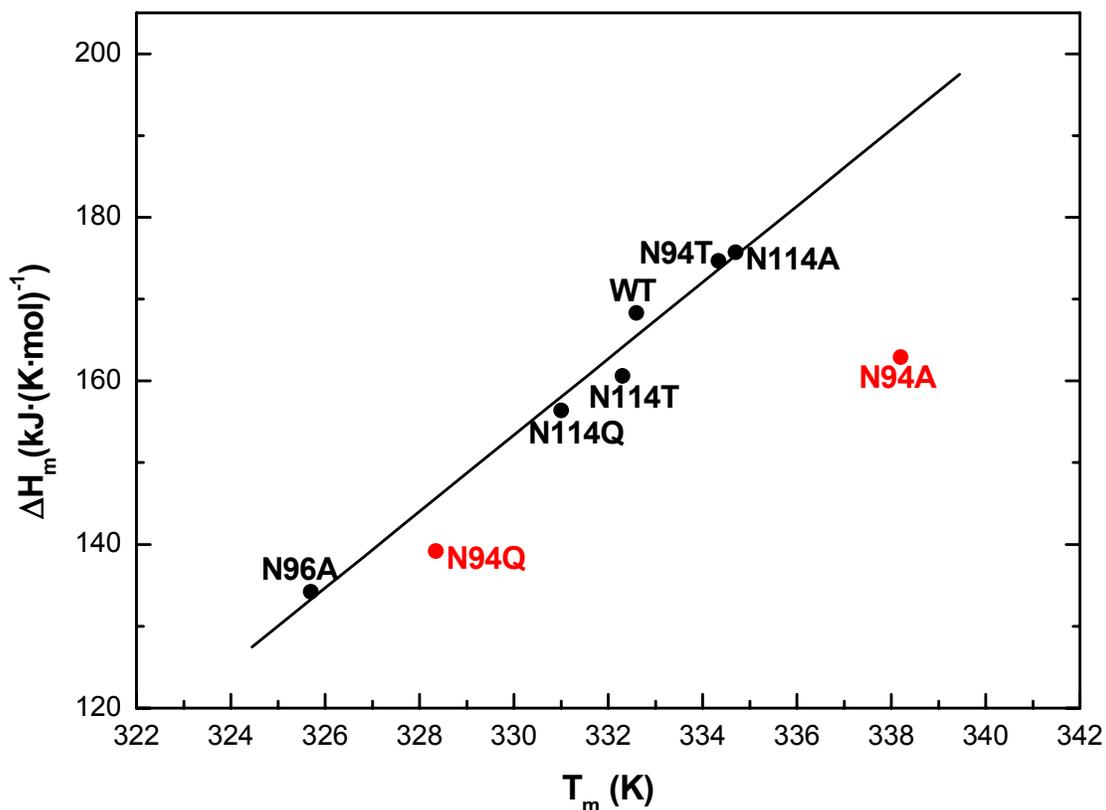
Hemos de tener en cuenta que la diferencia entre ambos valores de  $\Delta C_{p,m}$  podría ser consecuencia de la baja cooperatividad de las trazas experimentales y del pequeño intervalo de temperaturas de desplegamiento (10 grados aproximadamente) en el que se han calculado ambos valores, lo que hace que el cálculo de la pendiente de la recta sea susceptible de apreciable error.

Examinando la Figura 3.9, podríamos admitir que las mutaciones en la posición N114, en la hélice  $3_{10}$ , son las mejor toleradas desde un punto de vista conformacional pues presentan una función entalpía común con la proteína silvestre. De igual modo, las mutaciones en esta posición presentan efectos pequeños en estabilidad, siendo en todos los casos las diferencias en  $\Delta H_m$  y  $T_m$  inferiores a  $10 \text{ kJ}\cdot\text{mol}^{-1}$  y  $2 \text{ }^\circ\text{C}$  respectivamente con respecto al dominio silvestre, lo

## Resultados



**Figura 3.8.** Dependencia de la capacidad calorífica molar parcial con la temperatura para el dominio Abl-SH3 y el mutante N96A (panel A); Abl-SH3 y mutantes de la posición 94 (panel B); y Abl-SH3 y mutantes de la posición 114 (panel C), obtenidas a partir de los experimentos en glicina 50 mM pH 2.5. Los símbolos representan los datos experimentales, mientras que las líneas continuas corresponden al mejor ajuste al modelo de equilibrio de desplegamiento de dos estados (sección 2.2.10.3., Apéndice I).



**Figura 3.9.** Dependencia de entalpía molar parcial  $\Delta H_m$  frente a la temperatura de desplegamiento  $T_m$ , ambos obtenidos a partir del ajuste no lineal de las curvas de DSC al modelo de equilibrio de dos estados, suponiendo para las funciones capacidad calorífica molar parcial de estado nativo y desplegado una dependencia lineal con la temperatura (sección 2.2.10.3 y Apéndice I).

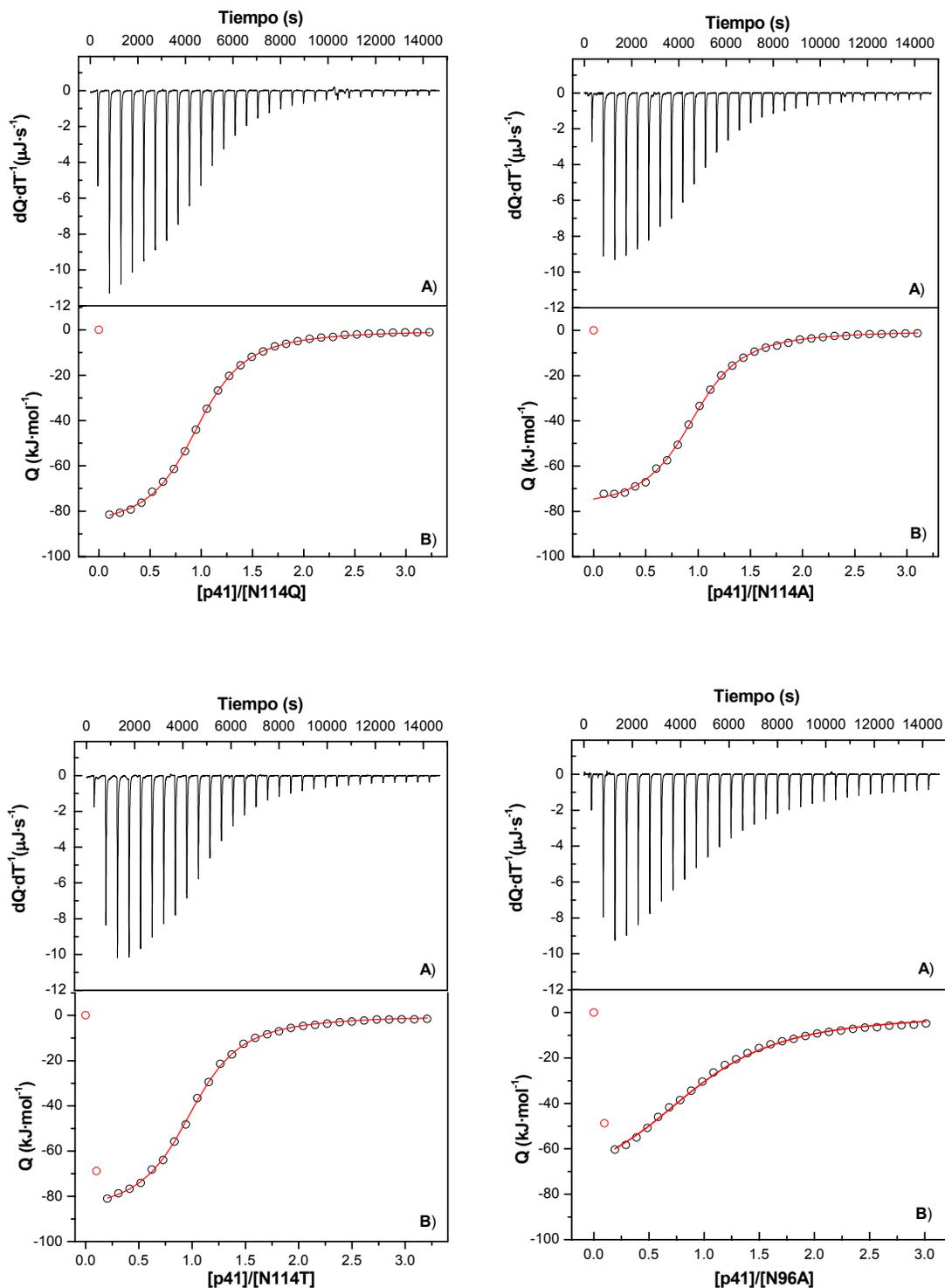
que se traduce en valores de  $\Delta\Delta G$  no superiores a  $2 \text{ kJ}\cdot\text{mol}^{-1}$ . Por el contrario, las mutaciones introducidas en la zona del lazo n-Src (posiciones N94 y N96) producen mayores perturbaciones en los parámetros termodinámicos. Así, aunque el mutante N96A correlaciona con los otros en la Figura 3.9, presenta una significativa disminución de 6 grados en el valor de  $T_m$  respecto al obtenido para la proteína silvestre, mientras que el mutante N94A muestra un aumento equivalente, además de una función entalpía diferente. Algo parecido también ocurre con N94Q, lo que indicaría que se pueden estar produciendo cambios conformacionales significativos como consecuencia de las mutaciones en ambas posiciones y que, en cualquier caso, las mutaciones afectan notablemente la estabilidad del dominio.

## **Resultados**

### **3.1.1.7. La sustitución de las cadenas laterales que coordinan las moléculas de agua interfaciales produce cambios significativos en la energética de unión**

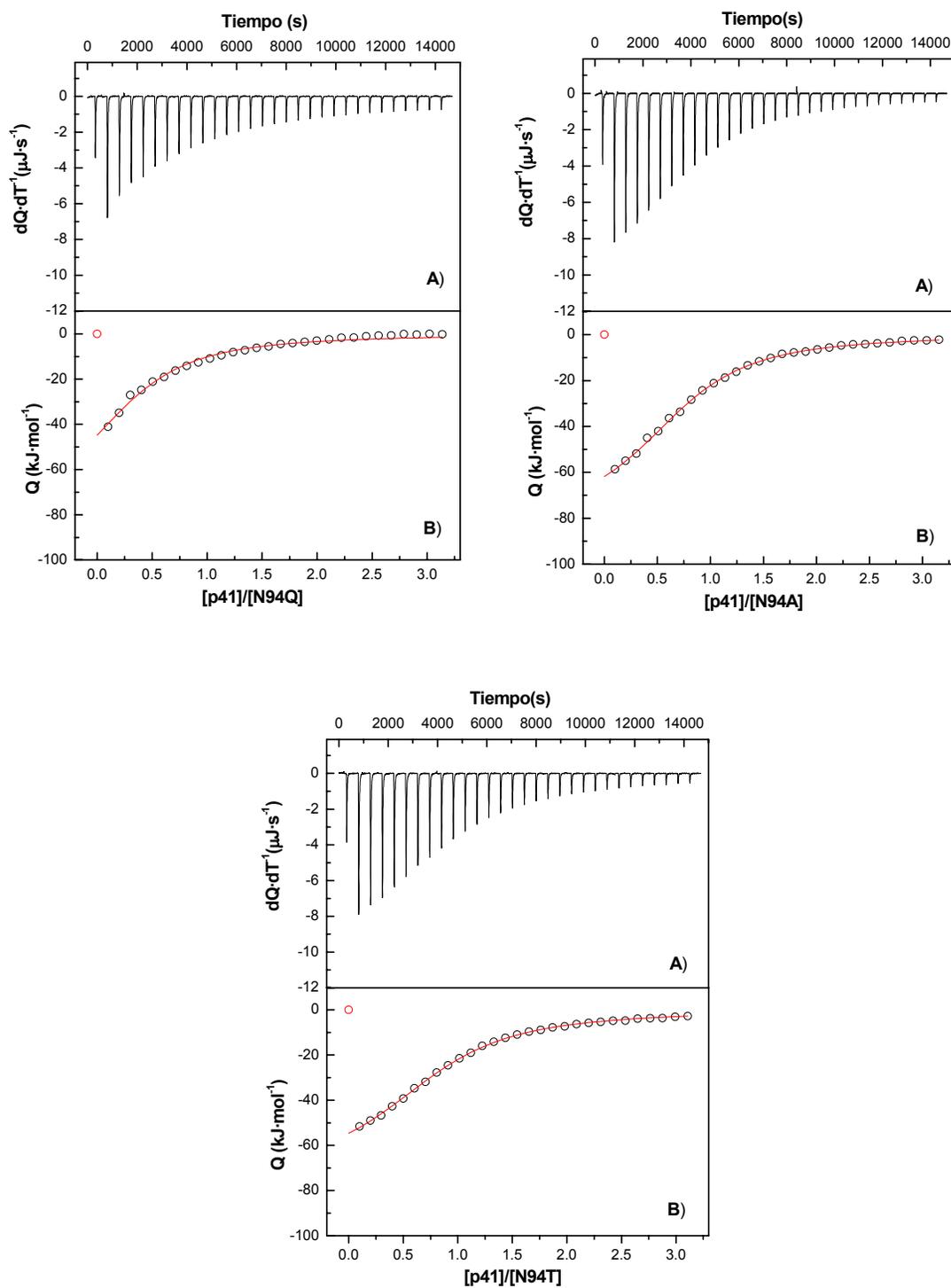
Una vez confirmada la integridad conformacional de las distintas variantes del dominio Abl-SH3, hemos procedido a evaluar el efecto de las mutaciones sobre la energética de unión al ligando p41 para lo que hemos hecho una caracterización termodinámica mediante calorimetría isotérmica de titulación. Hemos realizado los experimentos de ITC utilizando como condiciones estándar fosfato sódico 20 mM pH 7.0 y 25 °C. De todas las condiciones utilizadas en la caracterización termodinámica con la proteína silvestre el tampón fosfato es el que presenta el menor calor de protonación a pH 7.0 (prácticamente nulo), de modo que el valor de entalpía de unión aparente determinado experimentalmente puede ser considerado, dentro del error, como el valor de entalpía intrínseca de la interacción. Para los experimentos de ITC con los mutantes hemos utilizado las mismas concentraciones de proteína y ligando, así como el resto de parámetros instrumentales (número de inyecciones, etc.) que en el caso de la proteína silvestre.

En la Figura 3.10 se presentan los experimentos de titulación con el ligando p41 de los distintos mutantes, donde se muestran los termogramas (paneles superiores) junto con las curvas isothermas de unión y los correspondientes ajustes por regresión no lineal al modelo de “n” sitios idénticos e independientes (paneles inferiores). Los valores de los parámetros termodinámicos resultantes del análisis para los distintos mutantes se resumen en la Figura 3.11. En los mutantes N96A y N114A/T/Q el ligando p41 se une a un único sitio de unión tal y como se había descrito para la proteína silvestre, sin embargo para los mutantes de la posición 94 se han obtenido valores de 0.8 (N94A), 0.8 (N94T) y 0.5 (N94Q) lo que probablemente esté relacionado con el cambio en el equilibrio conformacional del dominio ya confirmado en los experimentos de plegamiento. Este cambio conformacional podría afectar a la termodinámica de la interacción con p41, por lo que los valores recogidos en la Figura 3.11 para estas mutaciones deberían ser interpretados con cierta cautela.



**Figura 3.10.** Experimentos de titulación mediante ITC del péptido p41 con los mutantes N96A, N114Q, N114A, N114T, en fosfato sódico 20 mM, pH 7.0 y 25 °C. **A:** Termograma de titulación formado por los calores por unidad de tiempo liberados tras la inyección de p41 (30 inyecciones de 7  $\mu\text{L}$  de una disolución de p41 a 0.6 mM) sobre el mutante (32  $\mu\text{M}$  en la célula calorimétrica). **B:** Isotherma de unión correspondiente al experimento en A, obtenida después de normalizar los calores de titulación y corregir por los calores de dilución. Los datos experimentales se representan como círculos, mientras que la línea continua corresponde al mejor ajuste para el modelo de "n" sitios de unión idénticos e independientes (ecuación 2.21).

## Resultados



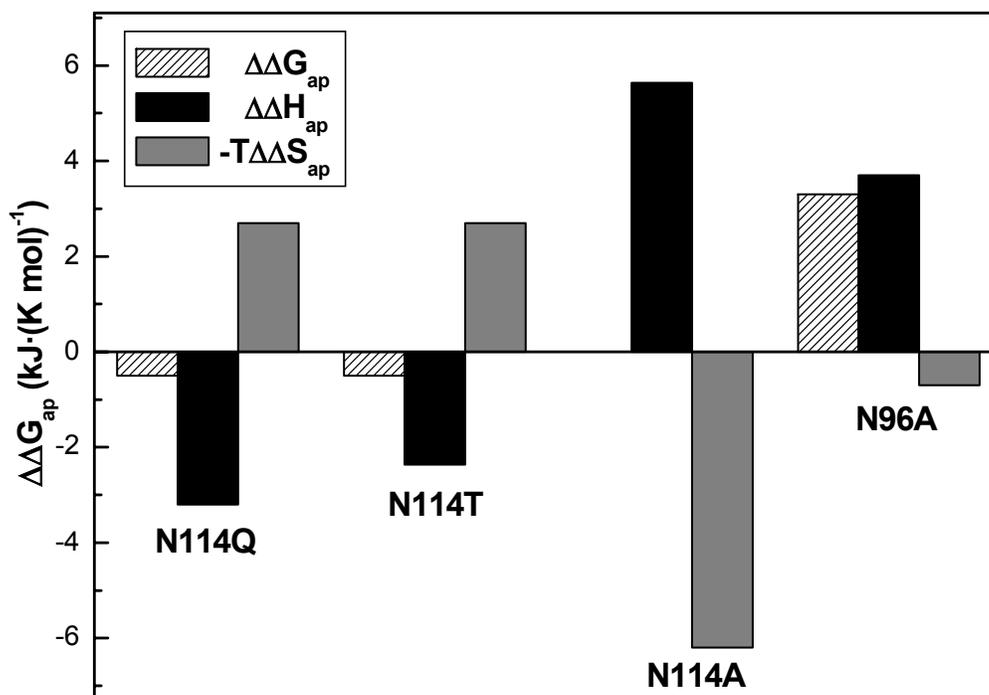
**Figura 3.10.** Experimentos de titulación mediante ITC del péptido p41 con los mutantes N94A, N94Q, N94T. La información mostrada en los paneles se corresponde a lo descrito en el pie de la figura anterior.

Como puede apreciarse en la Figura 3.11, en todos los casos el proceso de unión está caracterizado por un patrón termodinámico similar al observado para la proteína silvestre, de modo que la interacción está gobernada por una entalpía de unión muy favorable a la que se opone parcialmente una contribución entrópica desfavorable. Además, es destacable que a pesar de que ninguna de las posiciones mutadas establecería interacciones directas con el ligando según se desprende de la propia estructura cristalográfica, existen diferencias significativas en los parámetros termodinámicos de unión, lo que constituye una evidencia más de la contribución energética de las diferentes moléculas de agua ocluidas.

Centrándonos en las posiciones N96 y N114, en la Figura 3.11 se pueden observar las diferencias relativas al dominio silvestre, tanto en la energía libre de Gibbs, como en sus contribuciones entálpica y entrópica. La diferencia más notable en el valor de entalpía aparente se obtiene en las mutaciones por Ala, siendo esta diferencia en el caso de los mutantes a Gln y Thr casi despreciable o ligeramente más favorable ( $3.2 \text{ kJ}\cdot\text{mol}^{-1}$  más negativa para N114Q y  $2.4 \text{ kJ}\cdot\text{mol}^{-1}$  para N114T). También es destacable que en el caso de la posición N114 observamos que existe una compensación casi perfecta entre los cambios en los términos entálpico y entrópico, de modo que finalmente la afinidad de unión no varía significativamente respecto a la proteína silvestre. Sin embargo, en el mutante N96A la energía libre de Gibbs disminuye  $3.3 \text{ kJ}\cdot\text{mol}^{-1}$  con respecto a la proteína silvestre.

En resumen, estos resultados demuestran que la modificación de las cadenas laterales de los restos del dominio Abl-SH3 que coordinan las moléculas de agua en la interfase de unión puede producir efectos significativos en la energética de unión del ligando, aún cuando estos restos se encuentran en todos los casos a más de 7 Å de distancia de éste, no habiendo evidencia alguna de ningún contacto directo con p41. En el caso de aquellas mutaciones, como las sustituciones en posición N114, cuya estabilidad y equilibrio conformacional del dominio es muy similar al de la proteína silvestre, podemos interpretar estos efectos en términos exclusivamente de modulación de las interacciones mediadas por moléculas de agua. De este modo, como es de esperar, la eliminación de la cadena lateral de Asn en el mutante N114A se traduce en una disminución de la entalpía de unión, en este caso  $5.6 \text{ kJ}\cdot\text{mol}^{-1}$  menos favorable, y una contribución entrópica más favorable lo que, en el contexto de las Figuras 3.4 y 3.5, posiblemente esté relacionado con la pérdida de

## Resultados



**Figura 3.11.** Representación en forma de barras de las diferencias con respecto a la proteína silvestre Abl-SH3 de la energía libre de Gibbs, y de los términos entálpico y entrópico obtenidos en los experimentos de ITC con los mutantes seleccionados.

los puentes de hidrógeno que el resto N114 establece con las moléculas de agua 4 y 5, o incluso con la desaparición de alguna de ellas de la interfase de unión. Por el contrario, los mutantes N114T y N114Q muestran una contribución entálpica ligeramente más favorable que la proteína silvestre, lo que parece indicar que se conservan las interacciones establecidas con las moléculas de agua 4 y 5, o incluso llegarían a optimizarse ligeramente.

Las mutaciones introducidas en la región del lazo n-Src producirían efectos termodinámicos más complejos cuya interpretación no es trivial debido a la susceptibilidad de esta región frente a las mutaciones que, como hemos visto en el apartado anterior, en algunos casos alteran significativamente las propiedades conformacionales del dominio. Por lo tanto, nos centraremos en el estudio de las mutaciones en posición N114 para profundizar en el análisis del papel de las moléculas de agua en la interfase de unión de Abl-SH3/p41, que se describirá en los siguientes apartados.

**3.1.1.8. Las mutaciones en el resto N114 producen efectos locales, conservando la estructura e interacciones del dominio silvestre**

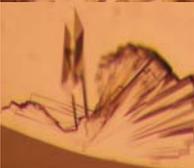
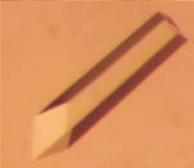
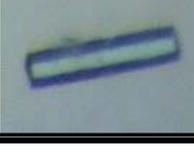
Puesto que nuestra intención es poder racionalizar las diferencias en los parámetros termodinámicos de unión al ligando p41 observadas entre los mutantes N114 y la proteína silvestre en base a la pérdida o modificación de las interacciones establecidas con las moléculas de agua en la interfase de unión, e incluso a la liberación de alguna de éstas, es imprescindible obtener información estructural de alta calidad sobre los mutantes y sus complejos que, además, nos permita confirmar la ausencia de cambios estructurales con respecto a la proteína silvestre.

Con este fin hemos abordado la cristalización y posterior resolución estructural mediante difracción de rayos X de los mutantes N114A, N114Q y N114T libres y unidos a p41. Estas estructuras han sido determinadas por la misma técnica para el dominio silvestre y se han recogido del banco de datos de proteínas (códigos Pdb: 1abq y 1bbz para Abl-SH3 y Abl-SH3/p41 respectivamente). En todos los casos, el proceso de cristalización se ha basado en las condiciones previamente descritas en bibliografía para la proteína silvestre (Musacchio *et al.* 1994; Pisabarro *et al.* 1998), que posteriormente se han optimizado realizando diferentes barridos de condiciones físico-químicas experimentales (sección 2.2.12.). Una vez obtenidos los cristales y recogidos los juegos de datos de difracción en el Laboratorio de Estudios Cristalográficos (CSIC, Granada), la Dra. Ana Cámara Artigas (Universidad de Almería) se ha encargado de su resolución estructural mediante métodos de reemplazamiento molecular.

Aunque en algunos casos se obtuvieron cristales para una misma proteína en una gran variedad de condiciones físico-químicas, en este apartado mostraremos sólo aquellas condiciones finales a partir de las cuales hemos recogidos las difracciones que nos han conducido a obtener posteriormente las correspondientes estructuras tridimensionales (Tabla 3.6). No obstante, la totalidad de condiciones experimentales donde hemos obtenido cristales se recogen en los Apéndices III y IV.

## Resultados

Tabla 3.6. Condiciones experimentales de cristalización y difracción para los mutantes del dominio Abl-SH3

Proteína	pH	Pptante.	Pptante.	Técnica	Aparición	T/ aditivos	Cristal
N114A	7.0	Formato sódico 4 M	PEG300 5%	Gota sentada	5-10 días	100 K Glicerol 10 %	
N114A	5.0	Sulfato amónico 2 M	PEG300 5%	Gota sentada	15-30 días	100 K Glicerol 10 %	
N114A	4.0	Sulfato amónico 2 M	PEG300 5%	Gota sentada	15-30 días	100 K Glicerol 10 %	
N114A	3.0	Sulfato amónico 2 M	PEG300 5%	Gota sentada	15-30 días	100 K Glicerol 10 %	
N114T	7.0	Sulfato amónico 2 M	PEG300 5%	Gota sentada	15-30 días	100 K Glicerol 10 %	
N114Q	3.0	Sulfato amónico 2 M	PEG300 5%	Gota sentada	15-30 días	100 K Glicerol 10 %	

Es interesante destacar que en el caso del mutante N114A se han obtenido cristales en varias condiciones de pH de suficiente calidad como para recoger sus juegos de datos por difracción de rayos X (Tabla 3.6). Como puede observarse existe una dependencia de la calidad del cristal con el pH y, por tanto, de la resolución final de la estructura, de modo que para pH 3.0 hemos conseguido llegar a una resolución de 1.3 Å.

En la Tabla 3.7 se muestran también los datos de recogida de las difracciones y la estadística para el proceso de refinamiento de todos estos mutantes.

**Tabla 3.7.** Datos de recogida de las difracciones y estadística del proceso de refinamiento para los mutantes N114A, N114T y N114Q del dominio Abl-SH3

	N114A pH 7.0			N114A pH 5.0		
Grupo espacial	P3 <sub>2</sub> 12			C222		
Dimensiones de celdilla unidad a b c	51.35	51.35	46.39	44.99	53.07	40.95
Rango de resolución (Å)	44.50-2.04			40.95-1.75		
Número de observaciones	34846			46003		
Reflexiones únicas	3936			4648		
Completitud de datos (%)	99.9 (100) <sup>a</sup>			89.4 (69.9) <sup>a</sup>		
$R_{\text{merge}}^b$ (%)	5.3 (30.3) <sup>a</sup>			3.1 (11.7) <sup>a</sup>		
$I/\sigma(I)$	27.5 (4.0) <sup>a</sup>			44.1 (5.5) <sup>a</sup>		
<b>Refinamiento*</b>						
Restos de proteína						
Moléculas de disolvente						
$R_{\text{work}}$ (%)						
$R_{\text{free}}$ (%)						
<b>Desviaciones de geometría ideal RMS*</b>						
Enlaces (Å)						
Ángulos (grados)						
B promedio para proteína (Å <sup>2</sup> )						
Restos en zonas permitidas del gráfico Ramachandran (%) <sup>c</sup>						

\* En proceso de refinamiento.

<sup>a</sup> Los valores en paréntesis corresponden al corte de mayor resolución.

<sup>b</sup>  $R_{\text{sym}} = \frac{\sum_h \sum_l |I_{hl} - \langle I_h \rangle|}{\sum_h \sum_l \langle I_h \rangle}$ , donde  $I_l$  es la  $l$  observación de reflexión  $h$  y  $\langle I_h \rangle$  es la intensidad promedio ponderada para todas las observaciones  $l$  de la reflexión  $h$ .

<sup>c</sup> Obtenidos con el programa PROCHECK (Laskowski et al. 1993).

	N114A pH 4.0*	N114A pH 3.0
Grupo especial	C222	C222
Dimensiones de celdilla unidad a b c	43.25 51.72 39.92	44.82 53.10 41.30
Rango de resolución (Å)	39.92-1.60	41.35-1.35
Número de observaciones	65067	152773
Reflexiones únicas	5542	10303
Compleitud de datos (%)	89.9 (80.9) <sup>a</sup>	92.3 (82.2) <sup>a</sup>
$R_{\text{merge}}^b$ (%)	4.6 (24.4) <sup>a</sup>	3.3 (13.3) <sup>a</sup>
$I/\sigma(I)$	21.4 (4.4) <sup>a</sup>	32.8 (7.3) <sup>a</sup>
<b>Refinamiento*</b>		
Restos de proteína		63
Moléculas de disolvente		37
$R_{\text{work}}$ (%)		20.8 (23.5) <sup>a</sup>
$R_{\text{free}}$ (%)		23.9 (34.2) <sup>a</sup>
<b>Desviaciones de geometría ideal RMS*</b>		
Enlaces (Å)		0.014
Ángulos (grados)		2.097
B promedio para proteína (Å <sup>2</sup> )		16.0
Restos en zonas permitidas del gráfico Ramachandran (%) <sup>c</sup>		100

\* Valores para N114A a pH 4.0 en proceso de refinamiento.

	N114T pH 7.0	N114Q pH 3.0
Grupo espacial	P3 <sub>2</sub> 12	C222
Dimensiones de celdilla unidad a b c	49.75 49.75 45.13	44.68 53.00 41.10
Rango de resolución (Å)	43.10-2.25	41.1-1.80
Número de observaciones	20588	29235
Reflexiones únicas	3146 (368)	4353 (471)
Compleitud de datos (%)	99.8 (98.4)	91.3 (67.9)
$R_{\text{merge}}^b$ (%)	3.57 (28.35)	3.3 (26.9)
$I/\sigma(I)$	22.75 (3.48)	26.1 (3.9)
<b>Refinamiento</b>		
Restos de proteína	56	63
Moléculas de disolvente	18	24
$R_{\text{work}}$ (%)	22.1 (31.40)	20.8 (22.8)
$R_{\text{free}}$ (%)	28.3 (44.6)	26.7 (30.6)
<b>Desviaciones de geometría ideal RMS</b>		
Enlaces (Å)	0.010	0.011
Ángulos (grados)	1.818	2.061
B promedio para proteína (Å <sup>2</sup> )	17.8	16.7
Restos en zonas permitidas del gráfico Ramachandran (%) <sup>c</sup>	100	100

## Resultados

Para la cristalización de los complejos hemos seguido la misma estrategia, partiendo de las condiciones descritas previamente para el complejo con la proteína silvestre (Pisabarro *et al.* 1998) y realizando diversos barridos de condiciones físico-químicas hasta obtener los mejores cristales (sección 2.2.12.).

Al igual que con los mutantes libres, aunque hemos obtenido cristales en varias condiciones, sólo mostramos en la Tabla 3.8 los que hemos utilizado para difracción y la obtención de estructuras cristalográficas. Los juegos de datos y la estadística del proceso de refinamiento aparecen recogidos en la Tabla 3.9. En este caso, tanto la técnica de gota sentada como la de contradifusión en capilares nos han permitido obtener las estructuras cristalográficas a una resolución lo suficientemente buena como para examinar las posibles diferencias estructurales con la proteína silvestre y localizar las moléculas de disolvente en la interfase de unión. A diferencia de los mutantes libres, todos los complejos cristalizados presentan el mismo grupo espacial ( $P2_12_12_1$ ) independientemente del pH utilizado. La única diferencia encontrada entre el complejo de la proteína silvestre y los mutantes es que la estructura del complejo silvestre previa (código Pdb: 1bbz) contenía cuatro moléculas en la unidad asimétrica, mientras que en todas nuestras estructuras de los complejos con los mutantes aparecen sólo dos.

En resumen, hemos resuelto las estructuras de los mutantes N114A (pH 3.0, resolución 1.3 Å), N114T (pH 7.0, resolución 2.2 Å) y N114Q (pH 3.0, resolución 1.7 Å) libres, así como las estructuras de los complejos con p41 de los mutantes N114A (pH 7.0, resolución 1.7 Å) y N114Q (pH 3.5, resolución 1.7 Å). Estas últimas se muestran en la Figura 3.14 en comparación con las estructuras correspondientes del dominio Abl-SH3 silvestre obtenidas del Pdb (códigos: 1abq y 1bbz). Como puede observarse, la sustitución de la cadena lateral del resto N114 por Ala, Gln o Thr no parece inducir cambios significativos en las estructuras de los dominios libres, que presentan valores de RMSD para los carbonos alfa que varían entre 0.46 y 0.94 Å entre sí. Las únicas diferencias significativas se observan en las regiones de los lazos RT y n-Src que muestran cierta variabilidad conformacional en concordancia con la elevada flexibilidad de estos lazos descrita previamente para éste y otros dominios SH3 (Arold *et al.* 1998; Martín-García *et al.* 2007; Wang *et al.* 2001). De hecho, esta flexibilidad se refleja en la pobre densidad electrónica que caracteriza la región del lazo RT en todas las estructuras de los dominios libres,

*Tabla 3.8. Condiciones experimentales de cristalización y difracción de los complejos formados entre el péptido p41 y los mutantes N114A y N114Q de Abl-SH3*

Complejo con p41	pH	Preptante.	Preptante.	Técnica	Tiempo Aparic.	T/ aditivos	Cristal
WT	3.1	Sulfato amónico 2 M	Cloruro sódico 0.2 M	Difusión vapor Gota colgante	-	100 K	Pisabarro et. al. 1998 JMB
N114A	7.0	Sulfato amónico 3 M		Contra-difusión capilares	5-10 días	273 K	
N114Q	3.5	Sulfato amónico 2 M	Cloruro sódico 0.4 M	Difusión vapor Gota sentada	15-30 días	100 K Glicerol 10 %	

incluidas aquellas como la del mutante N114A que han sido resueltas a muy alta resolución (1.3 Å).

Es interesante señalar que en las estructuras resueltas a pH ácido el lazo n-Src presenta una conformación que difiere significativamente de la adoptada en las estructuras resueltas a pH neutro y en los complejos. La conformación observada a más bajo pH se estabilizaría por el establecimiento de contactos cristalográficos intermoleculares entre el resto H95 de una molécula y los restos D60 y E62 de otra que desaparecen al protonarse la His a pH neutro. De igual modo, como se ilustra en la Figura 3.14, la cadena lateral del resto E98 adoptaría también dos conformaciones diferentes según el pH.

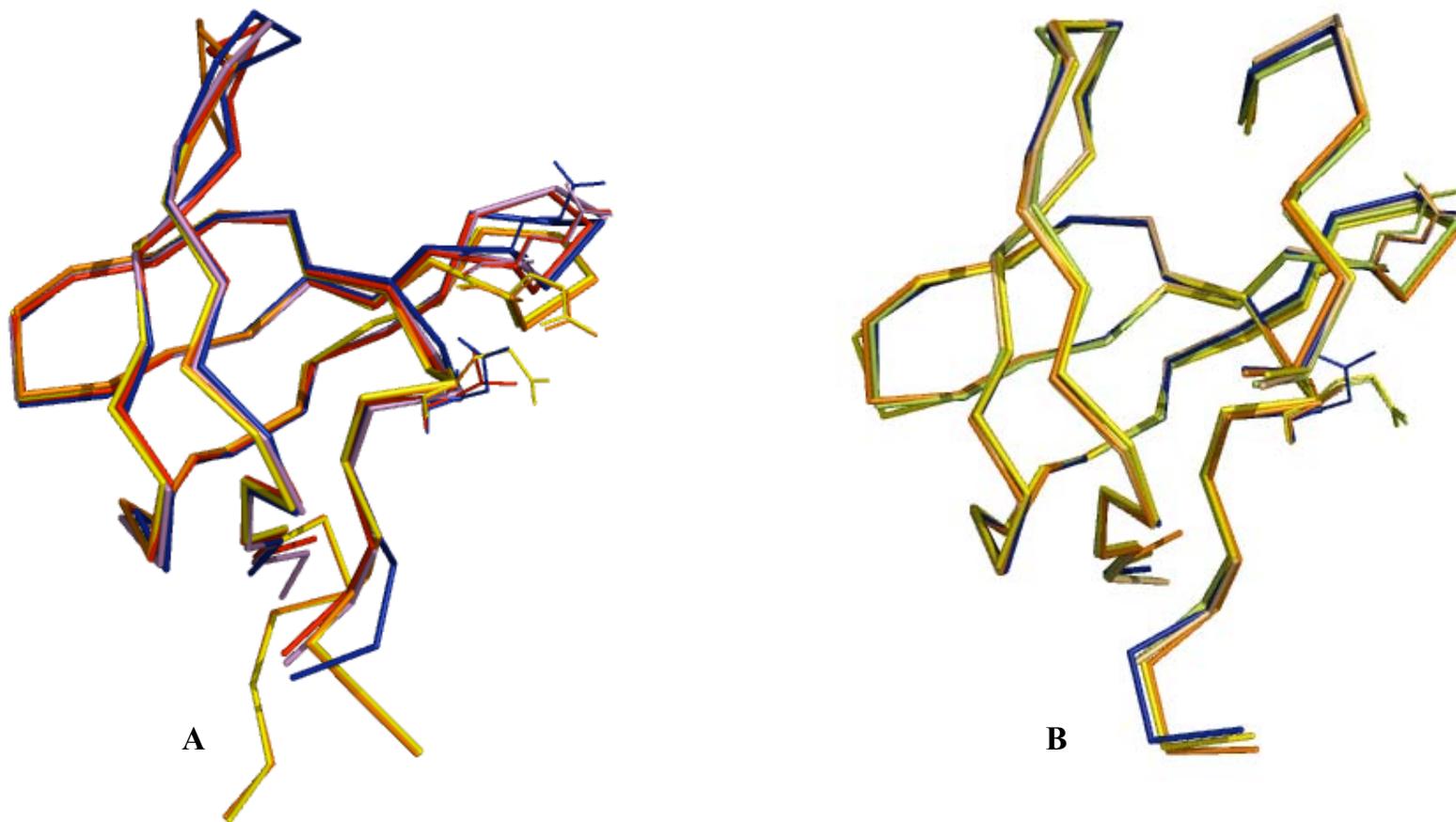
**Tabla 3.9.** Datos de recogida de rayos X y estadística del proceso de refinamiento para los complejos con el ligando p41 y los mutantes N114A y N114Q.

	Complejo N114Q:P41	Complejo N114A:P41
Grupo especial	$P2_12_12_1$	$P2_12_12_1$
Dimensiones de celdilla unidad a b c	45.996, 47.636, 55.662	48.170, 50.093, 56.431
Rango de resolución (Å)	50.0-1.85	50.0-1.75
Número de observaciones	83156	84702
Reflexiones únicas	10158 (1123) <sup>a</sup>	13236 (1638) <sup>a</sup>
Complejidad de datos (%)	92.8 (72.5)	92.2 (75.8)
$R_{\text{merge}}^b$ (%)	6.5 (24.3)	0.07 (0.24)
$I/\sigma(I)$	19.2 (3.5)	16.4 (4.4)
<b>Refinamiento</b>		
Restos de proteína	58 (cadena A), 56 (cadena B)	57 (cadena A), 55 (cadena B)
Moléculas de disolvente	56	69
$R_{\text{work}}$ (%)	19.2 (24.8) <sup>a</sup>	16.9 (20.5) <sup>a</sup>
$R_{\text{free}}$ (%)	20.2 (28.6) <sup>a</sup>	21.5 (26.8) <sup>a</sup>
<b>Desviaciones de geometría ideal RMS</b>		
Enlaces (Å)	0.012	0.008
Ángulos (grados)	2.106	1.740
B promedio para proteína (Å <sup>2</sup> )	16.6	24.1
Restos en zonas permitidas del gráfico Ramachandran (%) <sup>c</sup>	100	100

<sup>a</sup> Los valores en paréntesis corresponden al corte de mayor resolución.

<sup>b</sup>  $R_{\text{sym}} = \sum_h \sum_l |I_{hl} - \langle I_h \rangle| / \sum_h \sum_l \langle I_h \rangle$ , donde  $I_l$  es la  $l$  observación de reflexión  $h$  y  $\langle I_h \rangle$  es la intensidad promedio ponderada para todas las observaciones  $l$  de la reflexión  $h$ .

<sup>c</sup> Obtenidos con el programa PROCHECK (Laskowski et al. 1993).



**Figura 3.14.** Superposición del esqueleto peptídico de la proteína silvestre Abl-SH3 y los complejos de sus mutantes N114 con el péptido p41. **A.** Estructuras de las proteínas libres: Abl-SH3 (azul), mutante N114A (violeta) y N114T (rojo) a pH y 7.0; y N114A (naranja) y mutante N114Q (amarillo) a pH 3.0. **B.** Estructuras del complejo de la proteína silvestre (azul) y de los complejos N114A/p41, molécula a/c (naranja) y b/d (naranja claro) de la unidad asimétrica, y N114Q/p41, molécula a/c (amarillo) y b/d (lima) de la unidad asimétrica. El resto Asn114 en la proteína silvestre, y los restos Ala/Thr/Gln114 aparecen, junto con el Glu98 implicado en el contacto cristalográfico, en estructura de varillas.

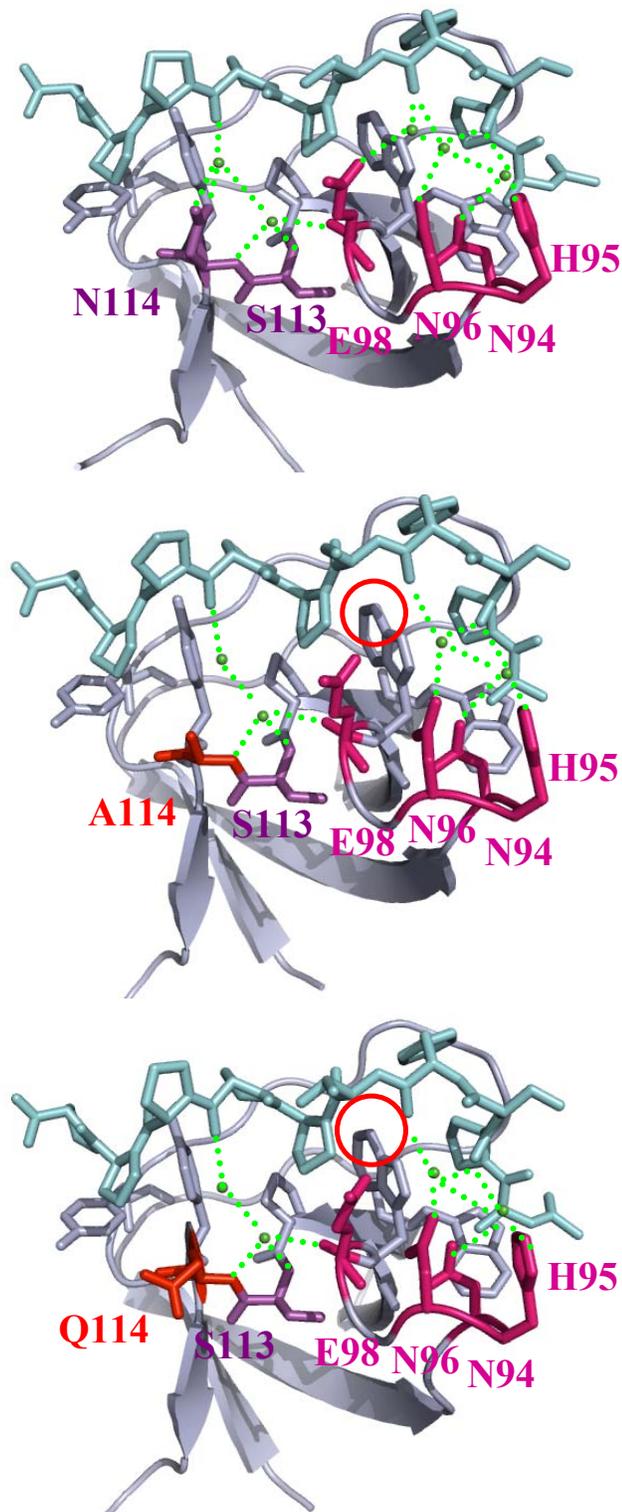
## **Resultados**

Por último, también se desprende de este estudio estructural que la variabilidad conformacional descrita para el dominio libre desaparece en las estructuras de los complejos con p41, de modo que ambos lazos, RT y n-Src, conformacionalmente flexibles en los dominios libres, se estabilizan en una única conformación en los complejos que es independiente de las condiciones de cristalización. De este modo, los valores de RMSD para los carbonos alfa entre las distintas estructuras de los complejos son aún más pequeños y varían entre 0.34 y 0.58 Å entre sí.

### **3.1.1.9. Cuatro de las cinco moléculas de agua interfaciales observadas en Abl-SH3 se conservan en las estructuras cristalográficas de los mutantes N114 unidos a p41**

Si nos centramos en la zona de la interfase de unión, su configuración tampoco se ve alterada por las mutaciones en posición N114 (Figura 3.15). Las conformaciones, tanto del péptido p41 como de los restos del dominio Abl-SH3 que definen el sitio de unión canónico, están muy conservadas en los tres complejos. Además, con la excepción de una ligera reorientación de la cadena lateral del resto E98 en la estructura N114Q/p41, las cadenas laterales que coordinan las moléculas de agua también adoptan orientaciones muy similares en todas las estructuras.

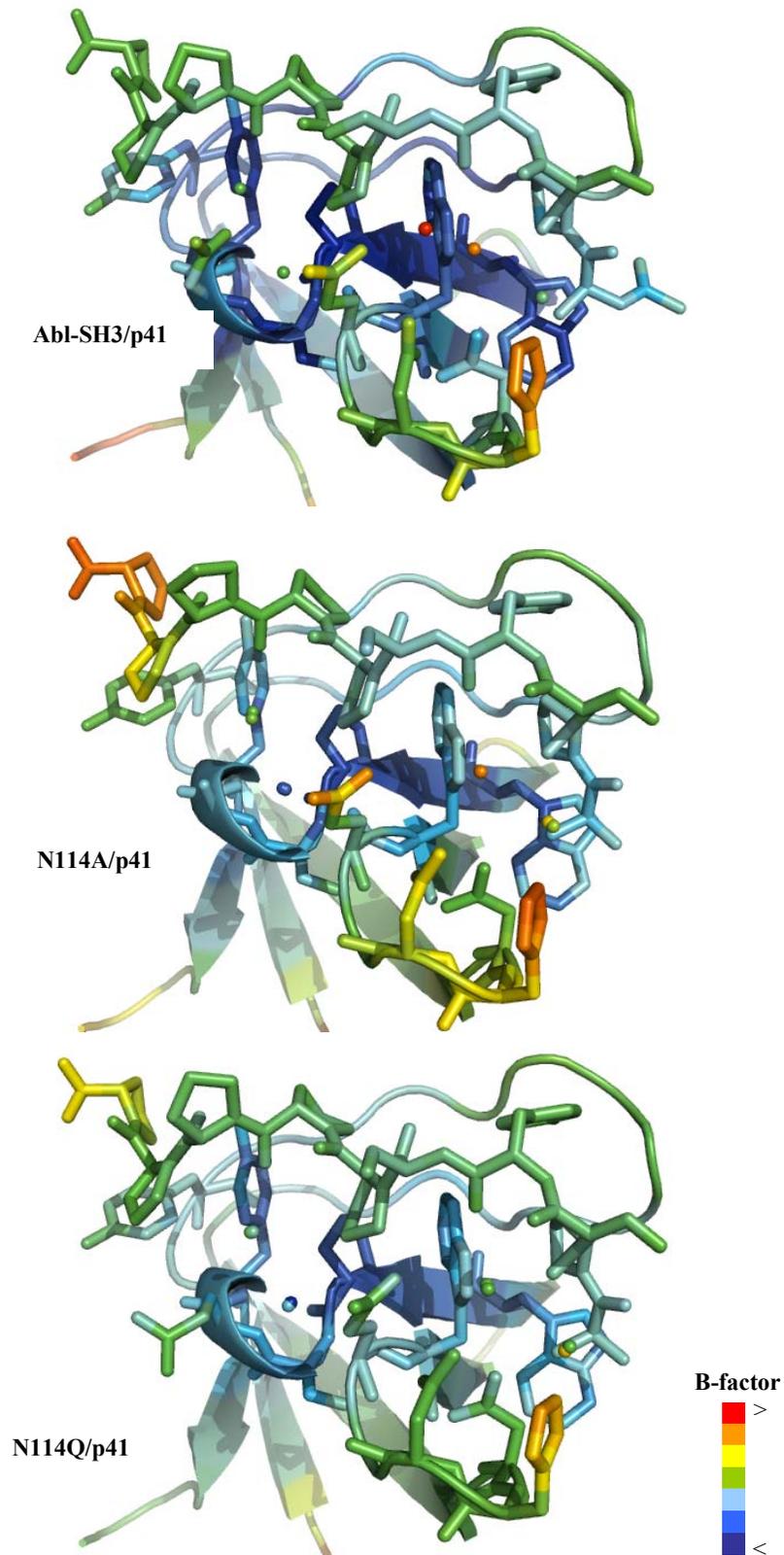
En cuanto a las moléculas de agua ocluidas en la interfase de unión, en las estructuras N114A/p41 y N114Q/p41 se observan moléculas de agua en posiciones equivalentes a las descritas para el complejo con el dominio silvestre, con la excepción de la molécula de agua en posición 3 (numeración según la Tabla 3.3 y Figura 3.4) que no se observa en ninguna de las estructuras de los mutantes libres. En la Tabla 3.10 se recogen las identidades y las áreas accesibles al disolvente de las distintas moléculas de agua observadas en las tres estructuras. En la Figura 3.16 se muestran las estructuras coloreadas según el valor de los factores B, de modo que las zonas más definidas y con menor movilidad (factores B bajos) se muestran en azul y las zonas con mayor movilidad (factores B altos) en rojo. Como puede observarse, en los tres complejos las moléculas de agua en posiciones 1, 2, 4 y 5 presentan factores B similares a los de los restos en el péptido y en el dominio SH3 que las coordinan, como es característico de aguas inmovilizadas {Park, 2005



**Figura 3.15.** Estructuras cristalográficas de los complejos formados con el péptido p41 y la proteína silvestre Abl-SH3 (arriba, PDB: 1bbz, (Pisabarro et al. 1998)); el mutante N114A (centro, PDB: 2o88); y el mutante N114Q (abajo, PDB: pendiente de depositar). El péptido p41 se ha representado en color turquesa y la proteína en azul claro. Los restos resaltados en varillas se corresponderían con los aminoácidos aromáticos que componen el sitio de unión de Abl-SH3 en azul claro, los restos implicados en interacciones con las moléculas de agua en el lazo n-Src en fucsia, y los respectivos de la hélice 3<sub>10</sub> en violeta. Los aminoácidos introducidos en la posición 114 se muestran en color rojo, y las moléculas de agua enterradas en la interfase de unión en color verde. La hipotética localización de la molécula de agua 3, ausente en ambos mutantes, se indica con un círculo en rojo y los puentes de hidrógeno establecidos por las moléculas de agua se representan como líneas verdes discontinuas.

*Tabla 3.10: Factores B y áreas accesibles al disolvente de las moléculas de agua en la interfase de unión de los complejos Abl-SH3/p41, N114A/p41, N114Q/p41 y completa de Abl.*

Aguas #	Abl-SH3 (a/b) <sup>1</sup>	Abl- SH3 (c/d) <sup>1</sup>	Abl- SH3 (e/f) <sup>1</sup>	Abl- SH3 (g/h) <sup>1</sup>	N114A (a/c) <sup>2</sup>	N114A (b/d) <sup>2</sup>	N114Q (a/c) <sup>3</sup>	N114Q (b/d) <sup>3</sup>	Quinasa
1	<b>w1015</b> 20.91 (0.03 Å <sup>2</sup> )	<b>w1036</b> 24.15 (1.12 Å <sup>2</sup> )	<b>w1019</b> 20.66 (3.76 Å <sup>2</sup> )	<b>w2036</b> 27.36 (2.85 Å <sup>2</sup> )	<b>W29</b> 28.22 (7.55 Å <sup>2</sup> )	<b>W45</b> 32.03 (15.53 Å <sup>2</sup> )	<b>W48</b> 32.68 (6.34 Å <sup>2</sup> )	<b>W12</b> 24.33 (10.18 Å <sup>2</sup> )	
2	<b>w1064</b> 36.30 (0.00 Å <sup>2</sup> )	<b>w1025</b> 22.04 (0.00 Å <sup>2</sup> )		<b>w1028</b> 32.24 (0.00 Å <sup>2</sup> )	<b>W55</b> 35.82 (5.93 Å <sup>2</sup> )		<b>w14</b> 24.83 (0.15 Å <sup>2</sup> )	<b>W36</b> 24.83 (0.02 Å <sup>2</sup> )	<b>w170</b> 35.04 (14.09 Å <sup>2</sup> )
3	<b>w2103</b> 44.07 (5.55 Å <sup>2</sup> )	<b>w2016</b> 47.21 (16.73 Å <sup>2</sup> )		<b>w1003</b> 37.88 (8.93 Å <sup>2</sup> )					
4	<b>W1082</b> 23.12 (2.37 Å <sup>2</sup> )	<b>w1067</b> 15.90 (4.89 Å <sup>2</sup> )	<b>w1097</b> 22.20 (0.00 Å <sup>2</sup> )	<b>w1001</b> 14.46 (6.25 Å <sup>2</sup> )	<b>w4</b> 14.16 (8.62 Å <sup>2</sup> )	<b>w3</b> 14.36 (7.00 Å <sup>2</sup> )	<b>w1</b> 10.10 (2.35 Å <sup>2</sup> )	<b>w5</b> 16.90 (2.39 Å <sup>2</sup> )	<b>w35</b> 27.50 (9.91 Å <sup>2</sup> )
5	<b>w1105</b> 22.56 (0.00 Å <sup>2</sup> )	<b>w1060</b> 21.70 (0.00 Å <sup>2</sup> )	<b>w1061</b> 13.32 (0.00 Å <sup>2</sup> )	<b>w1018</b> 21.74 (0.54 Å <sup>2</sup> )	<b>w9</b> 25.61 (8.42 Å <sup>2</sup> )	<b>w6</b> 26.74 (7.82 Å <sup>2</sup> )	<b>w7</b> 19.79 (4.90 Å <sup>2</sup> )	<b>w28</b> 21.66 (4.09 Å <sup>2</sup> )	<b>w44</b> 27.20 (5.06 Å <sup>2</sup> )
6			<b>w1089</b> 22.64 (0.00 Å <sup>2</sup> )						
Factor B rango para restos de la proteína	Min: 4.62 Max: 44.68	Min: 5.29 Max: 41.37	Min: 5.30 Max: 43.28	Min: 5.42 Max: 45.71	Min: 8.36 Max: 47.00	Min: 8.95 Max: 48.09	Min: 9.64 Max: 44.27	Min: 8.13 Max: 45.4	Min: 13.6 Max: 75.01



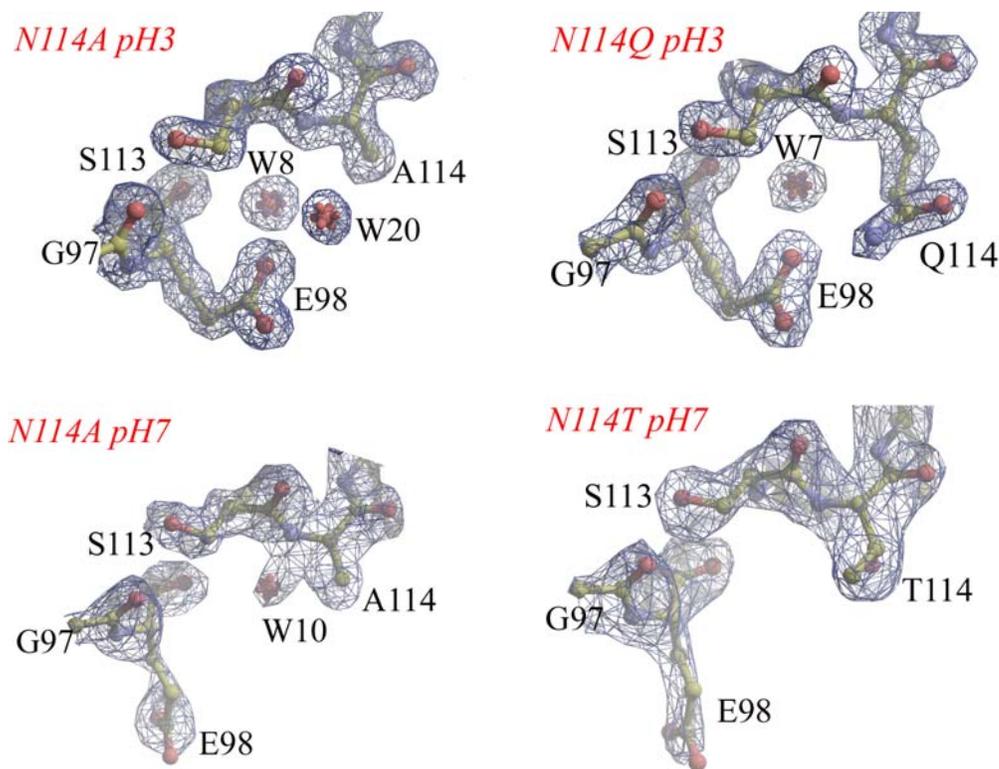
**Figura 3.16.** Representación de las estructuras cristalográficas de los complejos Abl-SH3/p41, N114A/p41 y N114Q/p41. Las moléculas de agua en la interfase de unión se representan como esferas. Los átomos de cada complejo y las moléculas de agua están coloreados en función de su valor de factor B en la estructura cristalográfica, donde el color rojo marca el valor máximo y el azul el valor mínimo. Para la realización de las figuras no se ha empleado el algoritmo TLS para la normalización de los factores B (Winn et al. 2001).

## **Resultados**

#120}. Es interesante señalar que la molécula de agua 3 está considerablemente más expuesta al disolvente que las otras cuatro moléculas de agua en el complejo silvestre y además está caracterizada por un factor B significativamente mayor que los restos de su alrededor, lo que indica que sus interacciones en la interfase de unión son probablemente más débiles y justificaría su ausencia en los complejos con los mutantes.

Es interesante señalar que la presencia de la mayoría de las moléculas de agua en la interfase de unión parece ser concomitante con la unión del ligando p41 puesto que sólo la molécula de agua en posición 4, que se encuentra fundamentalmente coordinada por los átomos del esqueleto peptídico de los restos E98, S113 y N114 es la única que se encuentra de forma sistemática en las estructuras de los dominios libres. En la Figura 3.17 mostramos a modo de ejemplo cómo esta molécula de agua está claramente definida en las estructuras de los mutantes N114A y N114Q obtenidas a pH 3.0 e incluso se puede modelar adecuadamente en la estructura de N114A obtenida a pH 7.0. Sin embargo, esta molécula no se observa en las estructuras obtenidas a pH neutro para el dominio silvestre (Pdb: 1abq) o en el mutante N114T, posiblemente a causa de la menor resolución ( $> 2.2 \text{ \AA}$ ) de las estructuras cristalográficas en estas condiciones. En cualquier caso, en las estructuras en las que se observa, esta molécula está caracterizada por factores B de magnitud comparable a los que caracterizan a la molécula de agua equivalente en las diferentes estructuras con p41. Estas observaciones parecen indicar que la molécula de agua en posición 4 forma parte estructural del dominio libre y, por tanto, debería ser tomada en cuenta para la definición del sitio de unión, lo que sería vital en cualquier estrategia para el diseño racional de ligandos.

Otro aspecto destacable es que el patrón de interacciones mediadas por moléculas de agua está también muy conservado entre las estructuras de los complejos con la proteína silvestre y los dos mutantes, de modo que los efectos de las sustituciones en posición N114 parecen estar restringidas exclusivamente al lugar de la mutación. Así, en N114A/p41 se observa que los dos puentes de hidrógeno establecidos entre la cadena lateral del resto N114 y la molécula de agua 5 desaparecen, de modo que esta molécula de agua se mantiene en su posición en



**Figura 3.17.** Densidad electrónica de la molécula de agua 4 y de los restos más cercanos en los mutantes libres N114A, N114Q y N114T, a varios valores de pH. La molécula de agua en posición 4 corresponde con la 8 en la estructura cristalográfica de N114A a pH 3.0, con la 10 a pH 7, y con la 7 en N114Q. En el caso de la estructura de N114A a pH 3.0, estructura con mejor resolución (1.3 Å), también aparece muy bien definida la molécula de agua 5 (W20). Como excepción, el mutante N114T, que fue obtenido a menor resolución (2.2 Å), no presenta la molécula de agua 4.

la interfase de unión sólo mediante dos enlaces de hidrógeno, uno establecido con el oxígeno carbonilo de P7 en el ligando y el otro con la molécula de agua en posición 4. Esta situación estaría en buena concordancia con los efectos termodinámicos derivados para esta mutación, ya que se traduciría en una disminución significativa de las contribuciones entálpicas a la unión de p41 como efectivamente ocurre (Figura 3.11).

Sin embargo, la interpretación de los resultados termodinámicos puede ser más complicada para el caso de N114Q/p41. En este caso observábamos un incremento en la contribución entálpica favorable (Figura 3.11) lo que apuntaría, por

## **Resultados**

tanto, hacia el establecimiento de interacciones similares a las observadas en el complejo silvestre o incluso a una mayor optimización de éstas. Sorprendentemente, en la estructura N114Q/p41 la cadena lateral del propio resto Q114 se encuentra orientada hacia el disolvente y hacia afuera del sitio de unión, adoptando una conformación que parece estar estabilizada por el establecimiento de contactos cristalográficos con el resto E98 de otra molécula de complejo en el cristal. Dado que la orientación de Q114 en la estructura parece ser debida a un efecto de la cristalización, no podemos extraer información concluyente acerca de las interacciones que establecería realmente en la interfase de unión cuando se encuentra en disolución.

**3.1.1.10. La dinámica molecular de las moléculas de agua interfaciales permite explicar los efectos termodinámicos observados en la interacción de p41 con Abl-SH3 y sus mutantes N114A y N114Q**

La información estructural obtenida para las distintas variantes del dominio Abl-SH3 y sus complejos ha puesto en evidencia la persistencia de cuatro de las cinco moléculas de agua en la interfase de unión y la robustez de las interacciones que establecen con el ligando y la proteína. Sin embargo, esta información no parece suficiente para interpretar adecuadamente su contribución a la energética de unión y su verdadero papel en la interacción. Es importante considerar que las proteínas y sus interacciones en disolución son sistemas complicados que, lejos de constituir un conjunto estático, presentan importantes propiedades dinámicas que pueden ser determinantes en su comportamiento (Helms 2007; Tarek & Tobias 2002). Por esta razón, hemos considerado interesante complementar la información estática obtenida de las estructuras cristalográficas con un estudio en profundidad de la dinámica molecular de esta interacción en disolución. Estos cálculos fueron realizados en colaboración con la Dra. María Teresa Pisabarro Díez (Instituto BioTec-Universidad Técnica de Dresden, Alemania).

Para ello hemos llevado a cabo simulaciones de dinámica molecular con las estructuras Abl-SH3/p41, N114A/p41 y N114Q/p41 de 12 ns de duración, lo que se considera un tiempo razonable para el análisis de las propiedades dinámicas de las moléculas de agua y de los complejos (Damjanovic *et al.* 2005; McCammon & Harvey 1997; Petrone & Garcia 2004). Este muestreo extensivo nos ha permitido investigar las interacciones, amplitud de las fluctuaciones y tiempos de residencia para las distintas moléculas de agua en la interfase de unión, así como los mecanismos de intercambio entre dicha interfase y el disolvente.

Las configuraciones iniciales para las dinámicas se basaron en las estructuras cristalográficas obtenidas, incluyendo aquellas moléculas de agua que se encuentran enterradas en la interfase de unión y mediando las interacciones entre el péptido p41 y el dominio SH3 en estas estructuras (moléculas 1 a 5 en el complejo silvestre y moléculas 1, 2, 4 y 5 en los mutantes). Es importante señalar que, en los tres casos, estas moléculas de agua cristalográficas son reemplazadas

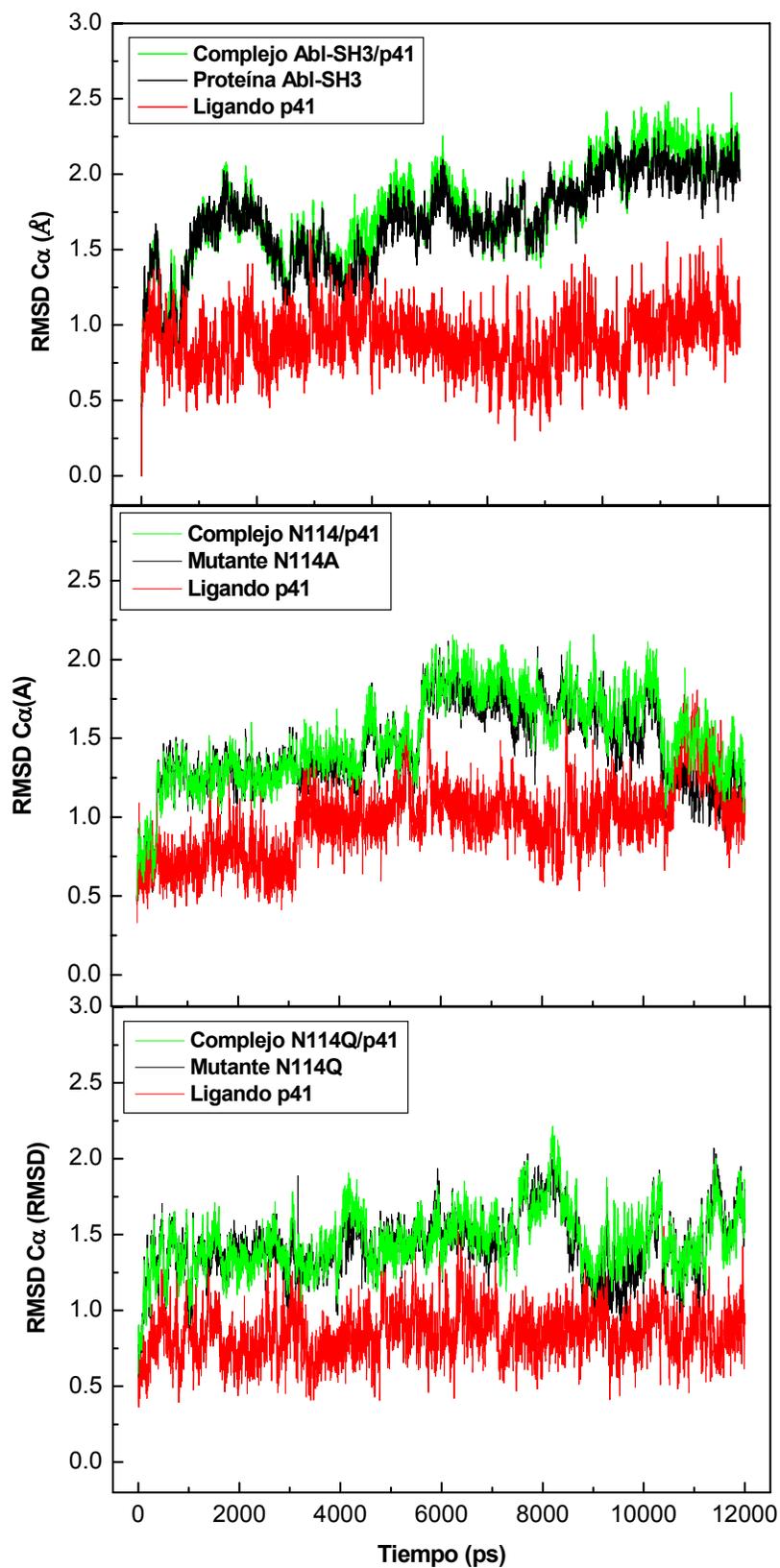
## **Resultados**

espontáneamente en etapas muy tempranas de la simulación por otras moléculas que penetran a la interfase desde el disolvente, las cuales ocuparían posiciones equivalentes a las anteriores. Así, aunque con distintas frecuencias, todas las moléculas en la interfase de unión se encuentran en un intercambio continuo y espontáneo con el disolvente a lo largo de toda la trayectoria de dinámica molecular.

En diferentes momentos de las simulaciones y a su finalización hemos comprobado que la integridad de las estructuras se mantiene y que las posiciones atómicas permanecen próximas a las que ocupan en las estructuras de partida. Una prueba de esto se recoge en la Figura 3.18 donde se muestra la evolución temporal para la desviación cuadrática media (RMSD) de los carbonos alfa para cada una de las simulaciones, tanto en el caso del complejo como para la proteína y el ligando individualmente. Como se puede observar, en todo momento el valor de RMSD es inferior a 2 Å, lo que se considera un valor aceptable considerando el tamaño de nuestro sistema (8 KDa) (Karplus & McCammon 2002; McCammon & Harvey 1997). Esta comprobación es importante, pues de este modo podremos asegurar que se cumple la hipótesis ergódica según la cual la evolución conformacional que se observa a lo largo del tiempo para una sola molécula refleja las propiedades conformacionales de una distribución molecular, única situación en la que estas simulaciones nos permitirán extraer conclusiones sobre el sistema en equilibrio.

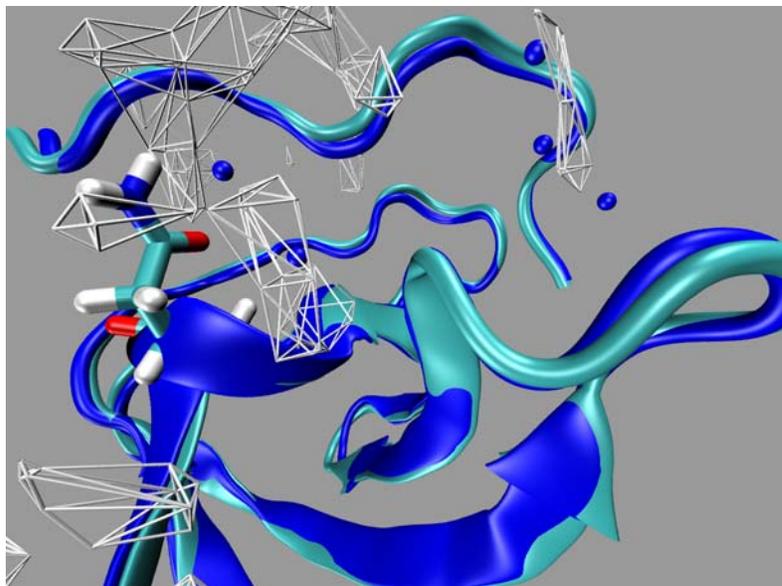
Para una evaluación inicial cualitativa de las trayectorias hemos generado los mapas de densidad de hidratación, que proporcionan información sobre las regiones en la interfase de unión en las que se observa la mayor probabilidad de encontrar una molécula de agua durante la simulación. En líneas generales, en las tres estructuras se observa que las posiciones ocupadas por las moléculas de agua son zonas con gran densidad de hidratación. No obstante, este estudio no permite comparar de un modo cuantitativo la densidad de hidratación entre los diferentes complejos. En la Figura 3.19 se muestra como ejemplo el mapa de densidad de hidratación tridimensional obtenido para Abl-SH3/p41.

Es también destacable que este análisis dinámico identifica claramente las dos regiones de hidratación descritas previamente a nivel estructural (Figuras 3.4 y



**Figura 3.18.** Representación de la desviación cuadrática media (RMSD) de los carbonos alfa durante la simulación para los complejos Abl-SH3/p41 (panel superior), N114A/p41 (panel central) y N114Q/p41 (panel inferior). La línea verde representa la rmsd para el complejo, la línea roja para el ligando, y la línea negra para la proteína libre.

## Resultados



**Figura 3.19.** Representación del mapa de densidad de hidratación calculado para el complejo Abl-SH3/p41. En la figura se ha superpuesto la estructura promedio durante la simulación (color turquesa) con la estructura cristalográfica original (color azul). El mapa de densidad de hidratación se ha representado en color gris. Las aguas observadas en las estructuras cristalográficas se han representado como esferas en color azul, y el resto N114, donde se introducen las mutaciones, el cual interacciona con las moléculas de agua 4 y 5, en estructura de varillas.

3.5), donde en una zona tendríamos las aguas 1, 2 y 3 de la estructura cristalográfica del complejo silvestre que median las interacciones entre el lazo n-Src y la región de especificidad del ligando p41; y en la otra zona localizaríamos las aguas 4 y 5 de la estructura del complejo silvestre que coordinan las interacciones entre los restos de la hélice  $3_{10}$  y la región del ligando que contiene la secuencia rica en prolina. Las dinámicas nos muestran que las aguas de ambas zonas se comportan independientemente durante la simulación y además presentan diferentes mecanismos de entrada y salida de las superficies de unión. Por consiguiente, nos referiremos a estas zonas o regiones de forma independiente.

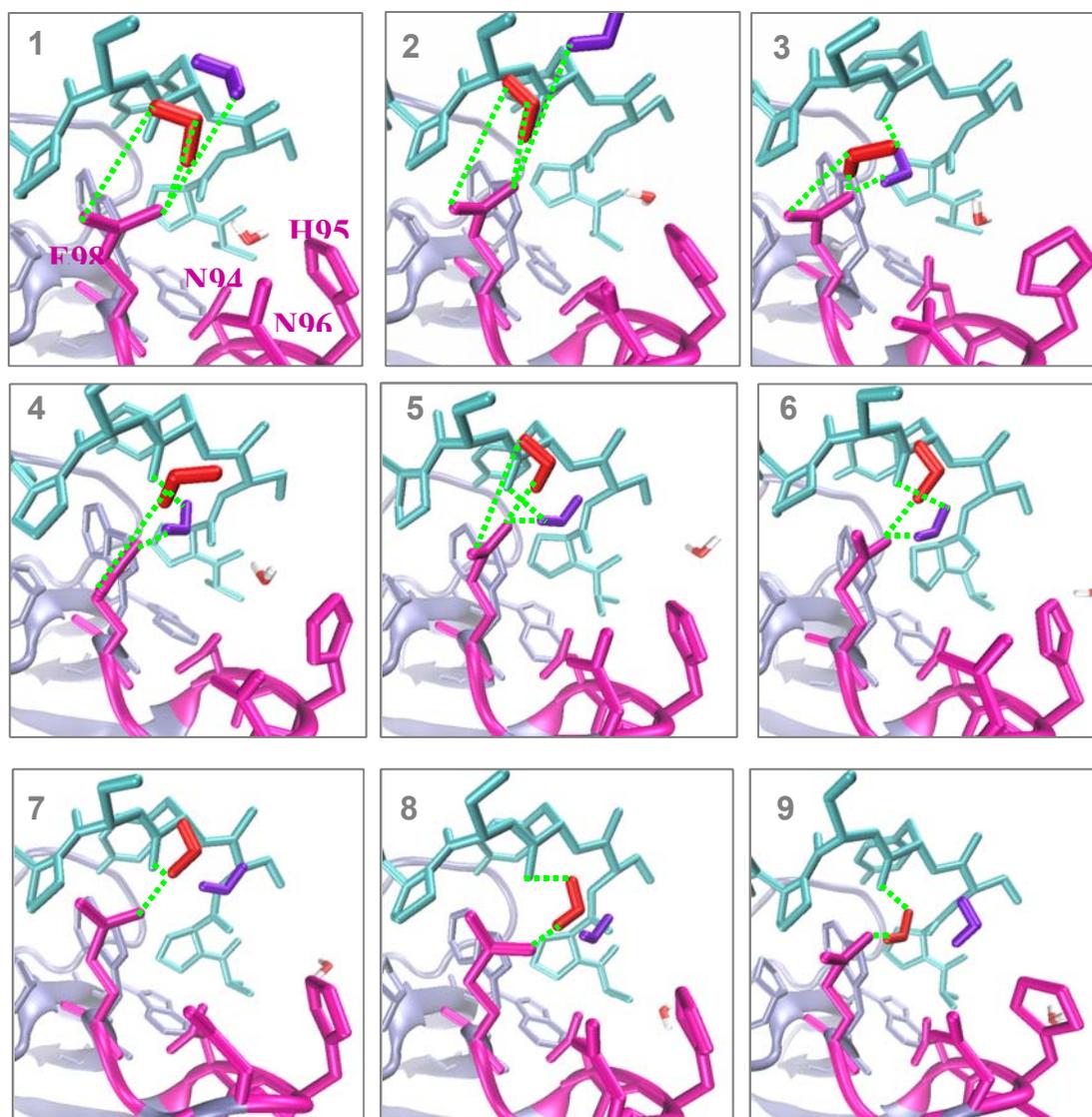
Para poder realizar un estudio cuantitativo de los efectos de las mutaciones en el comportamiento de las moléculas de agua y evaluar el papel que cada una de ellas juega en la interfase de unión, hemos calculado los tiempos de residencia de las interacciones establecidas por las moléculas de agua que ocupan los distintos sitios de hidratación. Desde este punto de vista, las dos regiones de hidratación

descritas estarían constituidas por cinco sitios (tres y dos respectivamente), que presentan diferentes propiedades en cuanto al porcentaje de tiempo que cada uno está ocupado por moléculas de agua (ocupancia) y a los tiempos de residencia y propiedades dinámicas características de las moléculas de agua que los ocupan, cuyos valores se resumen en la Tabla 3.11 para el complejo de la proteína silvestre y los mutantes N114A y N114Q con p41.

La **primera región** contiene, por tanto, dos sitios de hidratación principales y bien definidos (**posiciones 1 y 2**) que están ocupados por moléculas de agua durante más del 90% del tiempo de simulación para el caso Abl-SH3/p41. El reemplazo de moléculas de agua en estas dos posiciones está asistido por las interacciones establecidas por la cadena lateral del resto E98 en el lazo n-Src, que presenta una gran movilidad y que está cambiando de conformación continuamente a lo largo de la simulación, orientándose alternativamente hacia el interior o el exterior del sitio de unión y dirigiendo a las moléculas de agua con las que interacciona desde el disolvente hacia esta región interior, en concreto hacia la posición de hidratación 2 (Figura 3.20). Las moléculas de agua en esta región están, por tanto, en un intercambio continuo y relativamente rápido con el disolvente, de modo que durante los 12 ns observamos que un elevado número de moléculas de agua ocupa las posiciones 1 y 2 con tiempos de residencia promedio bajos.

Así, la **posición 1** es ocupada por un total de 98 moléculas de agua que permanecen en la interfase de unión durante un tiempo de residencia promedio de 100 ps, aunque encontramos una molécula con un tiempo de residencia máximo de 340 ps (Figura 3.21). El valor de RMSD calculado con respecto a la posición inicial del agua en la simulación es de 1.8 Å, lo que indica que es un agua que está fuertemente unida a la proteína, tal y como indican otros estudios anteriores en los que se describe que valores en torno a 2 Å son característicos de aguas enterradas en cavidades cuando son importantes para la estructura o interacciones que establece la proteína (Paschek *et al.* 2007).

## Resultados



**Figura 3.20.** Esquema secuencial de la asistencia de la cadena lateral de Glu98 para la entrada de aguas en las posiciones 2 y 3, tomado de las simulaciones de dinámica molecular. En las imágenes 1 a 3 se puede observar como la cadena lateral de E98 coordina dos moléculas de agua, una que ocupa la posición 3 (en rojo) y otra que llega a ocupar la posición 2 (violeta). En las imágenes 4 a 6 la molécula de agua 3 se desplaza hacia dentro de la interfase de unión situándose más cerca del agua 2. Finalmente, en las imágenes 7 a 9 la molécula de agua 3 pasa a ocupar la posición 2, desplazando el agua que había anteriormente. El resto E98 y otros restos que interaccionan con las moléculas de agua del grupo 1 se muestran en estructura de varillas en rosa, y los puentes de hidrógeno como líneas verdes discontinuas.

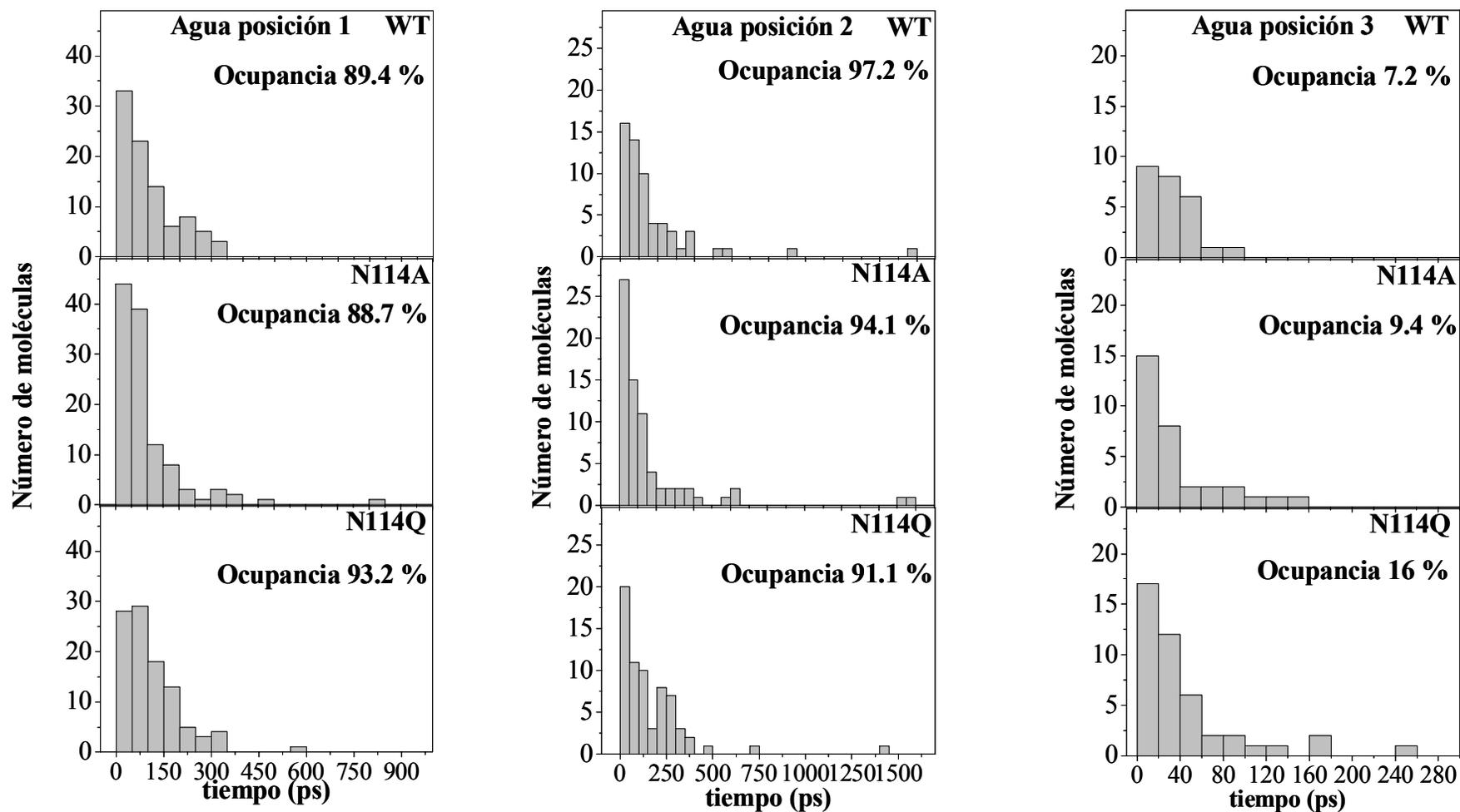
**Tabla 3.11.** Ocupancias, tiempos de residencia, RMSD promedio y factor B de las moléculas de agua en las cinco localizaciones de la interfase de unión de Abl-SH3/p41, N114A/p41 y N114Q/p41.

Posición	Complejo	Ocupancia <sup>1</sup> (%)	N <sup>o2</sup> moléc.	N <sup>o2</sup> (>500 ps)	N <sup>o2</sup> (>1000ps)	t <sub>máx</sub> <sup>3</sup> (ps)	<t> <sup>3</sup> (ps)	RMSD (Å)
<b>1</b>	Abl-SH3	89.4	92	0	0	340	100±85	1.80±2.34
	N114A	88.7	114	1	0	834	93±109	1.54±1.71
	N114Q	93.2	101	1	0	568	110±92	-
<b>2</b>	Abl-SH3	97.2	59	4	1	1584	171±250	1.33±0.99
	N114A	94.1	71	5	2	1560	159±273	1.31±0.78
	N114Q	91.1	67	2	1	1422	163±204	-
<b>3</b>	Abl-SH3	7.2	25	0	0	82	29±20	1.72±1.25
	N114A	9.4	32	0	0	148	35±39	1.86±1.21
	N114Q	16	44	0	0	256	44±51	-
<b>4</b>	Abl-SH3	97.6	12	4	4	2712	844±970	0.71±0.47
	N114A	97.2	9	5	5	3500	1300±1200	0.77±0.48
	N114Q	99.8	12	7	6	3332	997±887	-
<b>5</b>	Abl-SH3	79.8	46	2	1	1214	175±223	1.15±0.74
	N114A	54.3	88	0	0	386	74±71	1.66±2.84
	N114Q	93.3	168	0	0	382	66±63	-

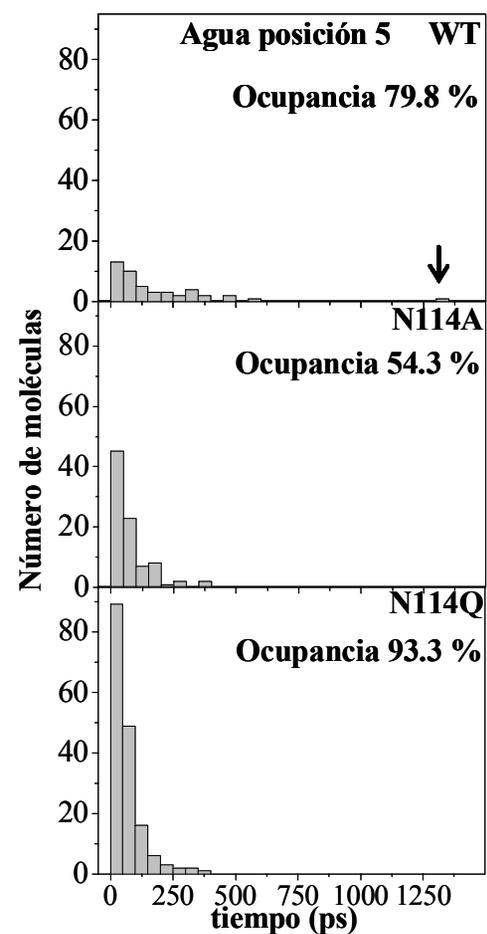
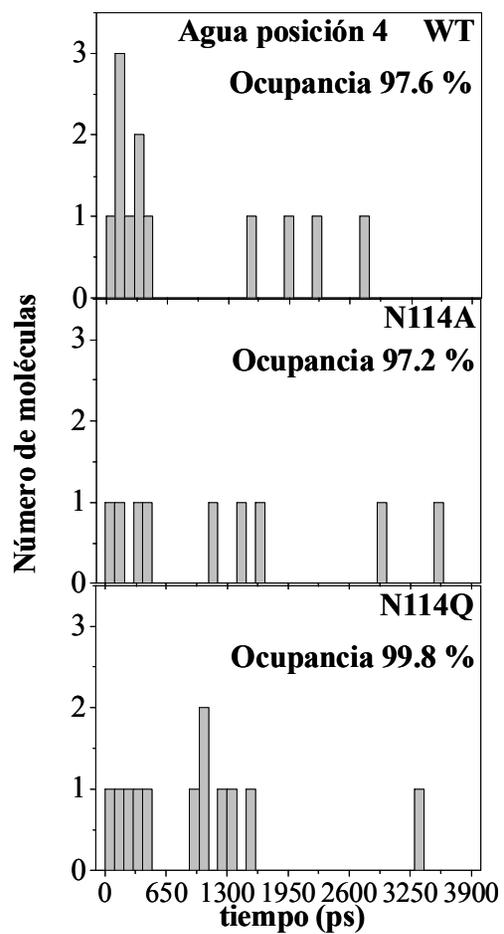
<sup>1</sup>La ocupancia corresponde con el tanto por ciento del tiempo total de la dinámica que una molécula de agua ocupa una posición.

<sup>2</sup>N es el número de moléculas que ocupan una posición determinada durante la simulación; N (> 500 ps) y N (>1000 ps) son aquellas con tiempos de residencia mayores de 500 y 1000 ps respectivamente.

<sup>3</sup>T<sub>max</sub> es el tiempo de residencia máximo y <t> el tiempo de residencia promedio para una determinada molécula de agua.



*Figura 3.21. Distribución del número de moléculas de agua en función de su tiempo de residencia para cada una de las posiciones de hidratación encontradas en la interfase de unión de Abl-SH3/p41, N114A/p41 y N114Q/p41. La ocupancia total para cada una de estas posiciones también se indica en el panel correspondiente.*



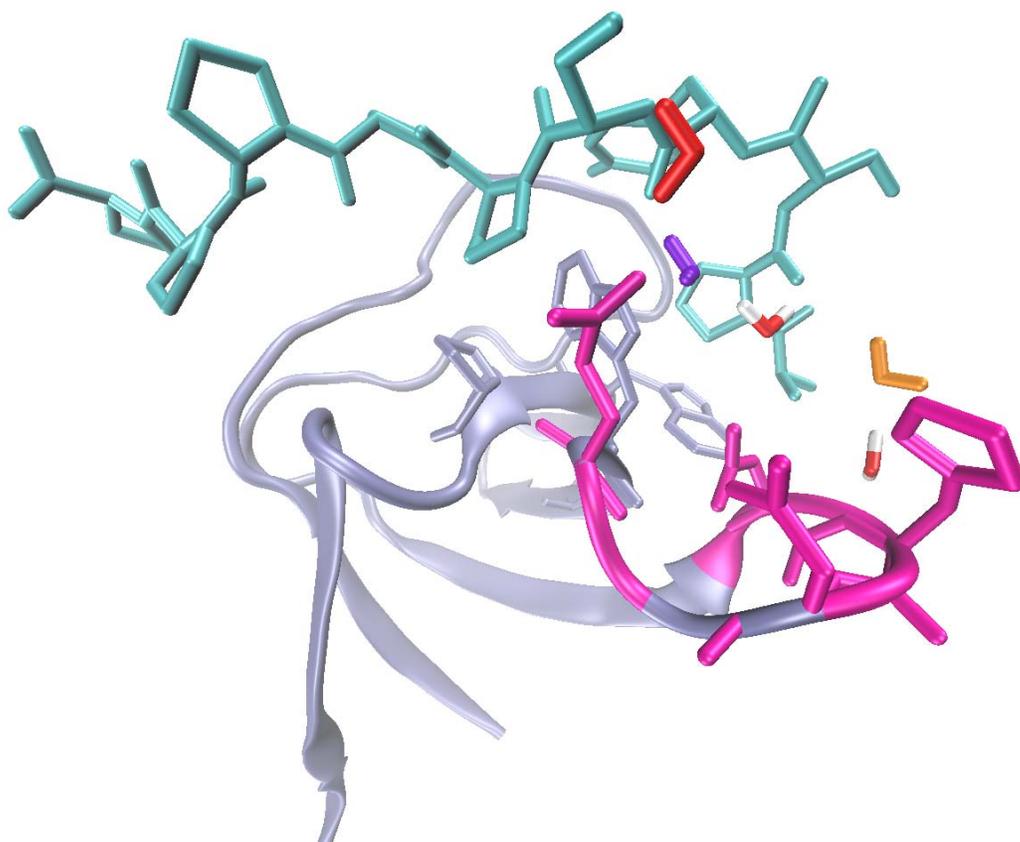
## Resultados

La **posición 2** se ocupa por 59 moléculas con un tiempo de residencia promedio de 171 ps y un valor de RMSD de 1.33 Å. En este caso, cuatro de estas moléculas presentan tiempos de residencia mayores de 500 ps, y una de ellas permanece en esta posición durante un tiempo superior a 1.5 ns (Figura 3.21), lo que indica que en esta posición las moléculas de agua están unidas más fuertemente al sitio de unión que las de la posición 1. Es destacable que las dinámicas muestran que, además de estos dos sitios de hidratación principales, esta región presenta otras posiciones que están ocupadas por moléculas de agua con menor frecuencia. Por ejemplo, la que denominamos posición 3 (Figuras 3.4 y 3.5) está ocupada por moléculas de agua sólo durante un 10% del tiempo de simulación en Abl-SH3/p41 y todas las moléculas de agua que ocupan esta posición presentan tiempos de residencia por debajo de 100 ps. Este comportamiento está de acuerdo con los elevados valores de accesibilidad al disolvente y de factor B que esta posición 3 presenta en la estructura Abl-SH3/p41, la única en la que se encuentra ocupada en algunas moléculas de la unidad asimétrica, excepto la e/f, mientras que en los otros casos, N114A/p41 y N114Q/p41, no llega ni a observarse ocupada. Ocasionalmente y de forma transitoria también se observan moléculas de agua en otras posiciones, de modo que en algunos momentos esta región puede llegar a contener hasta cinco sitios ocupados simultáneamente por moléculas de agua (Figura 3.22).

En la Tabla 3.11 hemos recogido las ocupancias, amplitud de fluctuaciones y tiempos de residencia para todas las moléculas en esta primera región. Observamos que estos valores no se ven afectados por las mutaciones introducidas en el resto 114, lo que refleja el comportamiento independiente de las dos regiones de hidratación.

La **segunda región** está constituida por dos sitios de hidratación muy bien definidos, correspondientes a las **posiciones de hidratación 4 y 5** en la estructura cristalográfica (Figuras 3.4 y 3.5), y que como hemos anticipado presentan propiedades dinámicas muy diferentes a las de la primera región.

En el caso de la **posición 4**, en la dinámica con Abl-SH3/p41 se observan únicamente 12 moléculas diferentes durante los 12 ns, con tiempos de residencia



**Figura 3.22.** Imagen tomada de un momento de la simulación de dinámica molecular con Abl-SH3/p41 donde llegan a situarse hasta un total de 5 moléculas de agua en la región primera. Las moléculas de agua 1, 2 y 3, se muestran en color naranja, violeta y rojo, respectivamente, y las dos moléculas de agua adicionales que ocupan posiciones intermedias entre las aguas 1 y 2 en código universal de colores. Los restos de la proteína que interaccionan con las moléculas de agua se muestran en estructura de varillas y en color rosa.

altos que mantienen este sitio ocupado durante casi la totalidad del tiempo total de la dinámica (98%). Así, el 83 % de la ocupancia se debe a la presencia de sólo 4 moléculas de agua caracterizadas por tiempos de residencia entre 1.5 y 2.7 ns y, además, se observan otras cuatro moléculas de agua con tiempos de residencia mayores de 0.5 ns. El tiempo de residencia promedio en este sitio de hidratación es 850 ps y el RMSD promedio 0.7 Å, lo que nos indica sin duda que las moléculas de agua en esta posición presentan muy baja movilidad y se encuentran fuertemente unidas a la proteína. A partir de los datos de la simulación, hemos calculado el factor B teórico para las distintas moléculas en este sitio de hidratación, lo que proporciona una indicación de su movilidad en la interfase de unión (Lounnas

## Resultados

& Pettitt 1994a; Lounnas & Pettitt 1994b). Para esta posición hemos obtenido un valor de factor B muy bajo ( $17 \text{ \AA}^2$ ), similar al que presentan los restos del dominio con los que esta molécula de agua interacciona y en muy buena concordancia con el factor B cristalográfico ( $18 \text{ \AA}^2$ ).

Las **posición 5** es ocupada por moléculas de agua durante un 80% del tiempo de simulación en el complejo silvestre, se intercambian con mayor frecuencia con el disolvente y presentan tiempos de residencia significativamente menores que la anterior, de modo que el tiempo de residencia promedio para las 46 moléculas de agua observadas en esta posición es 175 ps y sólo una molécula permanece durante un periodo mayor de 1 ns. Estas moléculas se caracterizan por un valor de RMSD de  $1.2 \text{ \AA}$  y un factor B teórico, calculado como promedio para todas aquellas moléculas de agua con tiempos de residencia mayores de 150 ps, de  $26 \text{ \AA}^2$ , también muy similar al obtenido en la estructura cristalográfica ( $20 \text{ \AA}^2$ ) y al de los restos de la proteína y el ligando que las coordinan. Estos datos indican que, aunque estas moléculas de agua están unidas fuertemente al complejo, presentan una movilidad mayor que las presentes en posición 4, probablemente debido a la mayor flexibilidad conformacional de los restos que las coordinan.

Al contrario que lo observado para los sitios de la primera región, las mutaciones introducidas en posición N114 producen importantes efectos en el comportamiento y las propiedades dinámicas de las moléculas de agua en la segunda región, especialmente en las moléculas en posición 5 que interaccionan directamente con los átomos de la cadena lateral de la Asn mutada. Los efectos en la posición 4 son muy sutiles y, como veremos a continuación, están relacionados con los cambios en el comportamiento de las moléculas de agua en posición 5. Así, el número de moléculas de agua en posición 4 (9 y 12 moléculas para N114A/p41 y N114Q/p41 respectivamente), y sus ocupancias cercanas al 100% en los dos casos, son muy similares a los obtenidos para Abl-SH3/p41 (Tabla 3.11 y Figura 3.21). El tiempo de residencia promedio es algo superior en el caso de N114A/p41 que en N114Q/p41 y Abl-SH3/p41. De igual forma, los factores B calculados para esta posición en N114A/p41 y N114Q/p41, fueron de 18 y  $17 \text{ \AA}^2$  respectivamente, similares a los del complejo silvestre y cercanos a los cristalográficos ( $14 \text{ \AA}^2$  para ambos mutantes).

Para las moléculas de agua en posición 5, que según la estructura cristalográfica establecen dos puentes de hidrógeno con la cadena lateral de N114, la eliminación de la cadena lateral en el mutante N114A tiene como resultado inmediato un descenso notable en el porcentaje de tiempo que esta posición está ocupada por moléculas de agua, que desciende desde el 80% hasta un 54%. Esta pérdida de interacciones también afecta al tiempo que cada molécula permanece en la interfase de unión, de forma que se observa también una disminución importante en los tiempos de residencia, en todos los casos inferiores a 400 ps, y un aumento en los valores de RMSD (que pasa de 1.5 Å en el complejo silvestre a 1.66 Å) y del factor B, tanto teórico como experimental (que pasan de 26 y 20 Å para el valor teórico y experimental respectivamente en Abl-SH3/p41 a 28 y 25 Å en N114A/p41). Todos estos datos indican, por tanto, un debilitamiento de las interacciones que concuerda con los efectos termodinámicos observados (entalpías de unión más desfavorables).

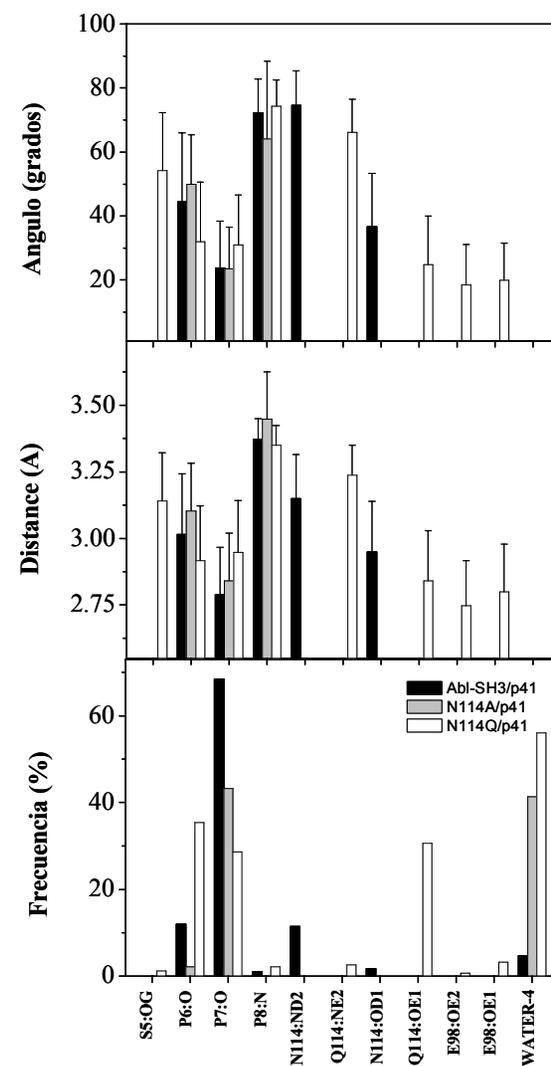
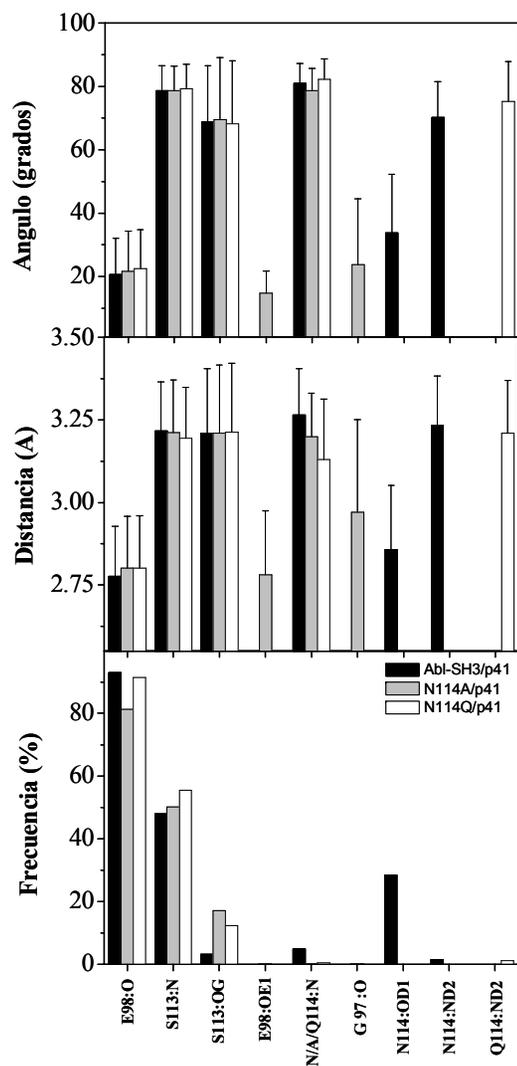
El impacto de la mutación N114Q es quizás algo más complicado. En este caso la ocupancia aumenta hasta un 93%, de forma que la posición 5 está ocupada por moléculas de agua durante prácticamente todo el tiempo de simulación, lo que se encuentra en buena concordancia con el aumento de las contribuciones entálpicas favorables. Sin embargo, se observa un cambio importante en las características dinámicas de estas moléculas de agua con respecto al complejo silvestre. Así, en N114Q/p41 la posición 5 está ocupada por un número muy grande de moléculas de agua (168 moléculas) que durante la dinámica permanecen periodos relativamente cortos de tiempo en la interfase de unión, con un tiempo de residencia promedio inferior a 70 ps, de magnitud similar al encontrado para la mutación N114A, lo que parece ser indicativo de que en este caso las interacciones no están tan optimizadas como lo estaban en Abl-SH3/p41. Esta situación se ilustra en la Figura 3.21, donde se muestra la distribución de las moléculas de agua en función de su tiempo de residencia para cada uno de los sitios de hidratación.

Para comprender mejor el origen de estas diferencias hemos analizado las frecuencias, distancias promedio (entre los dos átomos pesados) y ángulos de los puentes de hidrógeno que establecen las moléculas de agua en los sitios 4 y 5. Los resultados de este análisis se resumen en la Figura 3.23 para las tres dinámicas.

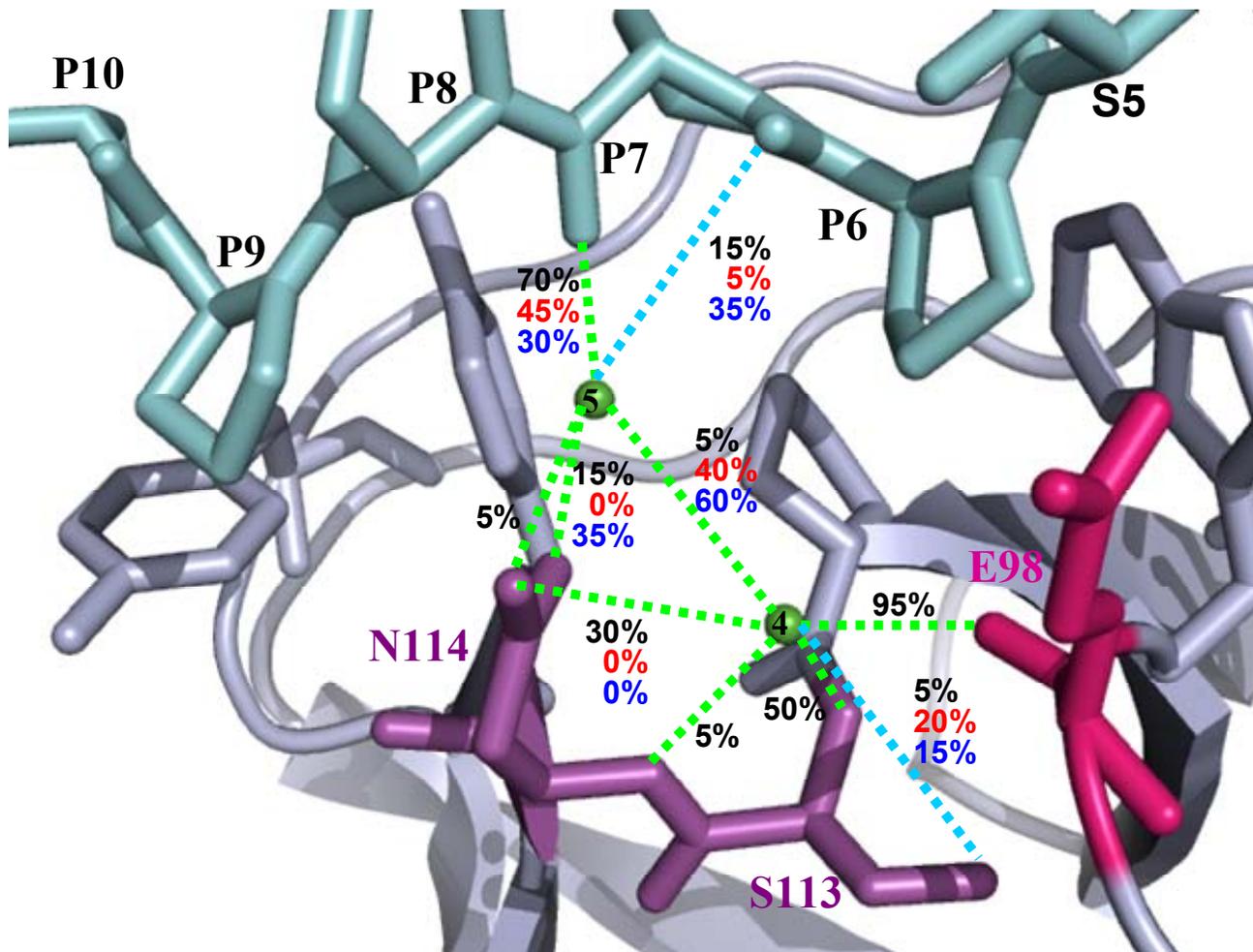
## Resultados

Para mayor claridad, en la Figura 3.24 se indica sobre la estructura tridimensional la frecuencia de formación de las interacciones más importantes establecidas por las moléculas de agua en las posiciones 4 y 5 en la proteína silvestre, así como las diferencias más importantes con los mutantes.

Como se puede observar en el caso de Abl-SH3/p41 las moléculas de agua en el sitio de hidratación 4 se mantienen en su posición mediante el establecimiento de enlaces de hidrógeno con los átomos del esqueleto peptídico de los restos E98 (presente durante un 90% del tiempo de simulación), S113 (50% del tiempo de simulación), y con menor frecuencia con la cadena lateral del aminoácido N114 (30% del tiempo de simulación). Adicionalmente, en el complejo silvestre se establece un puente de hidrógeno directo entre las dos moléculas de agua en posiciones 4 y 5 que es poco frecuente y se observa sólo durante un 5% del tiempo de simulación. Como se aprecia en la Figura 3.23 existe una buena correlación entre los valores de distancia y ángulo de enlace de una interacción en particular y la frecuencia con la que dicho enlace se establece. Así, cuando los valores de distancia y ángulo del enlace de hidrógeno son más cercanos a los valores óptimos para el puente de hidrógeno (distancia 2.8 Å entre los átomos pesados formadores del puente de hidrógeno, y ángulo de 50° definido entre el enlace covalente D-H y el puente de hidrógeno con A, donde D es el átomo donador y A el átomo aceptor del puente de hidrógeno) (Jiang *et al.* 2005), la frecuencia de dicho enlace es mayor. Por ejemplo, puede apreciarse que la interacción con E98 tiene una frecuencia de formación muy alta ya que su distancia (2.7 Å) y su ángulo (20°), están muy favorecidos. Por el contrario, el enlace de hidrógeno con S113 es menos frecuente, ya que su distancia y ángulo son menos favorables (3.3 Å y 70°, respectivamente). Las interacciones que mantienen las moléculas de agua en posición 5 en la interfase de unión son fundamentalmente las establecidas con los oxígenos carbonilo de los restos P7 (70% del tiempo de simulación) y P6 (15% del tiempo de simulación) en el péptido. Sorprendentemente, los puentes de hidrógeno entre las moléculas de agua en posición 5 y la cadena lateral del aminoácido N114 en el dominio SH3 sólo se establecen durante un 20% del tiempo de simulación. De nuevo se observa una muy buena correlación entre las propiedades geométricas del enlace y su frecuencia de formación, de modo que la interacción con P7 del ligando se caracteriza por ángulos y distancias de enlace óptimas (2.8 Å y 25° respectivamente), mientras que



**Figura 3.23.** Frecuencia de formación y propiedades geométricas (ángulo y distancia) de las interacciones establecidas por las moléculas de agua en los sitios de hidratación 4 (izquierda) y 5 (derecha) para la interfase de unión de los complejos Abl-SH3/p41 (barras negras), N114A/p41 (barras grises) y N114Q/p41 (barras blancas)



**Figura 3.24.** Frecuencia de interacciones establecidas en la interfase de unión de la estructura Abl-SH3/p41 observadas en las simulaciones de dinámica molecular. También se indican las diferencias en la frecuencia de formación para las mismas interacciones o equivalentes en el caso de N114A/p41 (en rojo) y N114Q/p41 (azul). Los puentes de hidrógeno que no se observan en las estructuras cristalográficas pero que tienen lugar en las simulaciones se representan como líneas turquesa discontinuas.

las interacciones minoritarias presentan geometrías menos favorables.

En cuanto a los efectos de las mutaciones en las interacciones establecidas en la interfase de unión, podemos observar en la Figura 3.24 que las mutaciones introducidas en la posición 114 tienen un notable impacto sobre las interacciones que establecen las moléculas de agua en el sitio 5, mientras que las establecidas por las moléculas de la posición 4 se ven menos afectadas. Como es previsible, en el mutante N114A se observa que las interacciones establecidas entre la cadena lateral de N114 y las moléculas de agua en posición 5 desaparecen y a su vez la frecuencia de las interacciones establecidas con los restos P6 y P7 del ligando es significativamente menor. Concretamente, la frecuencia del puente de hidrógeno con P6 es menor del 5% y, lo que es más relevante, las interacciones con P7 se establecen sólo durante un 45% del tiempo de simulación, mientras que para la proteína silvestre observábamos un 15 y 70 % respectivamente. Esta disminución en la frecuencia de interacciones se compensa mediante el fortalecimiento del puente de hidrógeno establecido entre las moléculas de agua de ambos sitios 4 y 5, muy minoritario en el complejo silvestre y que en el mutante N114A llega a formarse un 45 % del tiempo de simulación. En general, por tanto, podemos decir que existe un debilitamiento de las interacciones establecidas por esta molécula de agua en el caso de N114A/p41, que hace que su ocupancia del sitio sea menor y es coherente con el valor de entalpía de unión al ligando menos favorable.

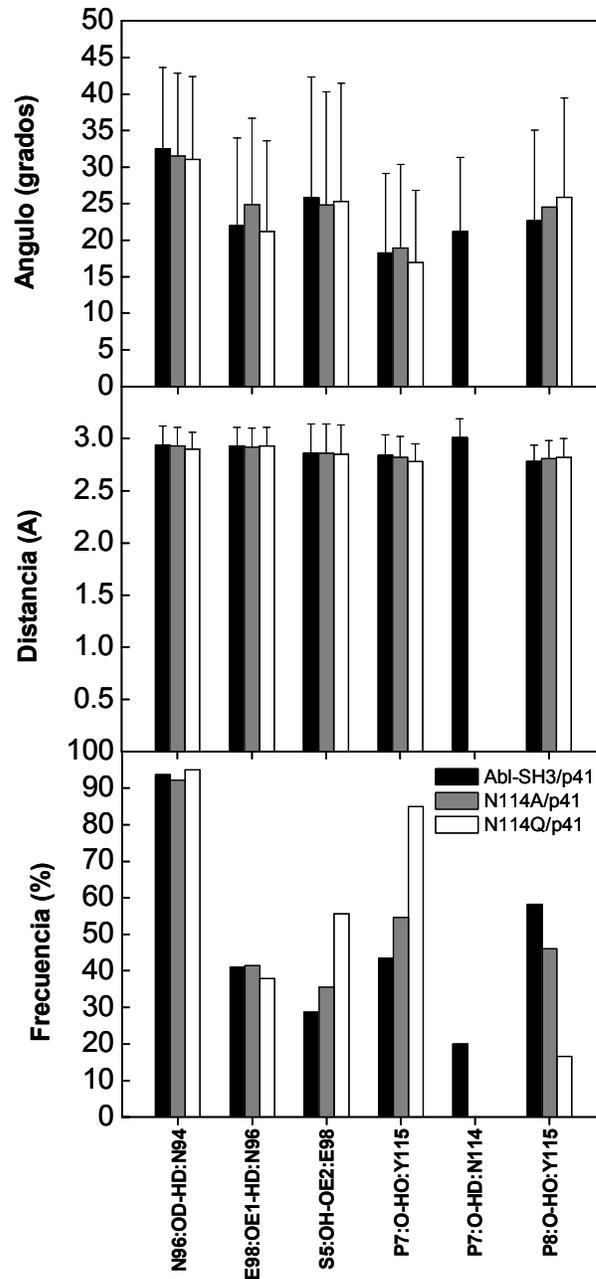
En el complejo N114Q/p41 se produce una situación similar con respecto a las interacciones de las moléculas de agua en posición 5 con el carbonilo de P7 del ligando, cuya frecuencia se reduce drásticamente (desde un 70% en el complejo silvestre hasta un 30%). Esta disminución, mayor que la observada en el caso N114A/p41, está compensada más eficientemente por un aumento en la frecuencia de formación de varias interacciones e incluso el establecimiento de algunas nuevas. Así, se observa un incremento significativo en el porcentaje de tiempo durante el cual se establecen dos puentes de hidrógeno por la molécula de agua en posición 5: tanto en el puente con el oxígeno carbonilo de P6 en el ligando como en el establecido con el oxígeno epsilon de la cadena lateral de Q114 aumenta en un 20% con respecto a la interacción equivalente establecida en Abl-SH3/p41 con la cadena lateral de N114. También el puente con las moléculas de agua en posición 4

## **Resultados**

aumenta dramáticamente desde un 5% en Abl-SH3/p41 hasta un 60% en N114Q/p41. Además de estas interacciones se establecen otras nuevas (con frecuencias del 5% o menores) que no se observan en Abl-SH3/p41: las interacciones establecidas con ambos oxígenos de la cadena lateral del resto E98 que suman una frecuencia en torno al 10%, y con la cadena lateral de S5 del ligando con frecuencias ligeramente menores del 5%. Aunque su contribución a la energética sea modesta, estas nuevas interacciones ilustran la mayor capacidad de la cadena lateral del resto Gln para establecer interacciones en el sitio de unión con respecto a Asn, posiblemente relacionada con su mayor longitud, y que se traduciría, como efectivamente ocurre, en un aumento de la ocupancia de la posición 5 y en una entalpía de unión más favorable.

Por último, hemos evaluado también la contribución de las interacciones no mediadas por aguas entre Abl-SH3 y p41. En la Figura 3.25 se muestra un análisis de las interacciones directas SH3-péptido con frecuencia de formación por encima del 5%, tanto intramoleculares como intermoleculares, para Abl-SH3/p41, N114A/p41 y N114Q/p41. Es interesante señalar que en las trayectorias de los tres complejos se observa el establecimiento de enlaces de hidrógeno intramoleculares en el dominio SH3 entre los átomos de la cadena lateral del resto N96, en el lazo n-Src, con los aminoácidos N94 y E98 y con frecuencias superiores al 95% y 40% respectivamente. Es posible que la pérdida de estas interacciones en los mutantes en la posición N94 sea en parte responsable de las alteraciones conformacionales descritas en apartados anteriores.

En cuanto a las interacciones intermoleculares establecidas directamente entre la proteína y el ligando podemos destacar tres principales: *i*) un enlace de hidrógeno entre la cadena lateral de N114 y el oxígeno carbonilo de P7 en el ligando, que no se observa en la estructura cristalográfica y que se establece durante un 20% del tiempo de simulación en Abl-SH3/p41 y desaparece en ambos mutantes, en el caso de N114A/p41 debido a la eliminación de la cadena lateral polar y en el mutante N114Q/p41 posiblemente desfavorecido por la mayor longitud de la cadena lateral de la Gln, que parece implicada en otras interacciones; *ii*) un enlace de hidrógeno entre el hidroxilo de la cadena lateral de Y115 y el carbonilo



**Figura 3.25.** Frecuencia de formación y propiedades geométricas (ángulo y distancia) de las interacciones directas establecidas entre proteína y ligando en la interfase de unión de los complejos Abl-SH3/p41 (barras negras), N114A/p41 (barras grises) y N114Q/p41 (barras blancas)

de P8 en el ligando, uno de los enlaces más conservados en todas las interacciones SH3-péptido, que se observa durante un 58% del tiempo de simulación en Abl-SH3/p41, y que el 45% del tiempo restante el enlace se establece con el oxígeno carbonilo de P7.

## **Resultados**

En las dinámicas con los complejos de los mutantes con p41 se observa que esta segunda interacción se favorece, bien ligeramente en el caso de N114A/p41 (45% para el enlace Y115-P8 frente al 55% para Y115-P7) y muy significativamente en el caso de N114Q/p41, donde la interacción con P7 se hace mayoritaria (15% para el enlace Y115-P8 frente al 85% para Y115-P7); *iii*) un puente de hidrógeno entre la cadena lateral de E98 y el hidroxilo de S5 en el ligando, que no aparece en la estructura cristalográfica pero que se observa en todas las dinámicas con frecuencias considerables (30, 35 y 60% para Abl-SH3/p41, N114A/p41 y N114Q/p41 respectivamente). Es interesante señalar que la posibilidad de formación de este último enlace se consideró en el proceso de diseño del ligando p41 (Pisabarro *et al.* 1994; Pisabarro & Serrano 1996), aunque en las estructuras cristalográficas la cadena lateral del resto S5 del ligando se encuentra implicada en un enlace de hidrógeno intramolecular con el carbonilo de P6 (Figura 3.5, página 137). Con el fin de comprobar la relevancia de estas interacciones, hemos llevado a cabo un análisis adicional de la dinámica del péptido p41 en disolución, que se presenta con detalle en el Apéndice IV, que confirma que la interacción intramolecular S5-P6 tampoco tiene lugar en el péptido libre y podría ser debida a un mero efecto de cristalización.

### **3.1.2. Análisis termodinámico-estructural de la interacción de Abl-SH3 con derivados mono-sustituídos de p41 y con el ligando 3BP1**

Con el fin de comprender la influencia del propio ligando p41 sobre la energética de unión hemos seleccionado una colección de ligandos que incluye el ligando natural 3BP1, secuencia origen en el proceso de diseño de p41, además de otros cuatro decapeptidos derivados de p41 (Pisabarro *et al.* 1994; Pisabarro & Serrano 1996). Estos últimos, cuyas secuencias se recogen en la Tabla 3.12, poseen sustituciones puntuales de las posiciones 5 y 8 con respecto al ligando p41 que, debido a la estructura PPII del péptido unido, están dirigidas hacia el exterior de la interfase de unión y no se encontrarían estableciendo interacción alguna con el dominio SH3 (Figura 3.4, página 135). Como consecuencia, en el caso de que tuviéramos un modo de unión que implicara únicamente la interacción directa entre dos superficies hidrofóbicas, sería de esperar que la sustitución de un resto aromático o alifático en estas posiciones por Pro originara un incremento en la afinidad de origen fundamentalmente entrópico debido a la reducción de grados de libertad del péptido libre en disolución, cuyo equilibrio conformacional se vería desplazado hacia la formación de la hélice PPII. Por otro lado, ya que estas cadenas laterales no están implicadas en ningún contacto directo con el dominio, sería de esperar que el efecto de las sustituciones en la entalpía de unión fuera muy pequeño.

Todos los experimentos fueron realizados en fosfato sódico 20 mM pH 7.0 a 25 °C, tampón con efecto despreciable sobre la entalpía intrínseca según comprobamos en la unión de p41 a la proteína silvestre. Los termogramas correspondientes a la titulación del dominio Abl-SH3 con los distintos ligandos así como los parámetros termodinámicos obtenidos se muestran en la Figura 3.26 y en la Tabla 3.13 respectivamente. Dado que las sustituciones no incluyen la eliminación ni la introducción de restos ionizables, es de esperar un comportamiento similar al descrito para el ligando p41 en lo que respecta a la ausencia de efectos asociados a equilibrios de ionización, acoplados a la unión, en los parámetros termodinámicos. Por consiguiente, los efectos caloríficos netos son los recogidos en la Tabla 3.13.

## Resultados

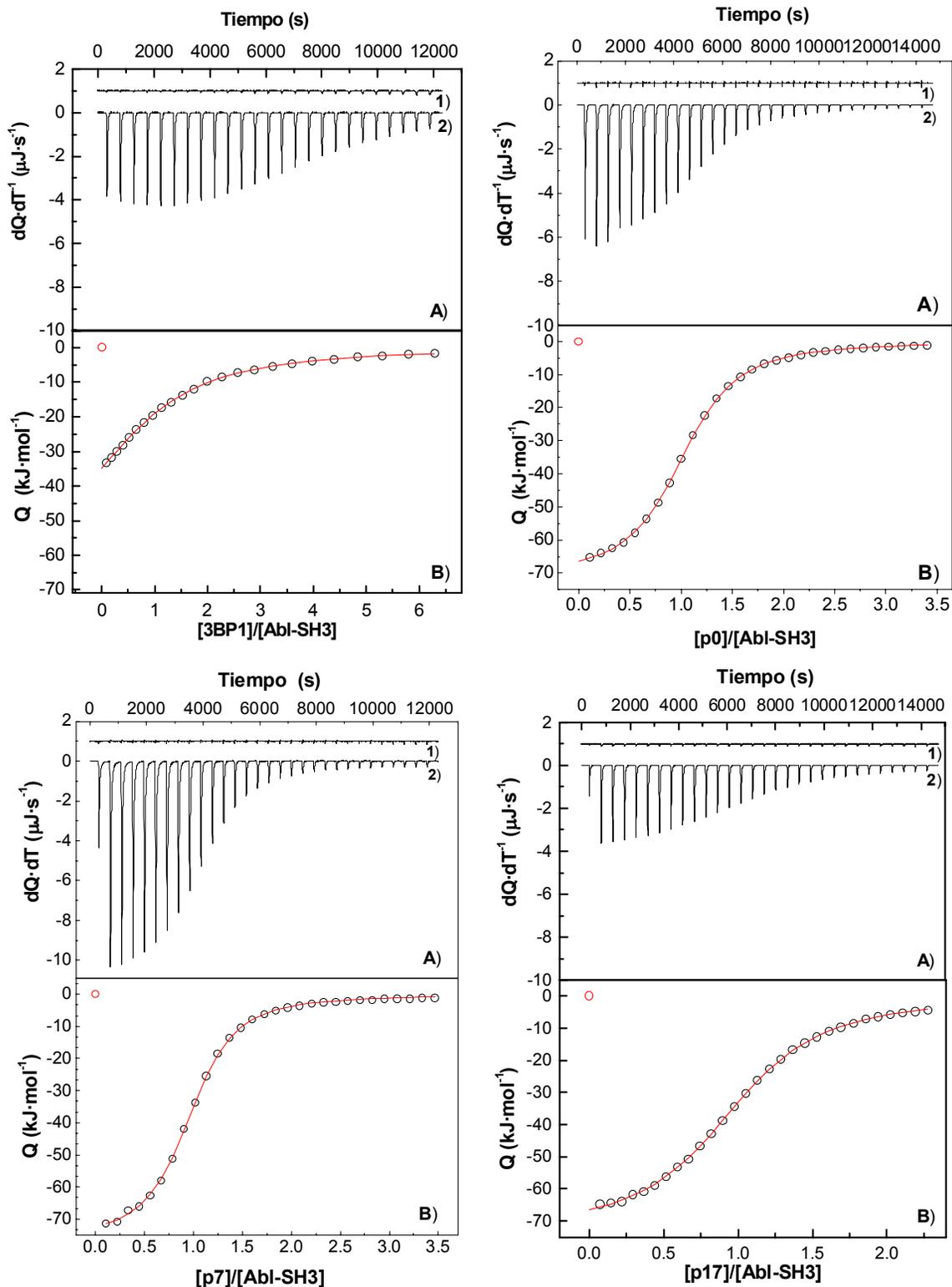
*Tabla 3.12 Secuencias de los ligandos derivados de p41 en las que las sustituciones en las posiciones expuestas al disolvente, 5 y 8, se muestran coloreadas en azul y rojo respectivamente*

Péptido	3BP1	p40	p7	p17	p0
Secuencia	APT <b>M</b> PP <b>L</b> PP	APT <b>Y</b> SPP <b>P</b> PP	APT <b>Y</b> PP <b>P</b> PP	APT <b>Y</b> SP <b>L</b> PP	APT <b>Y</b> PP <b>L</b> PP

En general, la unión de todos los ligandos se caracteriza por un patrón termodinámico muy similar al descrito para el ligando p41, caracterizado por una entalpía de unión favorable y una contribución entrópica desfavorable a la energía libre de Gibbs, las cuales se compensan parcialmente. No obstante, se aprecian diferencias significativas en las magnitudes de los distintos parámetros termodinámicos obtenidos para los diferentes ligandos, que se muestran esquemáticamente en la Figura 3.27.

### 3.1.2.1. Efecto de la sustitución P8L: comparación de los ligandos p40 frente a p17 y p7 frente a p0

El ligando p40 es el que posee mayor afinidad junto con p41, lo que coincide con lo que se había descrito previamente mediante experimentos de fluorescencia (Pisabarro & Serrano 1996). El aumento de afinidad obtenido para estos dos ligandos con respecto al resto de los estudiados aquí se ha explicado en base a una disminución del coste entrópico, ya que la introducción de un mayor número de restos Pro en la secuencia de estos ligandos provocaría una redistribución de las poblaciones en disolución hacia aquellas con la conformación en hélice PPII, que es la requerida para unirse al dominio. Sin embargo, tanto para p40 como para p41, el aumento de afinidad tiene un origen entálpico, muy parecido en ambos casos y que es el mayor observado para todos los ligandos (Tabla 3.13).



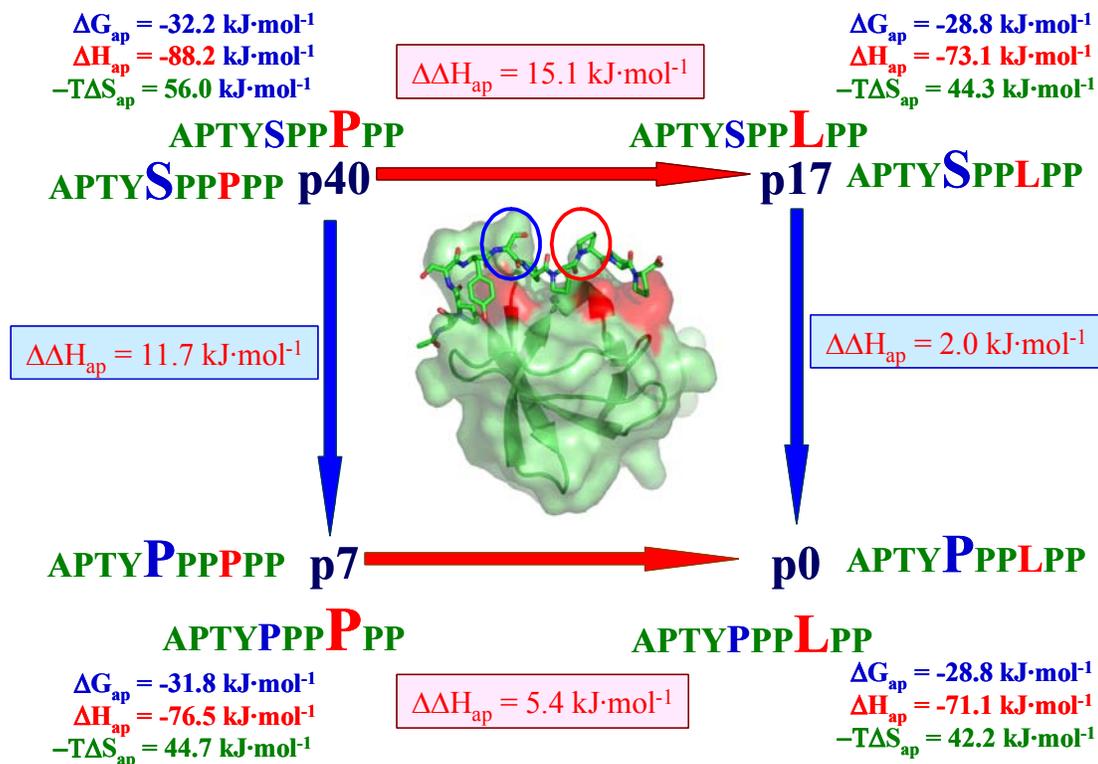
**Figura 3.26** Experimentos de titulación del dominio Abl-SH3 con ligandos derivados de p41 mediante ITC en fosfato 20 mM, pH 7.0, a 25 °C. **A.** (1) Termograma de titulación formado por los calores por unidad de tiempo liberados tras la inyección de ligando sobre la proteína Abl-SH3. (2) Termograma de dilución del ligando en las mismas condiciones que el experimento de titulación (desplazados en el eje para mayor claridad). **B.** Isotherma de unión correspondiente al experimento en A, obtenida después de normalizar los calores de titulación y corregir por los calores de dilución. Los datos experimentales se representan como círculos, mientras que la línea continua corresponde al mejor ajuste para el modelo de "n" sitios de unión idénticos e independientes (ecuación 2.21).

## Resultados

**Tabla 3.13** Valores de los parámetros termodinámicos característicos de la interacción de varios ligandos ricos en prolina con el dominio Abl-SH3. Todos los experimentos fueron realizados a 25 °C en fosfato sódico 20 mM pH 7.0.

Péptido	Secuencia	$\Delta G_{ap}$ (kJ·mol <sup>-1</sup> )	$\Delta H_{ap}$ (kJ·mol <sup>-1</sup> )	$-T\Delta S_{ap}$ (kJ·mol <sup>-1</sup> )
p41	APSYSPPPP	-32.2	-91.7	59.5
p40	APTYSPPPP	-32.2	-88.2	56.0
p7	APTYPPPPP	-31.8	-76.5	44.7
p17	APTYSPPLPP	-28.8	-73.1	44.3
p0	APTYPPPLPP	-28.8	-71.1	42.2
3BP1	APTMPPPLPP	-25.4	-69.5	44.1

Para analizar la sustitución Pro-Leu en la posición 8 del ligando podemos comparar las entalpías de los péptidos p17 (APTYSPPLPP) y p40 (APTYSPPPP) (flecha roja superior en Figura 3.27). Observamos una disminución de la entalpía aparente para p17 de 15.1 kJ·mol<sup>-1</sup> con respecto a p40. Dicha variación resulta sorprendente, no sólo porque sería la sustitución de Pro por Leu en una posición expuesta al disolvente y no implicada directamente en la interacción, sino porque además la magnitud del cambio (15 kJ·mol<sup>-1</sup>) es muy grande, además de ser la mayor observada para todos los ligandos en la Figura 3.27. Por otro lado, resulta sorprendente el hecho de que la introducción de Leu sea más favorable entrópicamente (44.3 frente a 56.0 kJ·mol) que Pro, cuando lo esperado sería lo contrario. Lo mismo ocurre cuando se analiza la sustitución Pro-Leu en el contexto de los péptidos p0 (APTYPPPLPP) y p7 (APTYPPPPP), aunque en este caso la magnitud de los cambios entálpicos y entrópicos es algo menor ( $\Delta\Delta H_{ap} = 5.4$  kJ·mol<sup>-1</sup> y  $-T\Delta\Delta S_{ap} = -2.5$  kJ·mol<sup>-1</sup>, flecha roja inferior en Figura 3.27). Es interesante señalar que la única diferencia entre las comparaciones p40-p17 y p7-p0 es la identidad del aminoácido en la posición 5, que es Ser en los péptidos p40 y p17, y Pro en p7 y p0, lo que indica que podría existir algún tipo de cooperatividad entre ambas posiciones.



**Figura 3.27** Ciclo termodinámico que resume los efectos de las mutaciones específicas en las posiciones 5 y 8 en los ligandos peptídicos sobre los parámetros termodinámicos de unión. En el centro de la figura se ha representado la estructura del complejo Abl-SH3/p41 donde se señalan con círculos azul y rojo las posiciones 5 y 8 del ligando respectivamente, que se encuentran expuestas al disolvente.

Finalmente cabe decir con respecto a las sustituciones asociadas a los ligandos p40-p17 como a los ligandos p7-p0 que los cambios en los términos entálpico y entrópico se traducen en un cambio de afinidad de unos  $3 \text{ kJ}\cdot\text{mol}^{-1}$

### 3.1.2.2. Efecto de la sustitución S5P: comparación de los ligandos p40 frente a p7 y p17 frente a p0

La introducción de Pro por Ser en el lugar 5 de p40 (APTYSPPPPP), que corresponde con el péptido p7 (APTYPPPPPP, flecha azul izquierda en Figura 3.27), genera una disminución en el valor de la entalpía aparente de  $11.7 \text{ kJ}\cdot\text{mol}^{-1}$ . Esta variación de entalpía está compensada por una variación de la contribución

## **Resultados**

entrópica de signo contrario ( $-11.3 \text{ kJ}\cdot\text{mol}^{-1}$ ), lo que se traduce en una variación de la afinidad prácticamente nula. La misma situación tiene lugar con la sustitución S5P en el contexto de los péptidos p0 (APTY**P**PPLPP, flecha azul derecha en la Figura 3.27) y p17 (APTY**S**PPLPP), que origina un cambio de entalpía aparente de  $2 \text{ kJ}\cdot\text{mol}^{-1}$ , compensada por una contribución entrópica equivalente de signo contrario, que hace que se mantenga la misma afinidad para ambos ligandos.

Observando la Figura 3.27, en el caso de los péptidos p40-p7 la magnitud de los cambios en los términos entálpico y entrópico (flecha azul izquierda) resulta comparable a la obtenida para los ligandos p40-p7 (flecha roja superior), aunque su valor es algo menor. Del mismo modo, los cambios obtenidos para los ligandos p17-p0 (flecha azul derecha) son comparables a los de los ligandos p7-p0 (flecha roja inferior), aunque también la magnitud es algo menor. A la luz de estos resultados, parece adquirir importancia la idea de que existe cierta cooperatividad entre los restos S5 y P8 del ligando, ya que la sustitución de cualquiera de ellos genera cambios en los términos entálpico y entrópico mucho mayores que cuando es otro aminoácido el que se sustituye en estas posiciones.

### **3.1.2.3. La sustitución por Tyr del cuarto resto en el ligando natural 3BP1 resulta en un aumento de afinidad por Abl-SH3**

Una de las estrategias utilizadas para el desarrollo y optimización de los ligandos que llevó a la obtención de p40 y p41 a partir del ligando natural 3BP1 fue la introducción de un resto Tyr en lugar del resto M4 del ligando natural 3BP1 (APT**M**PPPLPP). De acuerdo con la estructura cristalográfica del complejo Abl-SH3/p41 la introducción de Tyr en el ligando permite establecer un nuevo puente de hidrógeno con los aminoácidos S75 y D77 del lazo RT, lo que parece contribuir al aumento de afinidad obtenido en todos los ligandos analizados frente a 3BP1 (Tabla 3.13) (Pisabarro & Serrano 1996).

El análisis termodinámico que hemos realizado con estos ligandos indica, además, que el aumento de afinidad obtenido se debe a una optimización del término entálpico, el cual podría ser debido al establecimiento de los puentes de hidrógeno por parte de la Tyr, además del establecimiento de interacciones de tipo

polar por las moléculas de agua enterradas en la interfase de unión. Los términos entálpicos que caracterizan la interacción presentan cambios favorables con respecto a 3BP1 que van desde los  $-1.6 \text{ kJ}\cdot\text{mol}^{-1}$  para p0 hasta los  $-22.2 \text{ kJ}\cdot\text{mol}^{-1}$  para p41. Por otro lado, se generan cambios desfavorables de menor magnitud en los términos entrópicos que van desde los  $1.9 \text{ kJ}\cdot\text{mol}^{-1}$  para p0 hasta los  $15.4 \text{ kJ}\cdot\text{mol}^{-1}$  para p41 que, junto con los cambios entálpicos anteriores, se traducen en un aumento significativo de la afinidad respecto a 3BP1 que va desde  $-3.5 \text{ kJ}\cdot\text{mol}^{-1}$  para los ligandos p0 y p17 hasta  $-6.8 \text{ kJ}\cdot\text{mol}^{-1}$  para el ligando p41. Tal modificación constituye, por tanto, un claro ejemplo de sustitución ideal ya que con la introducción de un sólo aminoácido en la secuencia original del ligando obtenemos una optimización del término entálpico frente al entrópico, que contribuyen a un aumento significativo en la afinidad de unión.

### **3.2. RECONOCIMIENTO DE LIGANDOS RICOS EN PROLINA POR DOMINIOS WW**

Hemos abordado la caracterización termodinámica del reconocimiento de ligandos ricos en prolina por dominios WW con el fin de investigar si estas interacciones presentan un comportamiento termodinámico y un nivel de complejidad similares a los descritos para dominios SH3. Para ello, hemos seleccionado dos dominios WW de la clase de especificidad PPxY: el dominio **Yap65-WW**, considerado como prototipo de esta clase de especificidad y el más ampliamente estudiado hasta la fecha (Macias *et al.* 1996; Pires *et al.* 2001) y el dominio **Nedd4-WW3**, identificado como la diana celular de los dominios L víricos de tipo PPxY (Gottwein *et al.* 2003; Staub *et al.* 1996; Timmins *et al.* 2003). Como ligando propio del dominio Yap65-WW hemos seleccionado el péptido de secuencia **EYPPYPPYPSG**, que denominaremos **M1**, y que es el ligando de mayor afinidad descrito hasta la fecha para dominios WW (Pires *et al.* 2001). Como ligandos propios del dominio Nedd4-WW3 hemos seleccionado varias secuencias correspondientes a los **dominios L de los virus del ébola y de la leucemia humana de tipo I** en los que el motivo PPxY se encuentra en diferentes posiciones dentro del ligando y flanqueado por secuencias de longitud variable. En la Figura 3.28 se muestra un alineamiento de secuencias de los dominios L de varios virus y en la Tabla 3.14 se recogen las distintas secuencias de los ligandos utilizados y la nomenclatura que se utilizará para su identificación en esta Memoria. Es de esperar que el estudio conjunto de las interacciones de todos estos ligandos, incluyendo el péptido correspondiente al motivo **PPY** aislado, con ambos dominios, proporcione información sobre los determinantes termodinámicos de estas interacciones, no sólo en términos de afinidad sino también de especificidad de unión.

**Dominios L tipo PPXY**

**MOLONEY MURINE LEUKEMIA** - - L I D L L T E D P P P Y R D P R P P P S - -  
**HUMAN T-CELL LEUKEMIA 1** - - P P D S D P Q I P P P Y V E P T A P Q V - -  
**HUMAN T-CELL LEUKEMIA 2** - - P P P S H A E V P P P Y V E P T T T Q C - -  
**BOVINE LEUKEMIA** - - G A S A P E E Q P P P Y D P P A I L P I - -  
**GIBBON APE LEUKEMIA** - - L L L L S E P T P P P Y P A A L P P P L - -  
**VESICULAR STOMATITIS** - - K S K K L G I A P P P Y E E D T S M E Y - -  
**MASON-PFIZER** - - D W P P F L T R P P P Y N K A T P S A P - -  
**RABIES** - - D G D D L W L P P P E Y V P L K E L T S - -  
**MARBURG** - - N T Y M Q Y L N P P P Y A D H G A N Q L - -  
**ZAIRE EBOLA** - - R R V I L P T A P P E Y M E A I Y P V R - -  
**HUMAN IMMUNODEFICIENCY 1** - - Q S R P E P T A P P E E S F R S G V E T - -

**Dominios L tipo P(T/S)AP**

**HUMAN IMMUNODEFICIENCY 1** - - L Q S R P E P T A P P E E S F R  
**ZAIRE EBOLA** - - M R R V I L P T A P P E Y M E A  
**VESICULAR STOMATITIS** - - T S M E Y A P S A P I D K S Y F  
**GIBBON APE LEUKEMIA** - - R P V A G R P S A P P R P P T Y  
**HUMAN T-CELL LEUKAEMIA 1** - - P P P Y V E P T A P Q V L P V M  
**HUMAN T-CELL LEUKAEMIA 2** - - P P P Y V E P T T T Q C F P I L

**Figura 3.28.** Alineamiento de secuencias de dominios víricos tipo L que contienen secuencias específicas ricas en prolina que interactúan con diversas proteínas celulares. En la figura de arriba se muestra parte de la secuencia de los virus que contienen la secuencia consenso PPXY (donde X puede ser cualquier aminoácido) que interactúa con la proteína celular Nedd4-WW3. En la figura de abajo se muestran las secuencias de los virus que contienen la secuencia consenso PT/SAP que interactúa con la proteína Tsg101-UEV. Los alineamientos fueron realizados en base a homología de secuencias con la opción ClustaW del programa BioEdit (Hall 1999).

**Tabla 3.14.** Selección de ligandos de dominios L de los virus VIH, ébola y leucemia humana tipo I que hemos utilizado para la caracterización termodinámica de las interacciones que establecen con los dominios Yap65-WW y Nedd4-WW3

Ligando	Nombre	Secuencia*
ébola Zaire VP40	Ébola-11	IL <b>PTA</b> <b>P</b> PEYME
leucemia humana-gag	Leucemia-a	<b>PPPY</b> VEPTAP
leucemia humana-gag	Leucemia-b	SDPQI <b>PPPY</b> VEP
(Macias <i>et al.</i> 1996)	M1	EYPPY <b>PPPY</b> PSG
dominio L PPxY	PPPY	<b>PPPY</b>

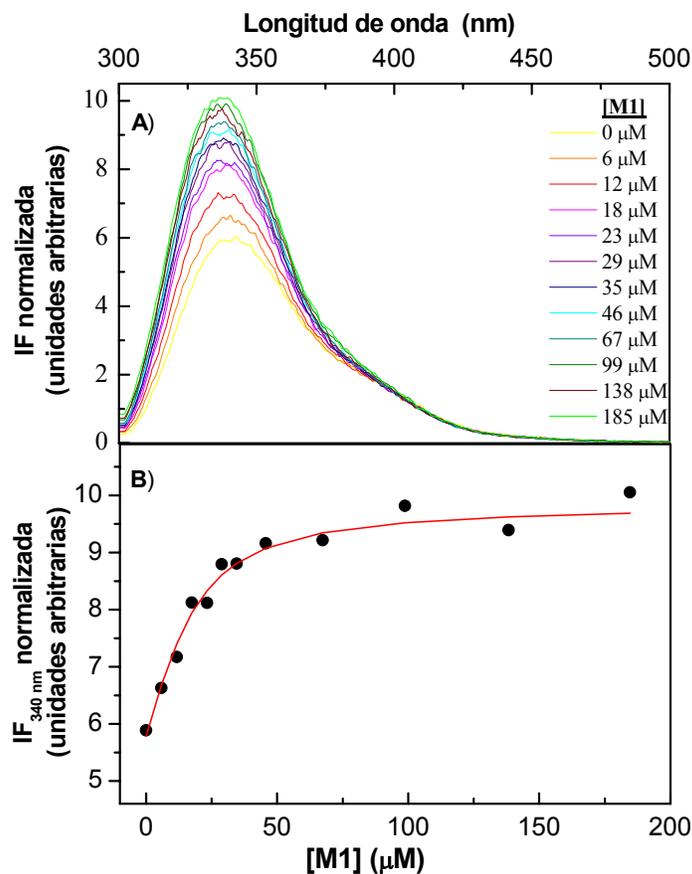
\*Para identificar más fácilmente la posición de los motivos PPxY y PTAP en la secuencia del ligando, éstos se muestran en rojo y azul respectivamente.

### 3.2.1. Interacción del ligando M1 con Yap65-WW y Nedd4-WW3

Aunque la interacción entre Yap65-WW y el ligando **M1** ha sido descrita con anterioridad (Pires *et al.* 2001) no hay información disponible sobre su interacción con Nedd4-WW3. Con el fin de obtener una estimación inicial de las constantes de asociación del ligando **M1** hemos analizado la interacción mediante espectroscopia de fluorescencia. Como se puede observar en la Figura Yap-complejo (página 32 antecedentes), donde se ilustra un complejo de Yap65-WW con un ligando muy parecido en secuencia a **M1**, uno de los dos Trp característicos de los dominios WW se encuentra en el sitio de unión (W39), por lo que es de esperar que la unión del ligando pueda alterar su entorno químico e inducir cambios en su espectro de fluorescencia. Aunque, como veremos a continuación, los efectos de la unión del ligando en la fluorescencia de estos dominios son muy variables de un complejo a otro y en algunos casos demasiado sutiles para un estudio cuantitativo, la interacción con **M1** induce un aumento en la intensidad del máximo de fluorescencia, así como un desplazamiento de éste desde 345 nm hasta 340 nm, que concuerda con lo descrito en estudios anteriores (Pires *et al.* 2001) y que permite la determinación de las constantes de asociación con suficiente fiabilidad.

En la Figura 3.29, se muestra como ejemplo la titulación del dominio Nedd4-WW3 con el ligando **M1** llevada a cabo en fosfato sódico 20 mM pH 7.0 a 25 °C, que son las mismas condiciones utilizadas en los estudios previos con Yap65-WW (Pires *et al.* 2001). Los resultados de estas titulaciones indican que el ligando **M1** interacciona con Yap65-WW con una constante de asociación de  $7.6 \cdot 10^4 \text{ M}^{-1}$ , en buena concordancia con los datos en bibliografía ( $K_a = 15 \cdot 10^4 \text{ M}^{-1}$ ), mientras que presenta un afinidad ligeramente superior para el dominio Nedd4-WW3 ( $K_a = 14.3 \cdot 10^4 \text{ M}^{-1}$ ).

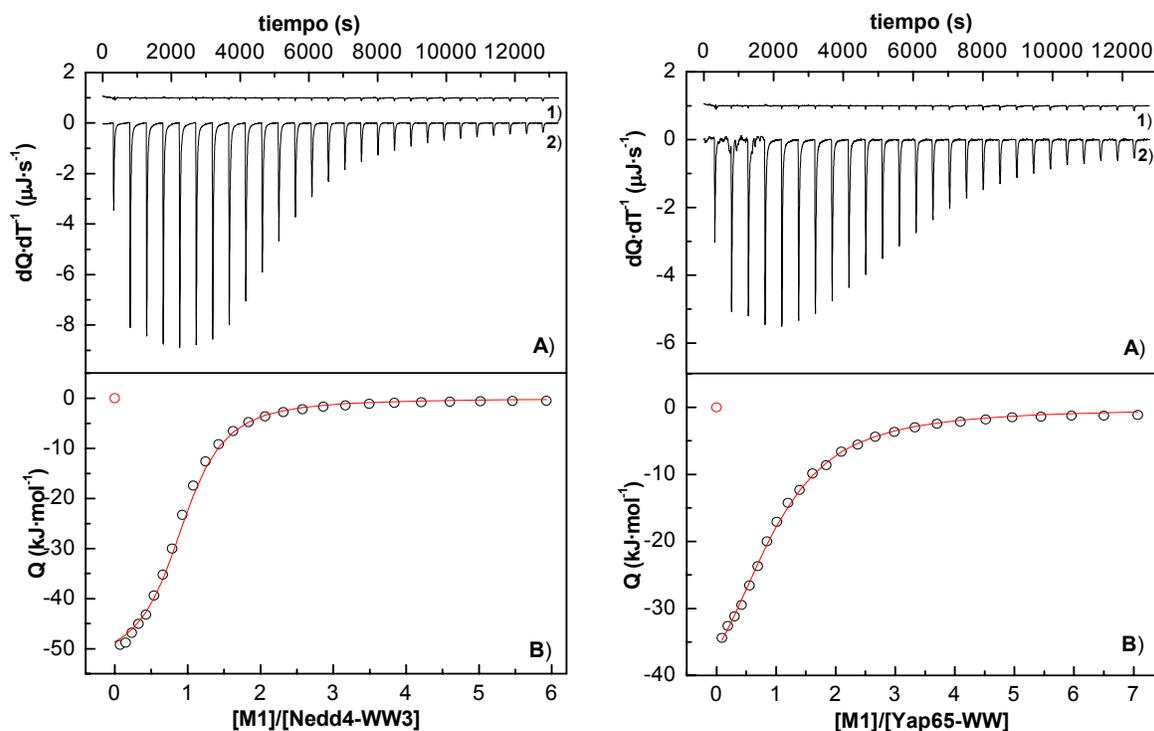
Una vez comprobado que la afinidad de ambos dominios por el ligando **M1** está en el intervalo adecuado hemos procedido a su caracterización termodinámica mediante calorimetría isotérmica de titulación. En la Figura 3.30 se muestran las titulaciones con ambos dominios WW en tampón fosfato sódico 20 mM pH 7.0 a 25 °C (paneles superiores), que se realizaron utilizando un perfil de volúmenes de



**Figura 3.29.** A) Espectros de fluorescencia tomados tras la inyección de ligando **M1** sobre la proteína *Nedd4-WW3*, una vez corregidos por la dilución del ligando y normalizados por la concentración, realizados en fosfato sódico 20 mM pH 7.0 a 25 °C. En el panel B) se han representado los valores de intensidad de fluorescencia a 340 nm de cada uno de los espectros anteriores normalizados por la concentración (círculos negros). La línea continua corresponde al mejor ajuste para el modelo de unión sencillo de un ligando a una proteína (ecuación 2.3, sección 2.2.8.3).

inyección crecientes (entre 4 y 21  $\mu\text{L}$ ) para optimizar la definición de la curva de unión, que se muestra en los paneles inferiores junto con el mejor ajuste a un modelo de “n” sitios idénticos e independientes. Adicionalmente, hemos llevado a cabo experimentos en las mismas condiciones pero a otros dos valores de temperatura, lo que nos ha permitido calcular el cambio de capacidad calorífica de unión. Los resultados obtenidos para todos estos experimentos se recogen en la Tabla 3.15. En ambos casos el ligando se une a un único sitio de unión con afinidades moderadas, muy parecidas a las determinadas mediante fluorescencia y a las descritas para otros complejos con dominios WW

## Resultados



**Figura 3.30.** Experimentos de calorimetría de titulación para la unión del ligando **M1** (EYPPY**PPY**PSG) con la proteína Nedd4-WW3 (izquierda) y con Yap65-WW (derecha), en fosfato sódico 20 mM pH 7.0 a 25 °C. **A)** (1) Termograma de dilución del ligando (desplazados en el eje para mayor claridad), y (2), termograma de titulación formado por los calores por unidad de tiempo liberados tras la inyección de ligando. **B)** Isoterma de unión de cada experimento, obtenida después de normalizar los calores de titulación y corregir por los calores de dilución. Los datos experimentales se representan como círculos, mientras que la línea continua corresponde al mejor ajuste para el modelo de "n" sitios de unión idénticos e independientes (ecuación 2.21, sección 2.2.9.3).

(Kanelis *et al.* 2006; Macias *et al.* 1996; Meiyappan *et al.* 2007; Morales *et al.* 2007; Pires *et al.* 2005; Pires *et al.* 2001). Como se puede observar, estas interacciones presentan un patrón termodinámico similar al descrito para dominios SH3 (entalpías de unión muy negativas compensadas parcialmente por contribuciones entálpicas desfavorables) y que, al igual que en estos dominios, es incompatible con el carácter altamente hidrofóbico de la interfase de unión.

La afinidad de unión así como las contribuciones entálpicas y entrópicas del ligando **M1** a ambos dominios son muy similares, aunque se observan diferencias significativas en cuanto a la capacidad calorífica de unión, calculada a partir de las gráficas mostradas en la Figura 3.31, de modo que el valor obtenido para el

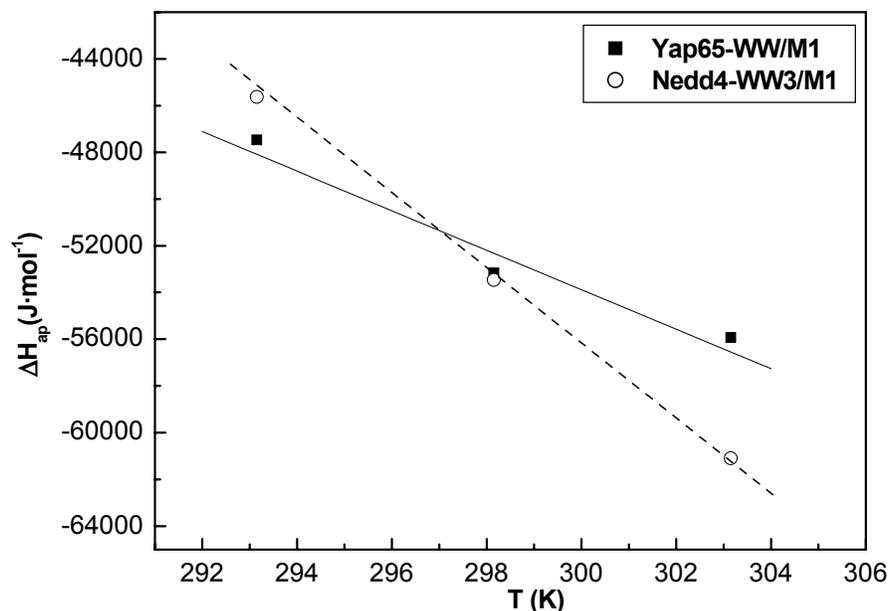
**Tabla 3.15.** Valores de los parámetros termodinámicos para la unión del ligando sintético **MI** (EYPPY**PPY**PSG) a los dominios WW3 de Nedd4 y WW de Yap65, obtenidos a partir de los experimentos de ITC realizados en fosfato sódico 20 mM pH 7.0 (\*).

Proteína	T (°C)	$K_a \times 10^4$ (M <sup>-1</sup> ) <sup>#</sup>	$\Delta G_{ap}$ (kJ·mol <sup>-1</sup> )	$\Delta H_{ap}$ (kJ·mol <sup>-1</sup> ) <sup>#</sup>	$-T\Delta S_{ap}$ (kJ·(K·mol <sup>-1</sup> ))	$\Delta C_{p,ap}$ (J·(K·mol <sup>-1</sup> ))
Yap65-WW	20	8.0	-27.5	<u>-47.3</u>	19.8	
	25	4.9 (7.6)*	-26.7	<u>-53.2</u>	26.5	<u>-846</u>
	30	3.3	-26.2	<u>-56.1</u>	29.9	
Nedd4-WW3	20	24.3	-30.2	<u>-45.6</u>	15.4	
	25	18.7 (14.3)*	-30.1	<u>-53.2</u>	23.1	<u>-1609</u>
	30	11.0	-29.2	<u>-61.1</u>	31.9	

\*Entre paréntesis se muestran los valores de las constantes de asociación obtenidos por fluorescencia.

<sup>#</sup>Los errores en los parámetros termodinámicos se estiman en torno al 5% para la entalpía de unión y al 10% para la constante de asociación. Los valores de entalpía aparente subrayados fueron utilizados para el cálculo de la capacidad calorífica de unión con cada dominio.

## Resultados

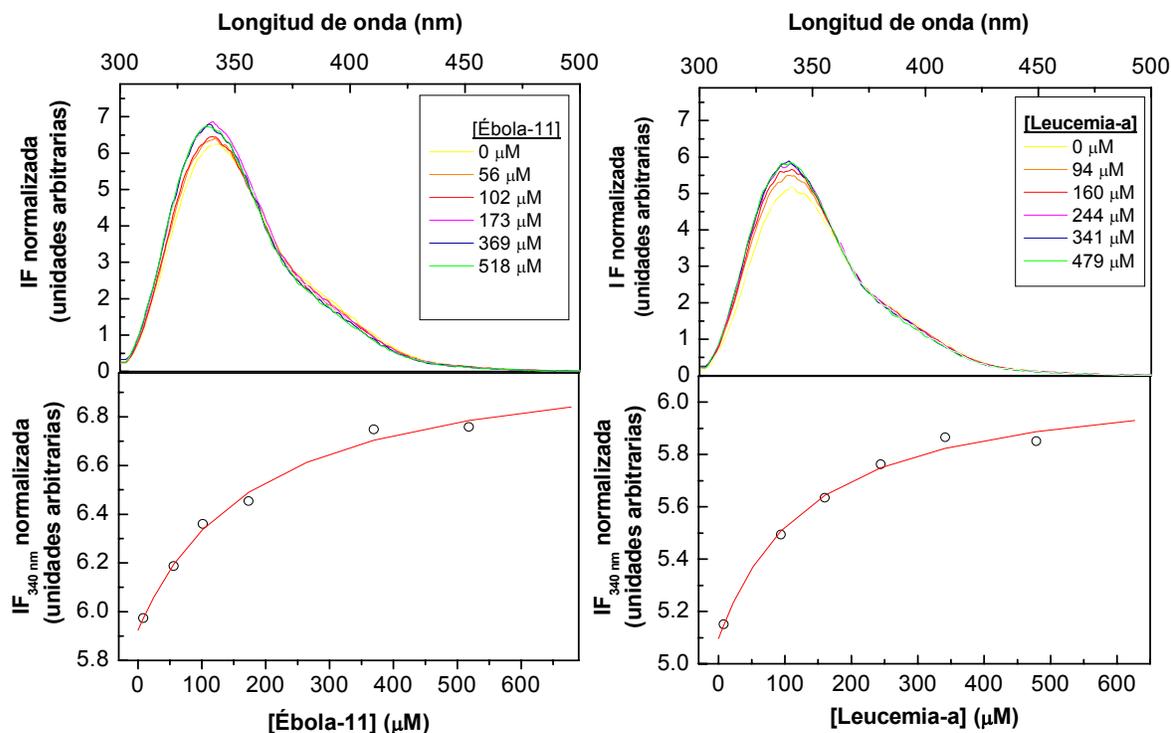


**Figura 3.31.** Representación de la entalpía aparente obtenida de los experimentos de calorimetría para la unión del ligando **M1** (EYPPY**PPY**PSG) con la proteína Nedd4-WW3 y con Yap65-WW en fosfato sódico 20 mM pH 7.0 a 20, 25 y 30 °C. Los datos experimentales se representan como círculos abiertos para Nedd4-WW3, y como cuadrados negros para Yap65-WW, mientras que las líneas corresponden con los respectivos ajustes por regresión lineal para ambas interacciones.

### 3.2.2. Interacción de dominios L víricos tipo PPxY con Nedd4-WW3 y Yap65-WW

Al igual que con el ligando **M1**, hemos realizado una caracterización preliminar mediante espectroscopia de fluorescencia de la interacción de los distintos ligandos peptídicos correspondientes a secuencias de dominios L víricos de tipo PPxY con Nedd4-WW3 y Yap65-WW, con el fin de confirmar la interacción y obtener una estimación de las constantes de asociación que nos permita diseñar correctamente los experimentos de titulación calorimétrica. En la Figura 3.32 se muestran como ejemplo las titulaciones de fluorescencia correspondientes a los péptidos denominados **Ébola-11** y **Leucemia-a** con el dominio Nedd4-WW3.

Como puede observarse en estos complejos los efectos de los ligandos en el espectro de fluorescencia son pequeños, lo que sólo nos ha permitido una estimación aproximada de la magnitud de las constantes de asociación para algunos de los ligandos que, en general, presentan afinidades significativamente menores



**Figura 3.32.** Panel superior: espectros de fluorescencia tomados tras la inyección de los ligandos de **Ébola-11** (ILPTA**PEY**ME) y **Leucemia-a** (PPP**VEPT**AP), sobre la proteína Nedd4-WW3, una vez corregidos por la dilución del ligando y normalizados por la concentración, realizados en fosfato sódico 20 mM pH 7.0 a 25 °C. Panel inferior: valores de intensidad de fluorescencia a 340 nm normalizada por la concentración (círculos negros). La línea continua corresponde al mejor ajuste para el modelo de unión sencillo de un ligando a una proteína (ecuación 2.3, sección 2.2.8.3).

que **M1**. En el caso de las diferentes interacciones con Nedd4-WW3 estas constantes están aún en el intervalo adecuado para su determinación mediante una titulación calorimétrica directa, pero en aquellos casos en los que las afinidades son demasiado bajas (interacciones con Yap65-WW y del ligando **PPPY** con ambos dominios) hemos determinado los parámetros termodinámicos mediante la realización de experimentos de competición en los que se ha titulado con el ligando **M1**, previamente caracterizado, sobre una disolución de dominio saturado con el

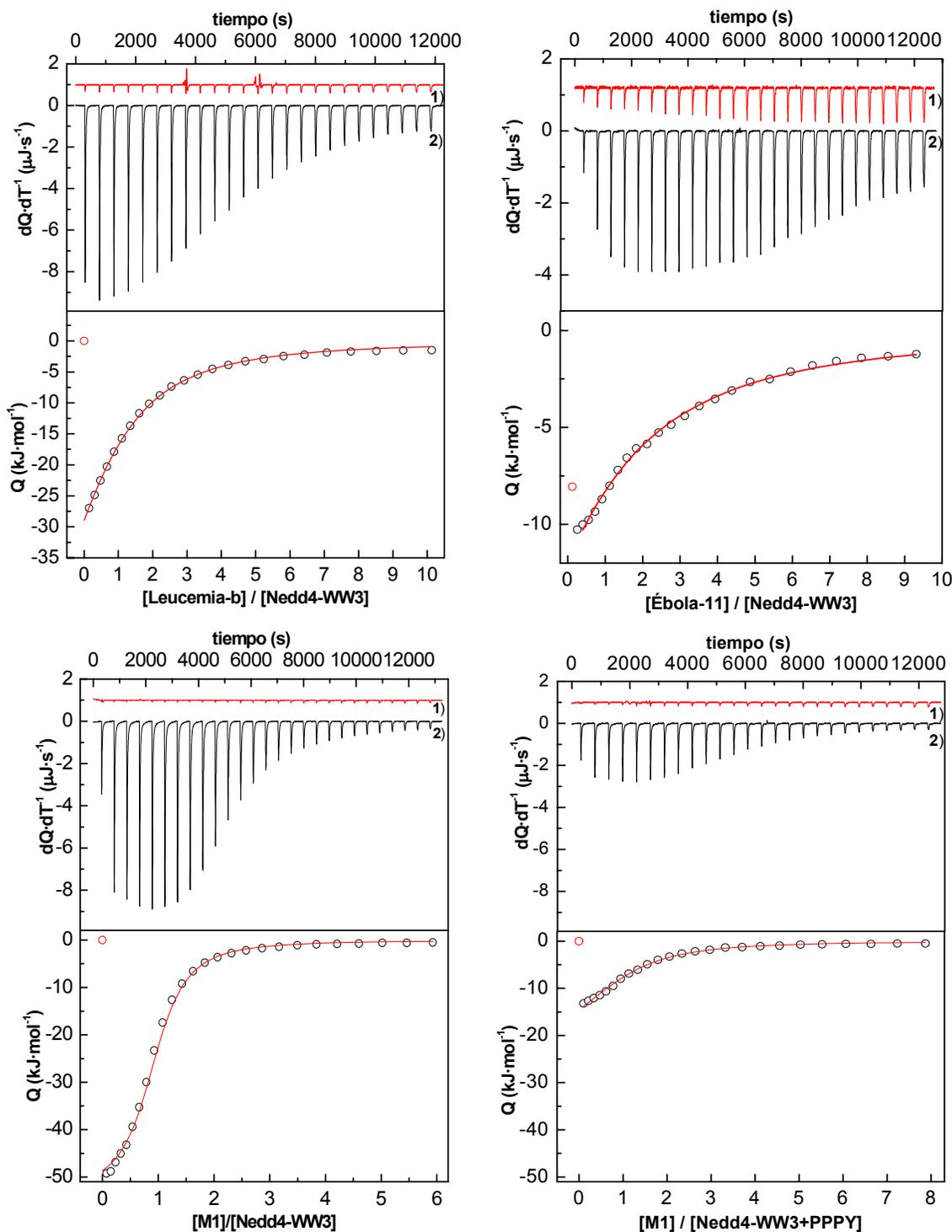
correspondiente ligando de baja afinidad (sección 2.2.9.4). En todos los casos las titulaciones se han preparado en tampón fosfato sódico 20 mM pH 7.0 a 25 °C. En la Figura 3.33 se muestra como ejemplo el experimento directo de titulación de los ligandos **Leucemia-b**, **Ébola-11** y el ligando **M1** con Nedd4-WW3 y el

**Tabla 3.16.** Parámetros termodinámicos obtenidos por ITC para la unión de ligandos viricos *Leucemia-a*, *Leucemia-b*, *Ébola-11*, y del ligando *PPPY* con los dominios *Nedd4-WW3* y *Yap65-WW* en fosfato sódico 20 mM pH 7.0 a 25 °C.

Proteína	Ligando/ secuencia	$K_a \times 10^4$ * (M <sup>-1</sup> )	$\Delta G_{ap}$ (kJ·mol <sup>-1</sup> )	$\Delta H_{ap}$ * (kJ·mol <sup>-1</sup> )	$-T\Delta S_{ap}$ (kJ·(K·mol <sup>-1</sup> ))
Yap65-WW	<b>Leucemia-a</b> <sup>1</sup> PPPYVEPTAP	0.13	-17.8	-34.7	16.9
	<b>Leucemia-b</b> <sup>1</sup> SDPQIPPPYVEP	0.37	-20.4	-35.6	15.2
	<b>Ébola-11</b> <sup>1</sup> ILPTA□PEYME	0.18	-18.6	-49.4	30.8
	<b>PPPY</b> <sup>1</sup>	0.16	-18.3	-40.2	21.9
Nedd4-WW3	<b>Leucemia-a</b> PPPYVEPTAP	0.56	-21.4	-70.3	48.9
	<b>Leucemia-b</b> SDPQIPPPYVEP	1.63	-24.0	-68.2	44.2
	<b>Ébola-11</b> ILPTA□PEYME	0.68	-21.9	-50.7	28.8
	<b>PPPY</b> <sup>1</sup>	0.47	-21.0	-38.1	17.1

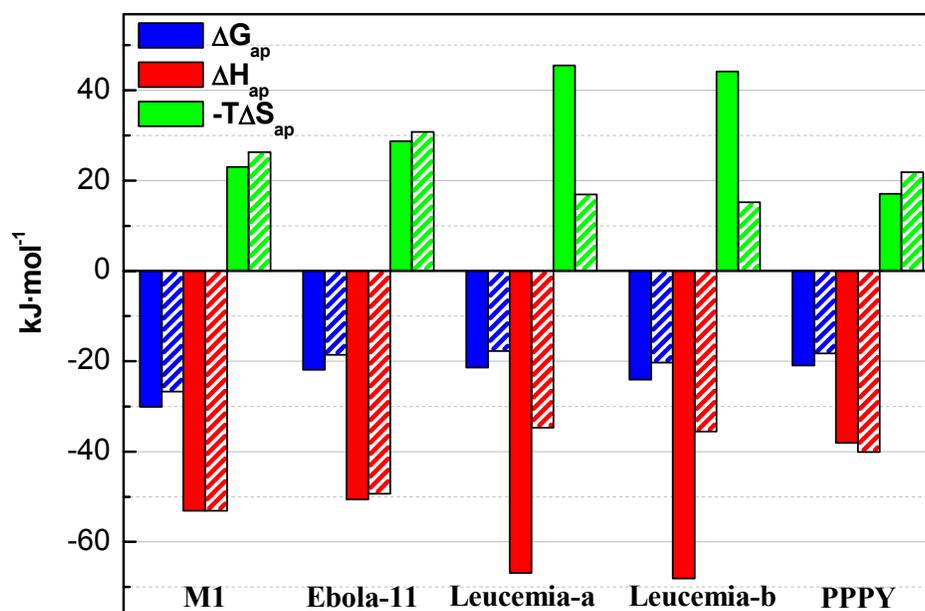
<sup>1</sup> Experimentos de competición con el ligando de afinidad moderada MI, para el que se han determinado previamente sus parámetros termodinámicos de unión a cada dominio.

\* El error en los parámetros termodinámicos obtenidos se estima en torno al 5% para la entalpía de unión y al 10% para la constante de asociación.



**Figura 3.33.** Experimentos de titulación directa (ligandos **Leucemia-b**, **Ébola-11**, y **M1**) e indirecta (ligando **PPPY**) con el dominio **Nedd4-WW3** en fosfato sódico 20 mM pH 7.0 a 25 °C. En la parte superior de cada gráfica, (1), muestra el termograma de dilución del ligando (desplazado en el eje para mayor claridad), y (2), el termograma de titulación formado por los calores por unidad de tiempo liberados tras la inyección de ligando. La parte inferior muestra la isoterma de unión de cada experimento obtenida después de normalizar los calores de titulación y corregir por los calores de dilución. Los datos experimentales se representan como círculos, mientras que la línea continua corresponde al mejor ajuste para el modelo de "n" sitios de unión idénticos e independientes, excepto para el ligando **PPPY** que se ajustan a un modelo de competición para "n" sitios de unión idénticos e independientes (sección 2.2.9.4). En el caso del experimento de competición se mantienen las mismas escalas que en la gráfica del experimento directo con **M1** para mostrar el efecto de inhibidor del ligando **PPPY**.

## Resultados



**Figura 3.34.** Representación en forma de barras de la energía libre de Gibbs y sus contribuciones entálpica y entrópica para la unión de todos los ligandos a la proteína Nedd4-WW3 (barras rellenas) y a Yap65-WW (barras rayadas). Los parámetros termodinámicos se han obtenido a partir de experimentos de calorimetría isotérmica de titulación en fosfato sódico 20 mM, pH 7.0 a 25 °C. Los ligandos **Leucemia-a** y **Leucemia-b** se corresponden con los que tienen la secuencia **PPPY** en el extremo N-terminal (**PPPYVEPTAP**) y centrada (**SDPQIPPPYVEP**) respectivamente.

experimento de competición correspondiente a **M1/PPPY** y Nedd4-WW3, donde se muestran los termogramas (paneles superiores) y las correspondientes curvas isotermas de unión junto con el mejor ajuste a un modelo de “n” sitios idénticos e independientes para los experimentos directos, y al modelo de desplazamiento completo (Sigurskjold 2000) para los experimentos de competición (sección 2.2.9.4.). Los valores para los parámetros termodinámicos de unión obtenidos de estos análisis se resumen en la Tabla 3.16.

Para facilitar la comparación, los parámetros termodinámicos de unión de todos los ligandos con ambos dominios se representan en la Figura 3.34. En todos los casos, los ligandos correspondientes a los dominios L víricos tienen constantes de asociación muy modestas aunque, como es de esperar, el dominio Nedd4-WW3 presenta sistemáticamente una afinidad ligeramente superior por estos ligandos que Yap65-WW. Como puede observarse, al igual que con el ligando **M1**, el proceso de unión está gobernado en todos los casos por un valor de entalpía favorable

(negativo) al que se opone parcialmente una contribución entrópica desfavorable (positiva). Es interesante hacer notar que esto ocurre incluso para el motivo **PPPY** aislado, que presenta un porcentaje muy considerable (entre un 60 y un 80%) de la entalpía de unión observada para los otros ligandos de mayor longitud. Aunque se observan diferencias significativas en cuanto a entalpías de unión, especialmente notables en el caso de los ligandos correspondientes a la leucemia humana (**Leucemia-a** y **Leucemia-b**), el efecto compensatorio de las contribuciones entrópicas hace que las diferencias en términos de energía de Gibbs de unión sean mucho menores.

### **3.3. RECONOCIMIENTO DE LIGANDOS RICOS EN PROLINA POR DOMINIOS UEV**

Para abordar el estudio termodinámico de las interacciones de dominios UEV con secuencias ricas en prolina hemos seleccionado el dominio Tsg101-UEV y varios ligandos peptídicos con secuencias correspondientes a dominios L víricos de tipo PTAP. Las secuencias y nomenclaturas de estos ligandos se recogen en la Tabla 3.17.

#### **3.3.1. Caracterización termodinámica de la interacción entre el dominio L de VIH y Tsg101-UEV**

En estudios previos se ha establecido que la secuencia PE**PTAP**PEE, perteneciente a la proteína p6-gag del virus VIH, presenta un comportamiento equivalente al de la proteína completa y contiene, por tanto, todos los elementos necesarios para la interacción con Tsg101-UEV (O. Pornillos *et al.* 2002; Owen Pornillos *et al.* 2002a). Medidas mediante la técnica de BIAcore indican que este péptido interacciona con Tsg101-UEV con una afinidad moderada ( $K_a = 10^5 \text{ M}^{-1}$ ) (Owen Pornillos *et al.* 2002a). Hemos seleccionado este péptido, que denominaremos **VIH-9**, como punto de partida para la caracterización termodinámica de las interacciones de dominios L de tipo PTAP con Tsg101-UEV. Adicionalmente, con el fin de facilitar la determinación precisa de la concentración de ligando, hemos

## Resultados

**Tabla 3.17.** Selección de ligandos de dominios L de los virus VIH, ébola y de la leucemia humana tipo I para la caracterización termodinámica de las interacciones que establecen con el dominio UEV de Tsg101.

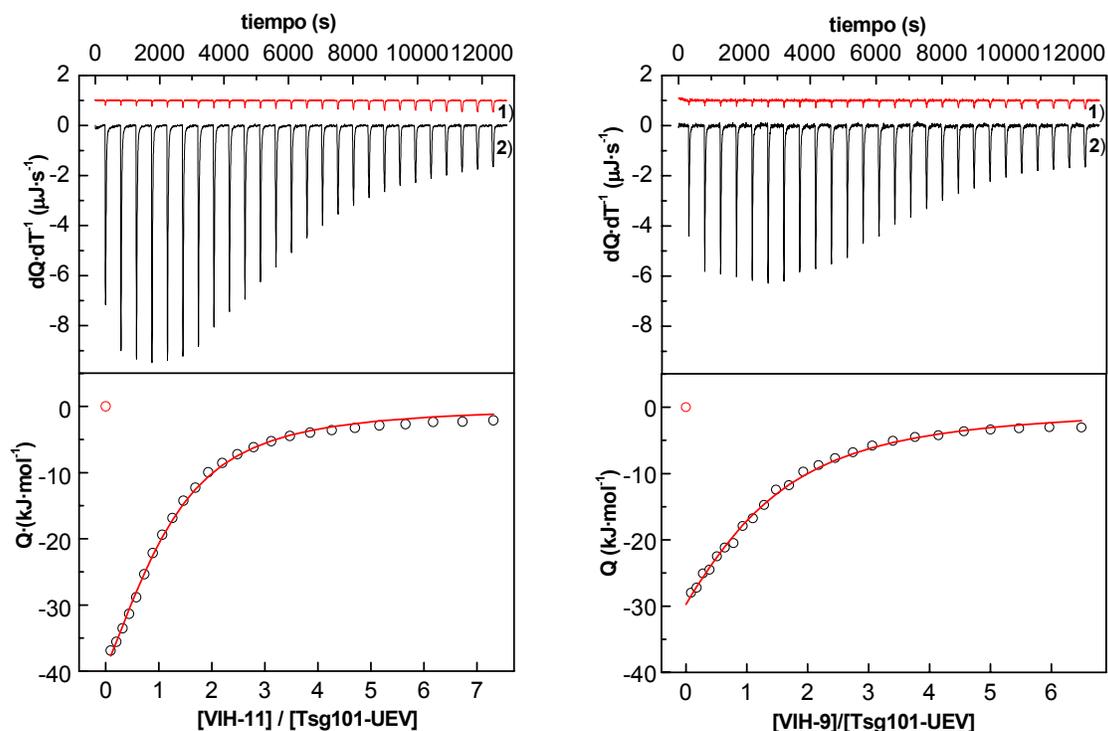
Ligando	Nombre	Secuencia*
VIH p6-gag	VIH-11	PEPTAPPEESF
VIH p6-gag	VIH-9	PEPTAPPEE
ébola Zaire VP40	Ébola-11	ILPTAPEYME
ébola Zaire VP40	Ébola-9	ILPTAPEY
leucemia humana-gag	Leucemia-a	PPPYVEPTAP
leucemia humana-gag	Leucemia-c	YVEPTAPQVL

\*Para identificar más fácilmente la posición de los motivos PPxY y PTAP en la secuencia del ligando, éstos se muestran en rojo y azul respectivamente.

analizado la termodinámica de unión de un segundo ligando **VIH-11**, de secuencia (PE**PTAP**PEESE), en el que se han incluido dos restos adicionales de la secuencia p6-gag en el extremo N-terminal del péptido **VIH-9**.

La energética de unión de ambos péptidos al dominio Tsg101-UEV se ha analizado directamente mediante calorimetría isotérmica de titulación ya que debido a la ausencia de triptófanos en el sitio de unión de Tsg101-UEV en este caso impide el estudio mediante fluorescencia. Teniendo en cuenta que las constantes de asociación parecen ser modestas, las titulaciones se han realizado utilizando un perfil de volúmenes de inyección crecientes (4-21  $\mu$ L) de una concentración de ligando 1.9 mM y una concentración de proteína en célula de 40-60  $\mu$ M.

Las titulaciones de ambos ligandos en fosfato sódico 20 mM, beta-mercaptoetanol 5 mM pH 7.2 a 25 °C se ilustran en la Figura 3.35, donde los paneles superiores muestran los termogramas de titulación y dilución, y los paneles inferiores



**Figura 3.35.** Experimentos de ITC para la unión de ligandos del dominio tipo L del virus VIH con la proteína Tsg101-UEV en fosfato sódico 20 mM, BME 5 mM pH 7.2 a 25 °C. En la parte superior de cada gráfica, (1), muestra el termograma de dilución del ligando (desplazado en el eje para mayor claridad), y (2), el termograma de titulación formado por los calores por unidad de tiempo liberados tras la inyección de ligando. La parte inferior de cada gráfica, muestra la isoterma de unión de cada experimento, obtenida después de normalizar los calores de titulación y corregir por los calores de dilución. Los datos experimentales se representan como círculos, mientras que la línea continua corresponde al mejor ajuste para el modelo de “n” sitios de unión idénticos e independientes (ecuación 2.21).

las correspondientes curvas isotermas de unión junto con los mejores ajustes a un modelo de unión de “n” sitios idénticos e independientes, a partir de los que se obtiene un número de sitios igual a uno. Como se puede observar en la Tabla 3.18, donde se recogen los resultados del análisis, las constantes de asociación determinadas mediante calorimetría ( $K_a = 2 \cdot 10^4 \text{ M}^{-1}$ ) son un orden de magnitud inferiores a las obtenidas en los experimentos de BIAcore ( $K_a = 3.5 \cdot 10^5 \text{ M}^{-1}$ ) técnica para la que en algunas ocasiones se han observado desviaciones significativas en la determinación de la constante de asociación (Kortt *et al.* 1997a; Kortt *et al.* 1997b). La interacción de estos ligandos con Tsg101-UEV presenta el mismo patrón termodinámico observado para la interacción de ligandos ricos en prolina con dominios SH3 y WW, de modo que la asociación presenta una elevada

## Resultados

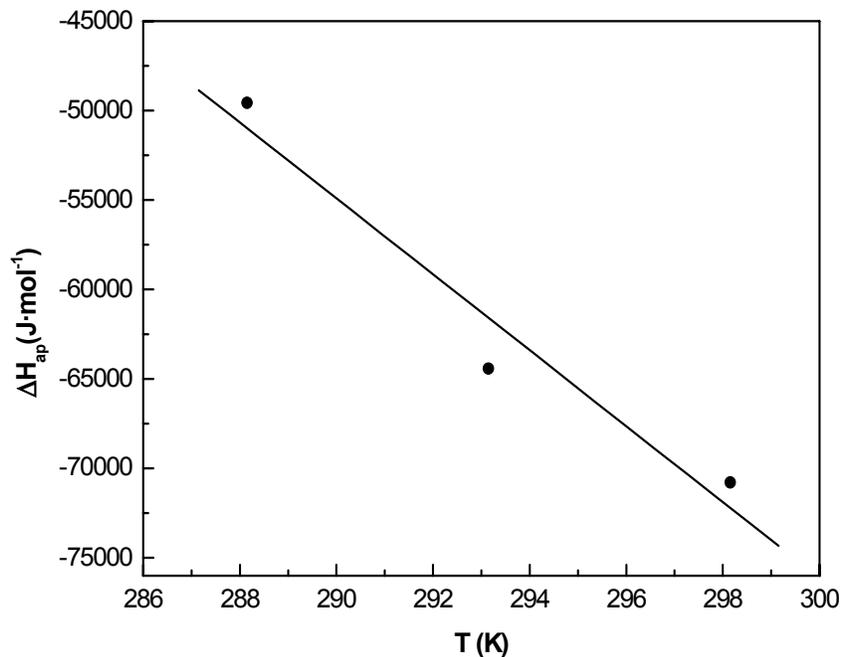
**Tabla 3.18.** Parámetros termodinámicos obtenidos por ITC para la unión del ligando **VIIH-11** (**PEPTAPPEESF**) con la proteína Tsg101-UEV en fosfato sódico 20 mM, BME 5mM, pH 7.2

pH	T (°C)	Tampón	$\Delta H_{ap}^*$ (kJ·mol <sup>-1</sup> )	$K_a^*$ (x10 <sup>4</sup> M <sup>-1</sup> )	$\Delta C_{p,ap}$ (kJ·(K·mol) <sup>-1</sup> )	$n_H^a$	$\Delta H_{int}^a$ (kJ·mol <sup>-1</sup> )
	15	Fosfato	<u>-49.4</u>	2.7			
7.2	20	Fosfato	<u>-64.5</u>	1.9	<u>-2.12</u>		
		Fosfato	<u>-70.3</u> (-69.5)*	2.0 (1.4)			
	25	PIPES	-65.3	1.2		0.5	-72.0
		Imidazol	-53.6	2.7			
5.0	25	Acetato	-63.3	15.4			
3.0	25	Glicina	-48.6	5.0			

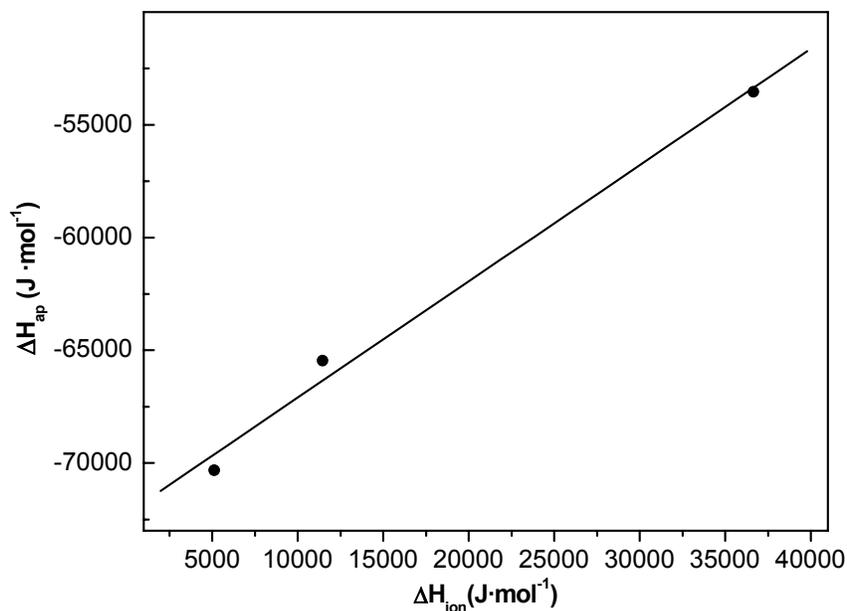
\* Valor de la constante de asociación correspondiente al experimento realizado con el ligando **HIV-9** (**PEPTAPPEE**).

<sup>a</sup> Número de protones intercambiados con el tampón y entalpía intrínseca de unión que se obtuvieron considerando los valores de ionización publicados (Christensen 1976; Fukada & Takahashi 1998) para fosfato sódico ( $\Delta H_{ion}=5128 \text{ J mol}^{-1}$ ), PIPES ( $\Delta H_{ion}=11450 \text{ J mol}^{-1}$ ) e imidazol ( $\Delta H_{ion}=36650 \text{ J mol}^{-1}$ ). Los valores de entalpía subrayados son los utilizados para el cálculo del cambio de capacidad calorífica de unión.

exotermicidad ( $\Delta H_{ap} = -69.5 \text{ kJ}\cdot\text{mol}^{-1}$ ) y contribuciones entrópicas desfavorables. Dado que los parámetros termodinámicos de unión son muy parecidos para los dos ligandos hemos decidido utilizar el ligando de mayor longitud para completar el análisis termodinámico, ya que la determinación de su concentración es menos susceptible de error. El análisis de la dependencia con la temperatura de la entalpía aparente de unión, que se muestra en la Figura 3.36, proporciona un valor para la capacidad calorífica de unión de  $-2.12 \text{ kJ}\cdot(\text{K}\cdot\text{mol})^{-1}$ , el cual es ligeramente superior al obtenido para otros dominios modulares como los



**Figura 3.36.** Representación de la entalpía aparente de unión del ligando **VIH-11** y Tsg101-UEV obtenida de los experimentos realizados en fosfato sódico 20 mM pH 7.0 a 15, 20 y 25°C, frente a su correspondiente valor de temperatura. De la pendiente de dicha recta obtenemos el cambio de capacidad calorífica aparente de unión.



**Figura 3.37.** Representación de la entalpía aparente de unión obtenida para la unión del ligando **VIH-11** a Tsg101-UEV en los tampones fosfato, imidazol y PIPES a 25 °C, frente a la entalpía de ionización de cada tampón. Si extrapolamos a calor de ionización cero obtenemos la entalpía intrínseca de unión, y de la pendiente de dicha recta obtenemos el número de protones intercambiados con el tampón. Dichos valores aparecen recogidos en la Tabla 3.18.

## Resultados

SH3 (Ferreon & Hilser 2004; Palencia *et al.* 2004) y WW estudiados en esta Memoria, cuyo  $\Delta C_{p,ap}$  varia entre 0.7 y 1.8  $\text{kJ}\cdot(\text{K mol})^{-1}$ .

Para evaluar el impacto de posibles equilibrios de ionización acoplados a la formación del complejo en la energética de unión hemos realizado una serie de experimentos en tampones con diferentes calores de ionización y a diferentes valores de pH, cuyos resultados se recogen en la Tabla 3.18. La dependencia de la entalpía aparente de unión con la entalpía de ionización del tampón se ilustra en la Figura 3.37. Como se puede observar, a pH 7.0 la unión del péptido **VIH-11** está asociada a la liberación de 0.5 protones por el tampón que son aceptados por restos de la proteína o el ligando en la interacción. El valor de la entalpía intrínseca de unión ( $-72.0 \text{ kJ}\cdot\text{mol}^{-1}$ ) sigue siendo muy negativo lo que confirma que, al igual que ocurría para dominios SH3, el elevado carácter exotérmico refleja la verdadera naturaleza de la interacción.

Las titulaciones realizadas a pH 5.0 y 3.0 también proporcionan valores de entalpías aparentes muy negativos que, teniendo en cuenta que los tampones utilizados están caracterizados por entalpías de ionización muy pequeñas, es de esperar que sean muy cercanas a las entalpías intrínsecas de unión a estos valores de pH. Como se puede observar la entalpía de unión disminuye considerablemente con el pH, lo que indica que posiblemente el equilibrio de ionización implique algún aminoácido con una entalpía de ionización grande, posiblemente un resto His en la proteína ( $\Delta H_{ion} = 28.5 \text{ kJ}\cdot\text{mol}^{-1}$ ,  $\text{pK}_a$  teórico = 6.0), cuyo  $\text{pK}_a$  puede ser alterado como consecuencia de la unión del ligando.

### 3.3.2. Caracterización termodinámica de las interacciones de los dominios L de ébola y leucemia humana tipo I con Tsg101-UEV

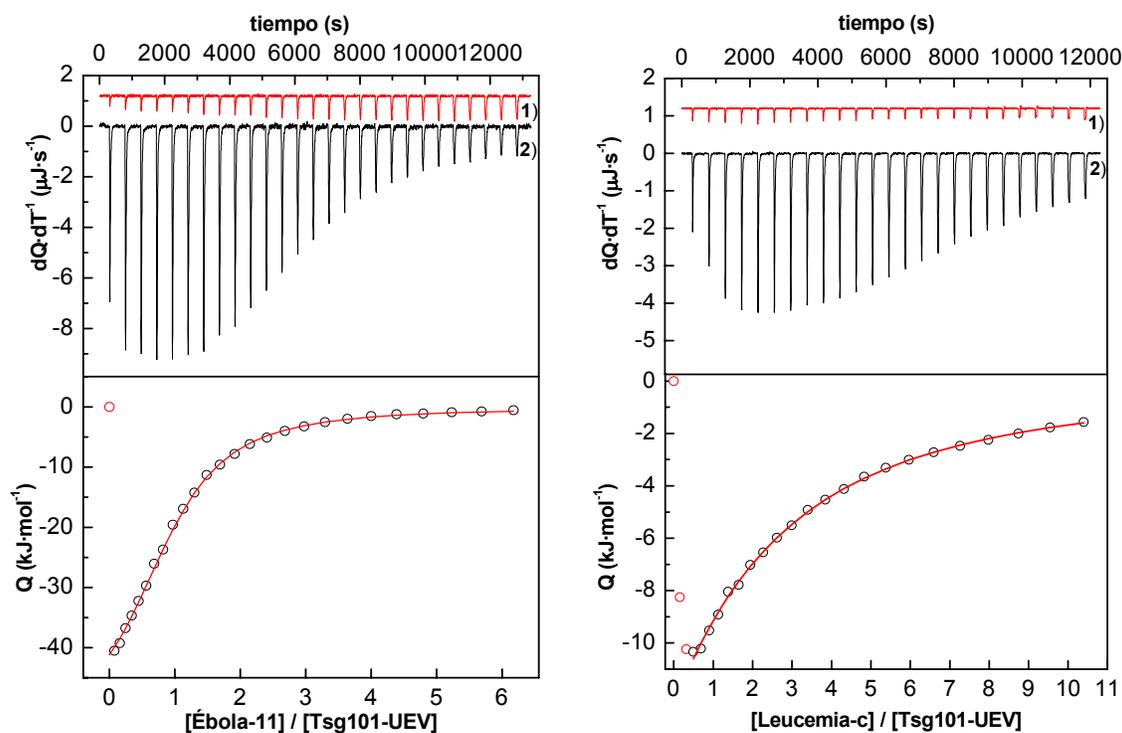
Hemos extendido el análisis termodinámico a los dominios L de los virus ébola (Zaire) y de la leucemia humana tipo I, para los que no existe ningún estudio de unión con el dominio UEV de Tsg101. Para el virus ébola hemos seleccionado dos ligandos de diferente longitud, cuyas secuencias **ILPTA**PEY**** (**Ébola-9**) e **ILPTA**PEY**ME** (**Ébola-11**) corresponden a los restos 5-13 y 5-15 respectivamente

de la proteína estructural VP40 de ébola (Martin-Serrano *et al.* 2001). Hemos diseñado estas secuencias de modo que tengamos el mismo número de restos a cada lado de la secuencia consenso PTAP que en los ligandos del virus VIH. En el caso del virus de la leucemia humana, hemos seleccionado también dos ligandos: el ligando denominado **Leucemia-a** (PPPYVEPTAP), que contiene los dos motivos PPxY y PTAP y cuya interacción con dominios WW se ha descrito en el apartado anterior, además el ligando **Leucemia-c**, que contiene únicamente el dominio PTAP centrado en la secuencia (YVEPTAPQVL); ambas secuencias corresponden a los restos 118 a 127 y 124 a 133 respectivamente de la proteína gag-p19 del virus de la leucemia humana tipo I (Sakurai *et al.* 2004). El análisis de ambos péptidos nos permitirá estudiar la posible influencia de la presencia del dominio L tipo PPPY y otros restos adyacentes al motivo conservado PTAP en la interacción.

Hemos realizado un análisis termodinámico de la interacción de estos ligandos mediante calorimetría isotérmica de titulación similar al descrito para el dominio L del virus VIH. Como condiciones estándar para la realización de los experimentos hemos escogido tampón fosfato 20 mM, beta-mercaptoetanol 5mM pH 7.2, condiciones bajo las que la contribución de los procesos de ionización acoplados a la unión del ligando **VIH-11** a Tsg101-UEV es pequeña ( $< 3\text{kJ}\cdot\text{mol}^{-1}$ ). En la Figura 3.38 se muestran como ejemplos las titulaciones de los ligandos **Ébola-11** y **Leucemia-c** a 25 °C. Los resultados del análisis de las isothermas de unión según el modelo de “n” sitios idénticos e independientes se recogen en la Tabla 3.19.

En el caso de la secuencia del virus ébola hemos obtenido un patrón termodinámico similar para ambos ligandos, aunque en el caso del ligando corto **VIH-9** la afinidad es algo menor ( $2.4\cdot 10^{-4}\text{ M}^{-1}$  frente a  $4.3\cdot 10^{-4}\text{ M}^{-1}$  para el ligando **Ébola-11**). Dado que en el caso del virus del VIH hemos centrado el estudio con el ligando **VIH-11**, en este caso utilizaremos el ligando **Ébola-11** ya que posee el mismo número de restos a cada lado de la secuencia consenso PTAP. En cuanto a los ligandos de la leucemia humana tipo I, el péptido **Leucemia-a** posee una afinidad muy baja ( $K_a < 10^3\text{ M}^{-1}$ ) que imposibilita el estudio directo mediante calorimetría isotérmica de titulación y, dado que para Tsg101-UEV no disponemos de ligandos de mayor afinidad que nos permitieran recurrir a experimentos de competición no ha sido posible su caracterización. Sin embargo, el ligando **Leucemia-c**, con la

## Resultados



**Figura 3.38.** Experimentos de ITC para la unión de los ligandos **Ébola-11** (ILPTA**R**PEYME) y **Leucemia-c** (YVEPTAPQVL) con el dominio Tsg101-UEV en fosfato sódico 20 mM, BME 5 mM pH 7.2 a 25 °C. La parte superior de cada gráfica, (1), muestra el termograma de dilución del ligando (desplazado en el eje para mayor claridad), y (2) el termograma de titulación formado por los calores por unidad de tiempo liberados tras cada inyección de ligando. La parte inferior de cada gráfica muestra la isoterma de unión de cada experimento, obtenida después de normalizar los calores de titulación y corregir por los calores de dilución. Los datos experimentales se representan como círculos, mientras que la línea continua corresponde al mejor ajuste para el modelo de "n" sitios de unión idénticos e independientes.

secuencia PTAP centrada, presenta mayor afinidad ( $K_a = 5.3 \cdot 10^3 \text{ M}^{-1}$ ). Este resultado nos indica, por tanto, *i*) que la presencia de restos adyacentes en el extremo C-terminal a la secuencia consenso PTAP es muy importante para la afinidad de unión, y *ii*) que la presencia del dominio L tipo PPPY en la secuencia del ligando no contribuye significativamente a la afinidad por Tsg101-UEV.

En la Figura 3.39 se representan los valores de los distintos parámetros termodinámicos (energía libre de Gibbs, entalpía y entropía) para la interacción de los distintos dominios L víricos con Tsg101-UEV. En todos los casos se observa el mismo patrón termodinámico, aunque existen diferencias notables entre los ligandos en cuanto a los valores de las constantes de asociación y de las contribuciones

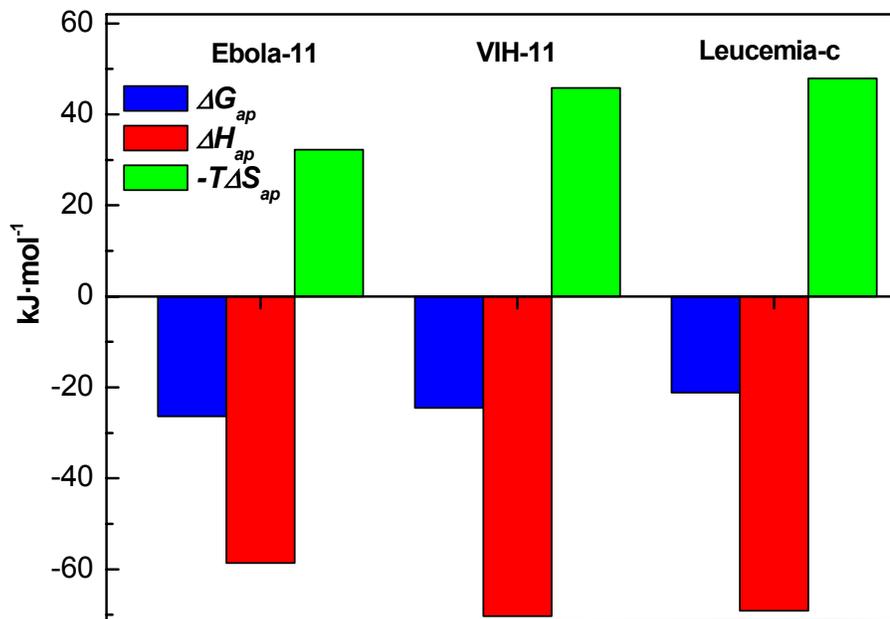
*Tabla 3.19. Parámetros termodinámicos obtenidos por calorimetría de titulación para la unión de los ligandos Ébola-11, Ébola-9, Leucemia-a y Leucemia-c con el dominio Tsg101-UEV en fosfato sódico 20 mM, beta-mercaptoetanol 5 mM pH 7.2 a 20 °C, 25 y 30 °C.*

Ligando/ secuencia	T (°C)	$K_a^*$ ( $\times 10^4 \text{ M}^{-1}$ )	$\Delta H_{ap}^*$ ( $\text{kJ}\cdot\text{mol}^{-1}$ )	$\Delta C_{p,ap}$ ( $\text{J}\cdot(\text{K}\cdot\text{mol})^{-1}$ )
<b>Ébola-9</b> ILPTA <b>P</b> PEY	25	2.4	-61.1	
<b>Ébola-11</b> ILPTA <b>P</b> PEYME	20	5.1	<u>-55.7</u>	
<b>Ébola-11</b> ILPTA <b>P</b> PEYME	25	4.3	<u>-58.6</u>	<u>-925</u>
<b>Ébola-11</b> ILPTA <b>P</b> PEYME	30	3.0	<u>-64.9</u>	
<b>Leucemia-a</b> PPPYVEPTAP	25	< 0.1	-	-
<b>Leucemia-c</b> YVEPTAPQVL	20	0.51	<u>-64.5</u>	
<b>Leucemia-c</b> YVEPTAPQVL	25	0.53	<u>-69.1</u>	<u>-694</u>
<b>Leucemia-c</b> YVEPTAPQVL	30	0.33	<u>-71.6</u>	

*\*El error en los parámetros termodinámicos se estima en torno al 5% en la entalpía de unión y 10 % en la constante de asociación. Los valores de entalpía aparente subrayados son los utilizados en cada ligando para el cálculo de capacidad calorífica de unión.*

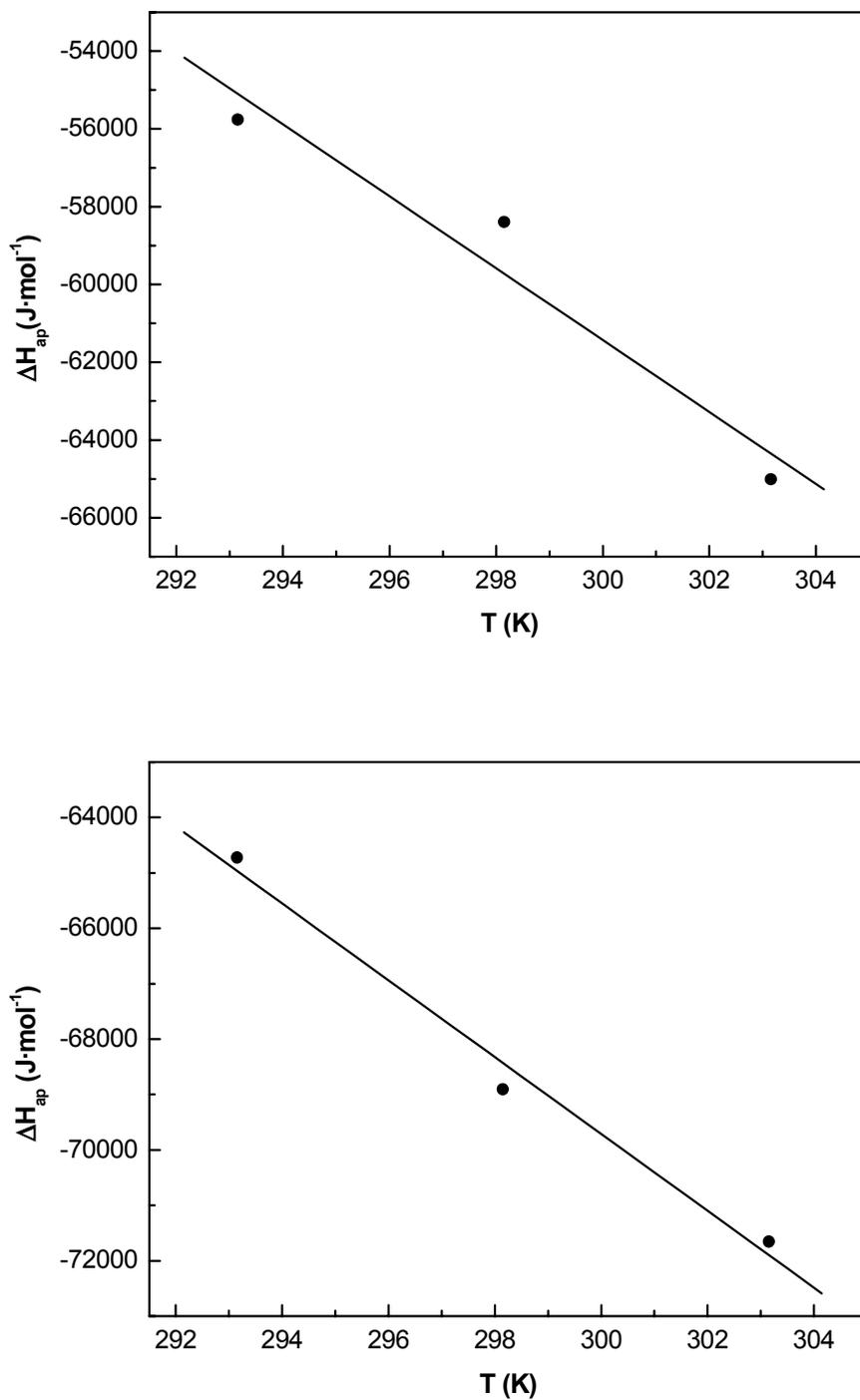
entálpicas y entrópicas. Así, los ligandos **VIH-11** y **Leucemia-c** presentan una entalpía significativamente más favorable que el ligando **Ébola-11**, aunque su contribución entrópica más desfavorable compensa esta diferencia, lo que hace que la afinidad de unión sea muy parecida para **VIH-11** y **Ébola-11**, e incluso significativamente menor en el caso de **Leucemia-c**. De igual forma, se observan diferencias en los valores de las capacidades caloríficas de unión de los tres ligandos,  $\Delta C_{p,ap}$  de -2.12, -0.9 y -0.7  $\text{kJ}\cdot(\text{K}\cdot\text{mol})^{-1}$ , obtenidas para los ligandos **VIH-11**, **Ébola-11** y **Leucemia-c** respectivamente a partir de las gráficas mostradas en las

## Resultados



**Figura 3.39** Representación en forma de barras de la energía libre de Gibbs y sus contribuciones entálpica y entrópica para la unión de los ligandos víricos **Ébola-11**, **VIH-11** y **Leucemia-c** al dominio Tsg101-UEV. Todos los experimentos fueron realizados en fosfato sódico 20 mM, BME 5 mM, pH 7.2 a 25 °C.

Figuras 3.36 y 3.40. Estas diferencias son llamativas teniendo en cuenta las similitudes en composición y longitud de los ligandos.



**Figura 3.40.** Representación de la entalpía aparente de unión de los ligandos **Ébola-11** (panel superior) y de **Leucemia-c** (panel inferior) a Tsg101-UEV obtenida en los experimentos realizados en fosfato sódico 20 mM, BME 5mM, pH 7.2 a 20, 25 y 30 °C, frente a su correspondiente valor de temperatura. De la pendiente de dicha recta obtenemos el cambio de capacidad calorífica aparente de unión.

## **Resultados**

### **3.3.3. Determinación estructural del dominio UEV de Tsg101 mediante difracción de rayos X**

Hemos abordado la cristalización del dominio Tsg101-UEV libre y de sus complejos con los distintos ligandos correspondientes a los dominios L víricos con el fin de obtener información estructural de alta resolución que nos permita racionalizar las diferencias observadas en los parámetros termodinámicos de unión y además investigar la posibilidad de que, al igual que ocurre con el dominio Abl-SH3, pudieran existir moléculas de agua en la interfase de unión que contribuyeran significativamente al carácter fuertemente exotérmico de estas interacciones.

Aunque por el momento no hemos obtenido cristales de suficiente calidad para los complejos, si ha sido posible cristalizar y resolver la estructura cristalográfica del dominio Tsg101-UEV libre (Palencia *et al.* 2006).

#### **3.3.3.1. Cristalización y experimentos de difracción del dominio UEV de Tsg101**

Para la cristalización del dominio UEV de Tsg101 hemos utilizado los reactivos “HR2-110” y “HR2-112” de Hampton Research, que permiten explorar una gran variedad de condiciones experimentales. Hemos realizado experimentos de cristalización de gota colgante (sección 2.2.12.3.) utilizando bandejas de 48 pocillos, mezclando 2  $\mu\text{L}$  de proteína con 2  $\mu\text{L}$  de disolución del reservorio, que contenía un volumen total de 200  $\mu\text{L}$ , a una temperatura constante de 15  $^{\circ}\text{C}$ .

Las condiciones preliminares que originaron microcristales, recogidas en la Tabla 3.20, fueron la nº 9 (PEG 4000 al 30%, sulfato amónico 0.2 M y citrato sódico a pH 5.6), la nº 15 (PEG 8000 al 30%, sulfato amónico 0.2 M, cacodilato sódico 0.1 M a pH 6.5) y la nº 16 (sulfato de litio 1.5 M, HEPES 0.1 M a pH 7.5). Puesto que los cristales de mayor tamaño fueron obtenidos con la condición nº 9, hemos centrado nuestros ensayos en torno a esta condición para obtener cristales de mayor calidad. Para ello, hemos explorado la influencia de algunas sales utilizadas como agente precipitante, la concentración de proteína y diferentes tipos de PEG, así como el

**Tabla 3.20** Condiciones experimentales para la cristalización del dominio UEV de la proteína Tsg101(\*).

Tampón	pH	Precipitante	Precipitante	Técnica	Tiempo aparición	Cristal
Citrato sódico 0.1 M	5.6	Sulfato amónico 0.2 M	PEG 4000 30 %	Difusión vapor Gota colgante	1-2 semanas-	
Cacodilato sódico 0.1 M	6.5	Sulfato amónico 0.2 M	PEG 8000 30 %	Difusión vapor Gota colgante	1-2 semanas	
HEPES 0.1 M	7.5	Sulfato litio 1.5 M	PEG 8000 30 %	Difusión vapor Gota colgante	1-2 semanas	-
TRIS 0.1 M	8.0	Sulfato amónico 0.2 M	PEG 4000 20 %	Difusión vapor Gota colgante	1-2 semanas	
MES 0.1 M	6.0	Sulfato amónico 0.2 M	PEG 4000 20 %	Difusión vapor Gota colgante	1-2 semanas	

(\*) Para todas las condiciones descritas hemos utilizado una temperatura constante de crecimiento de 15 °C.

valor de pH de la disolución a cristalizar. Para estas etapas de optimización hemos empleado bandejas de cristalización de 24 pocillos, en las que hemos mezclado 5  $\mu$ L de disolución de proteína con 5  $\mu$ L de mezcla de reservorio, que contenía 500  $\mu$ L en total, a una temperatura constante de 15 °C. Estos procesos de refinamiento nos han permitido encontrar unas condiciones experimentales idóneas que son las que aparecen recogidas en la Tabla 3.20, en las que al cabo de 1-2 semanas aparecen cristales que continúan creciendo hasta un tamaño aproximado de 1 mm tras 1-2 meses.

Aunque hemos obtenido cristales en dos condiciones casi similares, que se diferencian sólo en el tampón y valor de pH, finalmente los cristales obtenidos a pH 8.0 son los que se han utilizado para difracción de rayos X. Los juegos de datos de rayos X y la estadística del proceso de refinamiento aparecen recogidos en la Tabla

## Resultados

**Tabla 3.21.** Colección de datos de difracción de rayos X y estadística de refinamiento para el cristal obtenido en tampón TRIS a pH 8.0 mostrado en la Tabla 3.20.

<b>Datos de difracción de rayos-X</b>	
Grupo espacial	R3
Dimensiones de celdilla unidad a b c	97.9, 97.9, 110.6
Rango de resolución (Å)	60.0-2.25
Número de observaciones	101426
Reflexiones únicas	17160
Totalidad de datos (%)	92.4 (64.7) <sup>a</sup>
$R_{merge}^b$	0.04 (0.44) <sup>a</sup>
I/σ(I)	26.3 (2.4) <sup>a</sup>
<b>Refinamiento</b>	
Restos de proteína	143
$R_{work}$ (%)	18.9
$R_{free}$ (%)	24.3
<b>Rmsd respecto a la geometría ideal</b>	
Enlaces (Å)	0.020
Ángulos (grados)	1.785
<b>Factores B promedio (Å<sup>2</sup>)</b>	
Esqueleto peptídico	36.8
Cadenas laterales	37.3

<sup>a</sup> Los valores en paréntesis son para el rango de mayor resolución.

<sup>b</sup>  $R_{merge} = \frac{\sum_{hkl} \sum_i |I_i - \langle I \rangle|}{\sum_{hkl} I}$ , donde  $I_i$  es la intensidad para la medida "i" de una reflexión con índices hkl.

3.21. El mejor cristal difractó a una resolución de 2.25 Å y pertenece al grupo espacial R3 con los ejes a = b = 97.9 y c = 110.6 Å para la celdilla unidad, y ángulos alfa = beta = 90.00 grados y gamma = 120.00 grados. La unidad asimétrica obtenida finalmente contiene dos moléculas de proteína (A y B), con 143 restos cada una y 243 moléculas de agua. La calidad del refinamiento nos da, finalmente, un R-factor de 18.8% ( $R_{free} = 24.3\%$ ). El valor de RMSD entre A y B fue de 0.234 Å y todos los restos de la proteína se encuentran mayoritariamente en la región permitida del mapa de Ramachandran, estando el resto en la región favorable.

La resolución de esta estructura ha sido realizada por la Dra. Ana Cámara Artigas (Universidad de Almería, España), empleando el método de

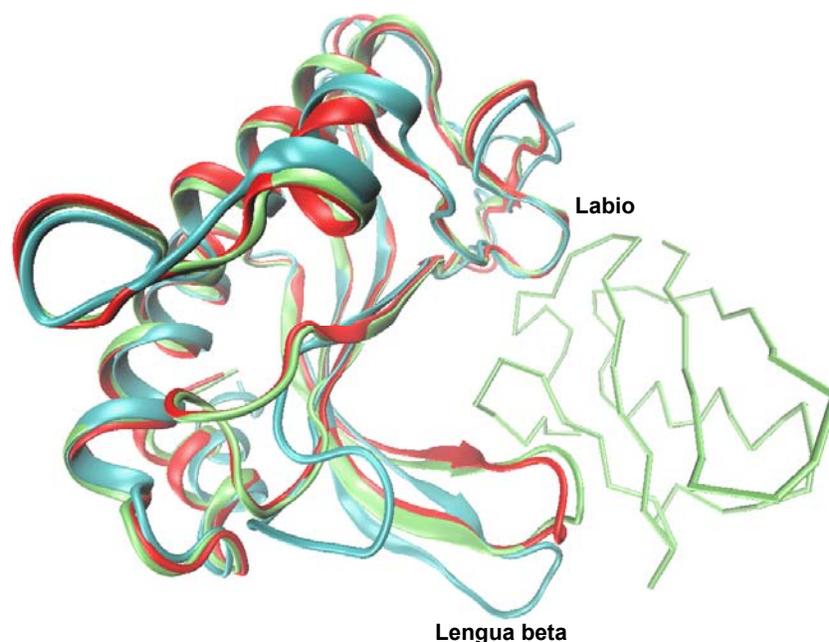
reemplazamiento molecular a partir de la estructura cristalográfica del complejo de Tsg101-UEV con ubiquitina (Pdb: 1s1q).

### **3.3.3.2. Descripción de la estructura del dominio UEV de Tsg101**

La estructura del dominio Tsg101-UEV presenta plegamiento tipo UBC según la base de datos SCOP (Murzin *et al.* 1995) y está compuesta de una lámina beta antiparalela formada por cuatro hebras y de cuatro hélices alfa que se empaquetan contra una de las caras de la lámina beta (Figura 1.10, página 36, antecedentes). Las hebras beta 1 y 2 están conectadas por un giro beta que en conjunto conforman lo que se conoce como “lengua” beta, y la hélice alfa 3 y la hebra beta 4 están conectadas por un lazo desordenado de mayor longitud, denominado “labio”. Los tres primeros restos de la secuencia de la proteína están desordenados y carecen de suficiente densidad electrónica, por lo que no se observan en ninguna de las dos estructuras de la unidad asimétrica.

En la Figura 3.41 se muestra la superposición de las distintas estructuras disponibles para el dominio Tsg101-UEV humano hasta el momento. En dicha figura se comparan las estructuras cristalográficas del dominio libre presentado en esta Memoria, la del complejo con ubiquitina (Pdb: 1s1q) (Sundquist *et al.* 2004) y la estructura en disolución por RMN del dominio libre (Pdb: 1kpq) (Owen Pornillos *et al.* 2002a). La comparación entre la estructura en disolución del dominio Tsg101-UEV libre resuelta por RMN (Pdb: 1kpq) y la estructura cristalográfica del complejo con ubiquitina (Pdb: 1s1q) pone de manifiesto desviaciones de hasta 10 Å en los restos 43-49 de la horquilla beta, lo que se ha interpretado como un posible cambio conformacional en la región de la “lengua” beta. Dado que este lazo está implicado en el reconocimiento de ubiquitina se ha propuesto incluso que este cambio conformacional podría tener relevancia biológica (Sundquist *et al.* 2004). Sin embargo, cuando comparamos ambas estructuras cristalográficas, la nuestra del dominio libre y la del complejo con ubiquitina (Figura 3.41), ambas son prácticamente idénticas y las diferencias, localizadas casi exclusivamente en la región de la lengua beta, son muy pequeñas, con valores de RMSD por debajo de 2 Å, representados en el panel superior de la Figura 3.42.

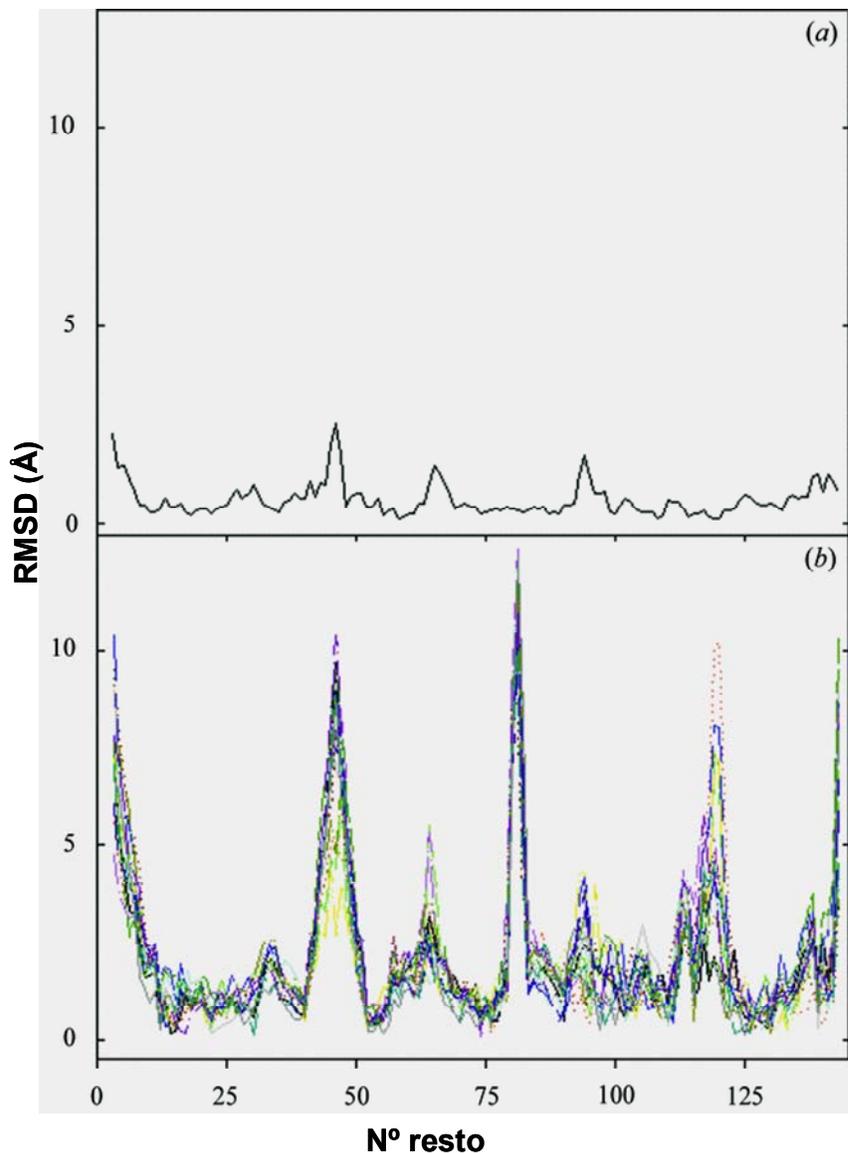
## Resultados



**Figura 3.41.** Superposición de las estructuras de difracción de rayos X de la proteína Tsg101-UEV libre determinada en esta Memoria (rojo, Pdb: 2f0r), formando complejo con ubiquitina (verde lima: Pdb: 1s1) y de la estructura libre en disolución determinada previamente por RMN (turquesa, Pdb: 1kpp).

Además, la comparación entre nuestra estructura de rayos X del dominio Tsg101-UEV libre con la del complejo entre un ligando y el dominio UEV de Vps23 (Pdb: 1uzx) (Teo *et al.* 2004), el homólogo en levaduras de Tsg101, tampoco muestra diferencias significativas en esta región. En el panel inferior de la Figura 3.42 se muestran los valores de RMSD obtenidos para la estructura del dominio Tsg101-UEV libre respecto a los 15 modelos propuestos para la estructura en disolución obtenida previamente. Como se puede observar, el valor promedio de RMSD es de 0.6 Å aunque se observan diferencias muy importantes de hasta 10 Å en la región del lazo 43-49, que, a juzgar por los valores de RMSD obtenidos para el conjunto de modelos de RMN (O. Pornillos *et al.* 2002), son de la misma magnitud que las diferencias observadas con la estructura cristalográfica.

Este desorden aparente no se observa en las estructuras cristalográficas, en las que los factores B de la región de la lengua beta son similares a los valores promedio de la estructura (36 Å<sup>2</sup>). Esta mayor estabilización de esta región en las



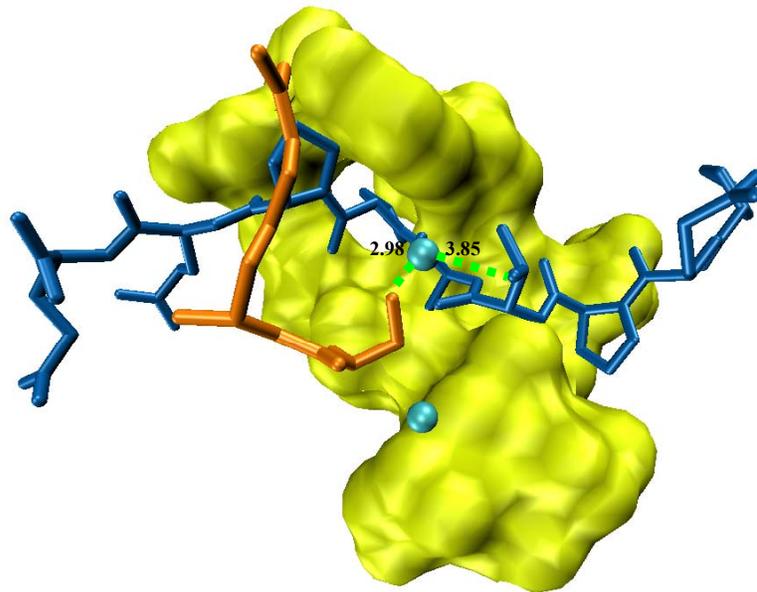
**Figura 3.42.** Comparación de las estructuras disponibles del dominio UEV de Tsg101 humano. a) valores de RMSD obtenidos para los restos de la proteína en la estructura cristalográfica determinada en este trabajo y la estructura cristalográfica del complejo con ubiquitina (Pdb: 1s1q). b) valores de RMSD obtenidos para las estructuras cristalográfica (2for) y en disolución (Pdb: 1kpq, 15 modelos propuestos por RMN) del dominio Tsg101-UEV libre.

estructuras respecto a la de RMN no se debe a efectos de empaquetamiento en el cristal, ya que el único contacto cristalográfico observado es una interacción débil a larga distancia ( $> 3.5 \text{ \AA}$ ) entre el resto N45 en la estructura A y el equivalente en la cadena B de la unidad asimétrica. Además, en el cristal hay espacio suficiente como para permitir que el lazo se mueva libremente y ocupe las posiciones

## Resultados

observadas en las estructuras de RMN. Inspeccionando la Figura 3.41 vemos que también hay diferencias significativas entre las estructuras de rayos X y de RMN en los lazos 78-84 y 116-124. Las mismas discrepancias se observan cuando comparamos nuestra estructura del dominio libre con la del complejo con el péptido **VIH-9** obtenida por RMN (O. Pornillos *et al.* 2002; Owen Pornillos *et al.* 2002a). Por tanto, la diferencia en la conformación de la lengua beta parece ser debida más bien a la propia dispersión proporcionada por los modelos de la estructura en disolución que a un cambio conformacional acoplado al reconocimiento de ubiquitina, como previamente se había postulado (Pornillos *et al.* 2003). Finalmente, hemos comparado las estructuras de la proteína libre en solución y la estructura cristalográfica obtenida en este trabajo con el servidor MOLPROBITY <http://kinemage.biochem.duke.edu/molprobity> (Richardson *et al.* 2003). Mientras que para nuestra estructura encontramos el 93% de los restos en la zona más favorable del mapa de Ramachandran, en el caso de la estructura en disolución tenemos sólo el 73%.

La buena resolución obtenida (2.2 Å) nos ha permitido, además, analizar la presencia de moléculas de agua en el sitio de unión para PTAP que pudieran ser relevantes para la formación del complejo. En la Figura 3.43 se muestra una superposición de nuestra estructura del dominio libre con la del complejo con el ligando **VIH-9** (PE**PTAP**PEE) resuelta por RMN (O. Pornillos *et al.* 2002; Owen Pornillos *et al.* 2002a). Hemos identificado dos moléculas de agua, 195 y 166 (Pdb: 2f0r), localizadas en el sitio de unión para el motivo PTAP que podrían mantenerse en el complejo. Estas moléculas de agua están en contacto con los restos S143 y R144 que, de acuerdo con la estructura de RMN del complejo con **VIH-9**, podrían estar implicadas en interacciones mediadas por moléculas de agua con el oxígeno gamma (donor) del resto T4 en el péptido, de forma similar a lo observado para el dominio Abl-SH3, ya que el puente de hidrógeno directo establecido entre S143 y T4 sólo se observa en algunos de los modelos. El agua 195 es idónea para el establecimiento de dos puentes de hidrógeno, uno con S143 en la proteína y otro con T4 del ligando. Mientras que la distancia entre T4 y S143 en el complejo en disolución es de 4.01 Å, las distancias que se predicen entre el agua 195 con S143 y T4 son 2.98 y 3.85 Å respectivamente, las cuales son más adecuadas para la formación de puentes de hidrógeno (Jiang *et al.* 2005).



**Figura 3.43.** Superposición de la estructura de rayos X de Tsg101-UEV libre (2f0r) y de la estructura en disolución del complejo con el péptido VIH-9 (PEPTAPPEE, Pdb: 1m4p), el cual ha sido representado en estructura de varillas en azul. Para mayor claridad se han ocultado las proteínas y se muestran sólo los restos más importantes para la interacción como superficie en amarillo. Las moléculas de agua 195 Y 166 en la interfase de unión, situadas muy cerca de los restos S143 y R144 (en naranja) se representan como esferas color turquesa. Los puentes de hidrógeno que se predicen con la molécula de agua 195 se representan como líneas verdes discontinuas y sus distancias en Å.

La molécula de agua 166 establece dos puentes de hidrógeno con el resto S143, uno con el oxígeno carbonílico y otro con el oxígeno gamma de su cadena lateral, y posee un factor B similar al de los átomos del resto S143 con el que interacciona.



## **DISCUSIÓN**



## **4. DISCUSION**

### **4.1. RECONOCIMIENTO DE LIGANDOS RICOS EN PROLINA POR DOMINIOS SH3**

#### **4.1.1. La interacción del dominio Abl-SH3 con ligandos ricos en prolina está caracterizada por un patrón termodinámico incongruente con la naturaleza hidrofóbica de la interfase de unión**

Desde la obtención de los primeros datos estructurales, los complejos de dominios SH3 con ligandos ricos en prolina se han presentado con frecuencia como el paradigma de uniones de tipo hidrofóbico basadas en la interacción directa de superficies predominantemente apolares (Mayer 2001; Nguyen *et al.* 1998). No obstante, como ya hemos comentado en la introducción, los escasos estudios termodinámicos llevados a cabo con estos sistemas revelan un patrón termodinámico incongruente con esta descripción (Arold *et al.* 1998; Ferreon & Hilser 2004; Renzoni *et al.* 1996; Wang *et al.* 2001) de modo que, en vez de estar dominadas por el efecto hidrofóbico como cabría esperar, estas interacciones están caracterizadas por entalpías de unión negativas parcialmente compensadas por contribuciones entrópicas desfavorables. Este comportamiento termodinámico sugiere una mayor complejidad en el mecanismo de reconocimiento de secuencias ricas en prolina por parte de dominios SH3, en el que deben entrar en juego factores adicionales a la mera interacción directa entre las interfases de unión que confieren a una unión aparentemente hidrofóbica el patrón característico de interacciones más polares.

Los complejos del dominio Abl-SH3 con ligandos ricos en prolina analizados en este trabajo constituyen un sistema especialmente interesante para el estudio de estos factores ya que, a pesar de presentar menos interacciones polares en la interfase de unión que otros complejos SH3, muestran este patrón termodinámico “anómalo” considerablemente más acentuado, con contribuciones entálpicas y entrópicas a la energía de Gibbs de magnitud muy superior a las obtenidas para otros dominios SH3. Así, la unión del ligando p41 a Abl-SH3 está caracterizada por una entalpía de unión de  $-92 \text{ kJ}\cdot\text{mol}^{-1}$ , unos  $80 \text{ kJ}\cdot\text{mol}^{-1}$  más negativa que lo que

## **Discusión**

cabría esperar para una interacción basada exclusivamente en el contacto entre las superficies hidrofóbicas del ligando y el dominio.

De entre los posibles factores que pudieran contribuir a la elevada exotermicidad de esta interacción, hemos descartado la existencia de equilibrios de ionización acoplados a la unión del ligando que pudieran estar enmascarando la naturaleza real de la interacción. De igual modo, el análisis de las estructuras disponibles para el dominio Abl-SH3 y para su complejo con p41 no revela la existencia de grandes cambios conformacionales en la estructura del dominio asociados a la unión del ligando. Únicamente se observan pequeñas desviaciones en los extremos de los lazos RT y n-Src, que, como se ha descrito para otros dominios SH3 (Arold *et al.* 1998; Wang *et al.* 2001) son las zonas más flexibles del dominio. De hecho, la elevada flexibilidad de estas regiones en el dominio libre está confirmada por la pobre densidad electrónica y la variabilidad conformacional con las condiciones de cristalización que presentan en las estructuras cristalográficas de los mutantes de Abl-SH3 libres obtenidas en este trabajo. La unión del ligando induce una rigidificación de estos lazos que están mucho mejor definidos y ocupando una posición única en las estructuras de los complejos. Efectos similares se han descrito en trabajos previos con otros dominios SH3 que demuestran que la unión del ligando induce cambios significativos en la dinámica de los lazos -RT y n-Src (Arold *et al.* 1998; Ferreon & Hilser 2003b; Wang *et al.* 2001). Es muy probable que la rigidificación de estos lazos, que supone una reorganización de la distribución conformacional del estado nativo del dominio (Freire 1999; Hilser *et al.* 1998; Hilser *et al.* 1997), tenga consecuencias energéticas, aún en ausencia de grandes cambios conformacionales. De hecho, estudios de resonancia magnética nuclear en los que se han analizado las propiedades dinámicas y el intercambio hidrógeno deuterio del dominio SH3 de Src, libre y unido a un péptido rico en prolina, indican que la unión del ligando provoca cambios significativos en la dinámica del esqueleto peptídico, en las longitudes de los puentes de hidrógeno y en los factores de protección al intercambio tanto en la vecindad del sitio de unión como en posiciones muy alejadas de éste, de modo que se produce una propagación efectiva de las interacciones del ligando a través del dominio SH3 (Cordier *et al.* 2000; Wang *et al.* 2001). La formación del complejo, por tanto, conlleva la estabilización selectiva de las conformaciones competentes para la unión, que se traduce en el estrechamiento de

la distribución conformacional y lleva asociado un cambio de entalpía favorable y una contribución entrópica desfavorable (Eftink *et al.* 1983). En resumen, los autores proponen que la modulación de la distribución conformacional del estado nativo de los dominios SH3 como consecuencia de la unión del ligando es uno de los factores determinantes del patrón termodinámico observado en estos sistemas. De hecho, resultados recientes obtenidos en nuestro grupo de investigación, que revelan efectos considerables en la entalpía de unión de ligandos ricos en prolina al dominio SH3 de Src asociados a mutaciones en el lazo distal, a más de 10 Å del sitio de unión (Martín-García, en preparación), apoyan esta hipótesis. Por tanto, la modulación del equilibrio conformacional del dominio SH3 como consecuencia de la unión de ligandos ricos en prolina parece ser un factor importante en la determinación del patrón termodinámico característico de estas interacciones. No obstante, la cuantificación de estas contribuciones no es trivial, por lo que con la información disponible actualmente no nos es posible evaluar si la mayor magnitud de la entalpía y entropía de unión observadas para los complejos de Abl-SH3 podría estar relacionada con estos efectos.

Es importante recordar que para unirse a un dominio SH3, el ligando debe adoptar la conformación PPII, de modo que, al igual que ocurre con la proteína, la formación del complejo supone una redistribución de las poblaciones de los distintos estados conformacionales accesibles al ligando debida a la estabilización selectiva de las conformaciones competentes para la unión. Claramente, esta modulación del equilibrio conformacional del ligando se traducirá en una contribución entrópica desfavorable a la unión. Recientemente, el grupo del Dr. Hilser en la Universidad de Texas ha demostrado, mediante un análisis termodinámico-estadístico combinado con estudios experimentales, que también contribuye favorablemente a la entalpía de unión. Los autores estiman que esta contribución es de unos  $-10 \text{ kJ}\cdot\text{mol}^{-1}$  para un ligando completamente desplegado en disolución (Ferreon & Hilser 2003b; Ferreon & Hilser 2004). En el caso de la unión del ligando p41 a Abl-SH3, es de esperar que estos efectos, ya modestos de por sí, sean mucho menores ya que, debido a su alto contenido en prolina, este péptido está muy estructurado en disolución (Pisabarro *et al.* 1998).

## ***Discusión***

Además de los efectos de ionización y de los efectos conformacionales hemos considerado la posibilidad de un tercer factor determinante del comportamiento termodinámico de los complejos de Abl-SH3: la existencia de interacciones polares mediadas por moléculas de agua ocluidas en la interfase de unión. En efecto, el análisis detallado de las estructuras cristalográficas disponibles para el complejo Abl-SH3/p41 aplicando criterios establecidos anteriormente en la literatura como los más adecuados para la identificación de moléculas de agua persistentes en las interfases de unión (Luque & Freire 2002) nos ha permitido identificar un conjunto de cinco moléculas de agua que participan en una compleja red de puentes de hidrógeno y que median las interacciones entre el ligando peptídico y un grupo de aminoácidos en el dominio, distintos de los que conforman el sitio de unión canónico para el reconocimiento de ligandos poliprolina. Estos restos se encuentran, en todos los casos, a distancias superiores a 5 Å del ligando, excepto Glu98 y Asn114, aunque en ningún caso establecen interacción directa con éste. Como consecuencia, la presencia de moléculas de agua se traduce en la definición de una interfase de unión más extensa y considerablemente enriquecida en restos polares. De este modo, el reconocimiento de p41 por parte del dominio SH3 de Abl parece tener lugar mediante un doble mecanismo que complementa la intercalación de restos apolares en los bolsillos hidrofóbicos con el establecimiento de un elevado número de puentes de hidrógeno mediados por moléculas de agua con restos polares en la periferia del sitio canónico de unión. La existencia de este segundo nivel de interacción, descrito por primera vez en este trabajo para dominios SH3, representa una ruptura con la descripción tradicional de las interacciones de dominios SH3 con sus ligandos y tiene importantes implicaciones, tanto desde el punto de vista energético como funcional. Por un lado, es de esperar que el complejo entramado de puentes de hidrógeno mediados por moléculas de agua contribuya de manera significativa a la energética de unión y, en particular, a la elevada exotermicidad presentada por los complejos de Abl-SH3. Por otro, la presencia de moléculas de agua abre nuevas posibilidades de diseño para la optimización de la afinidad y especificidad de unión en estos sistemas mediante la modulación de las interacciones que éstas establecen.

#### **4.1.2. Los sitios de hidratación observados en el complejo Abl-SH3/p41 están conservados en otros complejos Abl-SH3 y presentan una elevada ocupancia en disolución**

A pesar de que todas las macromoléculas biológicas, y por tanto, las proteínas, desempeñan sus funciones en medio acuoso, el papel del disolvente en el estudio de su estabilidad, dinámica y función ha sido tradicionalmente ignorado. Durante la última década, sin embargo, se han realizado numerosos trabajos sobre los detalles estructurales del agua (Smith *et al.* 2004; Wernet *et al.* 2004), sobre la estructura y propiedades de las moléculas de agua en la superficie de proteínas (Tanya M Raschke 2006), sobre su presencia en cavidades enterradas en el interior de proteínas globulares (Park & Saven 2005) y sobre su papel en la estabilidad y especificidad de complejos proteína-proteína y proteína-ligando (Rodier *et al.* 2005), utilizando una amplia variedad de técnicas (calorimetría diferencial de barrido, fluorescencia de femtosegundos, difracción de neutrones, espectroscopia de RMN o cristalografía de rayos X (Denisov & Halle 1995a; Denisov & Halle 1995b; Denisov & Halle 1996; Halle 2004; Otting *et al.* 1991a; S.K. Pal *et al.* 2002; S. K. Pal *et al.* 2002) que permiten medir la hidratación de proteínas y evaluar las escalas temporales de los procesos de hidratación en las superficies de proteínas, membranas y ácidos nucleicos (Bagchi 2005). Todos estos estudios han puesto en evidencia la versatilidad e importancia del papel que las moléculas de agua desempeñan, más allá del ampliamente estudiado efecto hidrofóbico (Levy & Onuchic 2006).

En el campo del reconocimiento molecular, el interés por el estudio de las moléculas de agua enterradas en las interfases de unión ha ido creciendo conforme se ha incrementado el número de estructuras cristalinas de alta resolución que, con frecuencia, presentan moléculas de agua enterradas en la interfase de unión de complejos proteína-proteína y proteína-ligando que median las interacciones entre ambos mediante el establecimiento de múltiples enlaces de hidrógeno (Bhat *et al.* 1994; Covell & Wallqvist 1997; Janin 1999; Wang *et al.* 1996a; Williams *et al.* 1994). Es interesante señalar que un 6% de todas las aguas cristalográficas se encuentran implicadas en interacciones con la proteína (Reddy *et al.* 2001). Es más, estudios recientes revelan que en complejos proteína-proteína se observa, en promedio, una

## Discusión

molécula de agua por cada 100 Å<sup>2</sup> de superficie (Lo Conte *et al.* 1999) que las interacciones polares mediadas por moléculas de agua son tan abundantes como los enlaces de hidrógeno directos (Rodier *et al.* 2005) y que el 40.1% de los aminoácidos en la interfases de unión interaccionan a través de moléculas de agua (Teyra *et al.* 2006; Teyra & Teresa Pisabarro 2007). A la vista de estos resultados, resulta evidente que el agua, lejos de ser un simple medio inerte, es parte del código de reconocimiento molecular por lo que al ignorar sus contribuciones se corre un serio riesgo de obtener una descripción incompleta e imprecisa de las superficies de unión. El complejo del dominio Abl-SH3 con p41 constituye, sin duda, un claro ejemplo de esta situación, ya que la consideración o no de las moléculas de agua en la interfase de unión proporciona descripción muy distinta de la interacción, tanto en cuanto a su mecanismo como en cuanto a los principios que gobiernan el proceso de asociación.

Sin embargo, aunque numerosos grupos de investigación han abordado el estudio de estas cuestiones en otros sistemas utilizando tanto técnicas experimentales como computacionales (Damjanovic *et al.* 2005; Denisov & Halle 1995a; Denisov & Halle 1995b; Denisov *et al.* 1996; Garcia & Hummer 2000; Zhang *et al.* 2001) hoy en día la cuantificación de la influencia de las moléculas de agua interfaciales en las propiedades termodinámicas de la interacción supone aún un reto importante. En primer lugar, la identificación y localización de moléculas de agua en las interfases de unión no es trivial. Con frecuencia, las estructuras resueltas en disolución presentan importantes interferencias del disolvente y la resolución temporal de la técnica experimental, generalmente RMN, es raramente suficiente para la detección de agua unida a la macromolécula. Conviene, por tanto, disponer de estructuras cristalográficas de alta calidad, con resoluciones preferiblemente superiores a los 2 Å, tanto del complejo como de las proteínas libres. Aún así, hay que tener presente que, en ocasiones, las moléculas de agua observadas en las estructuras pueden ser el resultado de artefactos de cristalización y refinado (Karplus & Sali 1995; Meyer 1992). En este sentido la estructura cristalográfica del complejo Abl-SH3/p41 (código Pdb 1bbz) (Pisabarro *et al.* 1998) se ha obtenido con una muy buena resolución (1.65 Å). Lo mismo ocurre con las dos estructuras de los complejos N114A/p41 y N114Q/p41 resueltas en este trabajo, de resolución 1.75 y 1.85 Å respectivamente, en las que las densidades electrónicas

correspondientes a las moléculas de agua permiten establecer su presencia en el cristal de forma inequívoca.

Una de las estrategias más comunes para evaluar la importancia de cada una de las moléculas de agua en la interacción es la comparación de las posiciones de éstas en varias estructuras de alta resolución, obtenidas para diferentes miembros de una misma familia de proteínas evolutivamente relacionados o bien para la misma proteína bajo diferentes condiciones de cristalización (Loris *et al.* 1999; Nakasako 1999; Sadasivan *et al.* 1998). En este sentido, es importante recordar que las posiciones ocupadas por moléculas de agua en la interfase de unión están muy conservadas en todas las estructuras de complejos con Abl-SH3 disponibles hasta el momento (complejos con los ligandos p41 y 3BP1 e interacción en el contexto de la estructura de la quinasa completa del dominio SH3 con la secuencia de conexión entre el dominio SH2 y el dominio catalítico), como se muestra en las Tablas 3.3 y 3.10 de la sección de Resultados. Así, las posiciones 1 y 2 están ocupadas por moléculas de agua en 9 de las 11 estructuras disponibles y la posición 5 en 10 de ellas. No obstante, las moléculas de agua en posición 4 parecen estar desempeñando un papel central, posiblemente no sólo en la interacción sino también en la propia estabilidad del dominio SH3 ya que no sólo se observan en todas y cada una de las estructuras de complejos Abl-SH3 sino que además están presentes en todas las estructuras del dominio libre resueltas con resoluciones superiores a los 2 Å independientemente de las condiciones de cristalización (N114A pH 3.0, 4.0, 5.0 y 7.0; y N114Q a pH 3.0)

Desgraciadamente carecemos de evidencias experimentales directas acerca de la persistencia de estas moléculas de agua en disolución. No obstante, todas las moléculas de agua observadas en estas estructuras, a excepción de la que ocupa la posición 3 en el complejo WT, presentan factores B muy bajos y similares a los de los aminoácidos con los que interaccionan, lo que es característico de aguas inmovilizadas e indica que la dinámica de las aguas en posición 1, 2, 4 y 5 está determinada por la de las regiones de la proteína y el ligando que las rodean. Es más, los resultados de las simulaciones de dinámica molecular con el complejo WT muestran claramente que estas posiciones están ocupadas por moléculas de agua durante porcentajes muy elevados, alrededor del 90% del tiempo de simulación, en

## **Discusión**

todos los casos. Es interesante señalar que los factores B teóricos calculados a partir de la simulación para las moléculas de agua en posiciones 4 y 5 (17 y 26 Å respectivamente) del complejo WT son muy similares a los experimentales (18 y 20 Å) lo que constituye un buen indicador de la calidad y fiabilidad de las simulaciones computacionales realizadas. La posición 3 presenta una ocupación mucho menor, alrededor de un 10% del tiempo de simulación, lo que está en muy buena concordancia con los mayores valores de ASA y factores B característicos de las moléculas de agua observadas en esta posición, que sólo está ocupada en cuatro de las 11 estructuras cristalográficas. La buena concordancia entre el análisis estructural y los resultados de dinámica molecular también nos confirma que los criterios utilizados para la identificación inicial de estas moléculas de agua a partir de la estructura cristalográfica del complejo (moléculas de agua enterradas en la interfase de unión, a unos 5-7 Å de distancia del ligando y mediando las interacciones entre éste y el sitio de unión mediante el establecimiento de múltiples enlaces de hidrógeno (Luque & Freire 2002; Velazquez-Campoy *et al.* 2000b) proporcionan una buena aproximación a la situación en disolución.

### **4.1.3. Las interacciones mediadas por moléculas de agua contribuyen significativamente al elevado carácter exotérmico de la interacción entre Abl-SH3 y p41**

El análisis detallado de la información estructural disponible para los distintos complejos de Abl-SH3 y los resultados de las simulaciones de dinámica molecular confirman la relevancia de las moléculas de agua en la interfase de unión del complejo con p41 y refuerzan la hipótesis de un mecanismo de unión dual en el que las interacciones mediadas por moléculas de agua parecen desempeñar un papel importante en la interacción. En este punto surge la cuestión sobre cuál es la naturaleza de ese papel y cuál es la magnitud de su impacto en la interacción, tanto en términos termodinámicos como en cuanto a aspectos funcionales como la especificidad de unión. No obstante, a pesar de que en los últimos años se ha avanzado notablemente, tanto experimentalmente (Denisov & Halle 1995a; Denisov *et al.* 1996; S.K. Pal *et al.* 2002; S. K. Pal *et al.* 2002) como desde el punto de vista teórico (Ben-Naim 2002; Ben-Naim 2003; Cooper 2005), en el estudio del papel de las moléculas de agua en las interfases de unión de proteínas, responder a estas

preguntas no es trivial. Dado que estas contribuciones no pueden, en general, deducirse de forma fiable y cuantitativa únicamente a partir de datos estructurales y dinámicos, es necesario combinar este tipo de estudio con el análisis termodinámico detallado de las interacciones para establecer de forma cuantitativa el papel energético de las moléculas de agua.

Las moléculas de agua funcionan como adaptadores que rellenan espacios vacíos, optimizan las interacciones de van der Waals, participan en el establecimiento de puentes de hidrógeno con la proteína y el ligando y pueden contribuir a la disipación de las cargas. Es de esperar que todos estos términos contribuyan favorablemente a la entalpía de unión. Por otro lado, la localización de una molécula de agua en la interfase de unión supone un coste entrópico que se opone a los efectos entálpicos favorables, de modo que, los efectos de las moléculas de agua enterradas en la energía de Gibbs son generalmente modestos. A partir de datos derivados del análisis de sales hidratadas en estado cristalino, se ha estimado que la inclusión de una única molécula en una cavidad o en la interfase de unión de una proteína, llevaría asociado un coste entrópico máximo de unas 8.5 kJ·mol<sup>-1</sup> a 25 °C así como una ganancia entálpica de -16 kJ·mol<sup>-1</sup>, lo que se traduciría en un incremento de -7.5 kJ·mol<sup>-1</sup> en la energía de Gibbs, o lo que es lo mismo, en un aumento en la afinidad de unión de casi un orden de magnitud (Dunitz 1994; Dunitz 1995). Estas predicciones están en buena concordancia con los resultados obtenidos mediante el uso de modelos reticulares sencillos (Cooper 2005) que estiman que cada molécula de agua atrapada en el complejo contribuiría a una entalpía de unión más exotérmica en unos -6 a -12 kJ·mol<sup>-1</sup>, a un aumento en el coste entrópico de unos -11 J·(K·mol)<sup>-1</sup> (3.3 kJ·(K·mol)<sup>-1</sup> a 25 °C) y a un cambio en capacidad calorífica más negativo (-75 J·(K·mol)<sup>-1</sup>). Es de esperar, sin embargo, que, en la práctica, la contribución de cada molécula de agua sea diferente dependiendo de su grado de inmovilización en el sitio de unión y del número y calidad de las interacciones que establezca. La teoría de solvatación de fluidos no homogéneos (Lazaridis 1998a; Lazaridis 1998b), aplicada recientemente al estudio de aguas interfaciales, permite la evaluación de las contribuciones entrópicas y entálpicas de cada molécula de agua y predice que el coste entrópico de las moléculas de agua varía entre los 0 y -41 kJ·mol<sup>-1</sup>, dependiendo de su grado de inmovilización, superando incluso el límite propuesto por Dunitz. De igual manera, esta teoría

## Discusión

estima que la presencia de las moléculas de agua tiene siempre un beneficio entálpico que en algunos casos es muy elevado ( $-63 \text{ kJ}\cdot\text{mol}^{-1}$  para una de las aguas en la interfase de unión de la proteasa del VIH y uno de sus inhibidores (Li & Lazaridis 2003) y en otros mucho más modesto ( $-8 \text{ kJ}\cdot\text{mol}^{-1}$  para una de las aguas en el complejo ciclofilina-ciclosporina A (Li & Lazaridis 2006). Esta descripción es más cercana a las observaciones experimentales que muestran un amplio abanico de efectos que dependen de las características específicas de cada sistema, lo que incrementa la dificultad de predecir las contribuciones de una molécula determinada y pone de manifiesto la necesidad de un análisis termodinámico detallado de cada sistema.

La estrategia más generalizada para el estudio experimental de los efectos de las moléculas de agua interfaciales se basa en la perturbación de las interacciones establecidas por éstas mediante la introducción de modificaciones en el ligando o en la proteína destinadas a la introducción o expulsión de una molécula de agua del complejo y a la posterior evaluación de los efectos de estas modificaciones en la termodinámica de unión. En sistemas como la proteasa del VIH (Lam *et al.* 1994), scytalone deshidratasa (Chen *et al.* 1998), la proteína de unión a péptidos OppA (Sleigh *et al.* 1999) o estreptavidina (Weber *et al.* 1994), se han descrito modificaciones en el ligando que resultan en la eliminación de una o más moléculas de agua del sitio de unión y que se traducen en incrementos significativos en la afinidad de unión, generalmente de origen entrópico. Efectos similares se han descrito asociados a mutaciones en la proteína, como es el caso de las mutaciones Y82F en FKBP12 (Connelly *et al.* 1994), Y120F en MUP-I (Sharrow *et al.* 2005) o F10L en la proteína de unión a zinc ZFY (Lachenmann *et al.* 2004). En general se observa que la eliminación de moléculas de agua de la interfase de unión conlleva beneficios entrópicos asociados con el incremento en los grados de libertad de éstas y está penalizado entálpicamente, con cambios en las entalpías de unión que oscilan entre  $9.5$  (Li & Lazaridis 2005) y  $30 \text{ kJ}\cdot\text{mol}^{-1}$  (Sharrow *et al.* 2005). Sin embargo, el resultado neto en la afinidad de unión es variable, de modo que en algunos casos el desplazamiento de moléculas de agua es energéticamente favorable (Chen *et al.* 1998; Lam *et al.* 1994; Sleigh *et al.* 1999; Weber *et al.* 1994), en otros desfavorable (Chen *et al.* 1998; Clarke *et al.* 2001; Connelly *et al.* 1994; Lachenmann *et al.* 2004;

Lam *et al.* 1994; Sharrow *et al.* 2005; Sleight *et al.* 1999; Weber *et al.* 1994) y en ocasiones neutro (Dall'Acqua *et al.* 1998; Watson *et al.* 1994).

Siguiendo esta estrategia hemos generado un juego de mutantes en el complejo Abl-SH3/p41 destinados a introducir distorsiones en la red de interacciones establecidas por las moléculas de agua en la interfase de unión, bien mediante la eliminación de las cadenas laterales que las coordinan o bien cambiando su longitud. En todos los casos estas mutaciones originan efectos significativos en las contribuciones entálpica y entrópica que se compensan entre sí y resultan en efectos muy pequeños, dentro del error experimental, en la afinidad de unión. La mutación N96A representa una excepción interesante, ya que, se traduce en una disminución muy significativa, unas cuatro veces, en la afinidad de unión. No obstante, es importante recordar que la zona de la proteína en la que se encuentra, el lazo n-Src, ha resultado ser muy sensible a las mutaciones. En particular, las mutaciones en posición 94 parecen alterar dramáticamente el equilibrio conformacional del dominio, como lo demuestran los experimentos de DSC y el número de sitios anormalmente bajo obtenido en los experimentos de titulación. Como se ilustra en la Figura 3.5, página 161, la cadena lateral del resto Asn94, además de coordinar la molécula de agua en posición 1, se encuentra implicada en varios enlaces de hidrógeno intramoleculares con átomos del esqueleto peptídico del resto Trp99 y con la cadena lateral de Asn96 que parecen ser muy importantes para la estabilidad y propiedades de Abl-SH3. De hecho, este enlace de hidrógeno Asn94-Asn96 se observa durante un porcentaje muy elevado del tiempo de simulación en las dinámicas moleculares de Abl-SH3 y de los dos mutantes N114A y N114Q. No obstante, los efectos observados en posición N114, en la zona de la hélice 3<sub>10</sub> son más fáciles de interpretar. Las mutaciones N114A, N114Q y N114T introducidas en esta posición no provocan alteraciones significativas en la estructura y comportamiento del dominio, lo que nos permite, por tanto, atribuir limpiamente sus efectos en la termodinámica de unión en términos de su impacto en las interacciones mediadas por las moléculas de agua. En este sentido, la combinación del análisis termodinámico con la caracterización estructural y el estudio de la dinámica del complejo WT y de los mutantes, nos ha proporcionado una descripción bastante detallada del papel que desempeñan las moléculas de agua en esta región del complejo.

## **Discusión**

De acuerdo con la estructura cristalográfica del complejo Abl-SH3/p41, la eliminación de la cadena lateral de Asn en el mutante N114A, tendría como resultado directo la pérdida de tres enlaces de hidrógeno, uno establecido con la molécula de agua en posición 5 y caracterizado por una distancia de enlace pequeña, sobre 2.5 Å en promedio entre las cuatro estructuras, y otros dos con distancias de enlace menos favorables establecidos con las moléculas de agua en posición 4 (3.0 Å) y 5 (2.8 Å). Sería de esperar que esta pérdida de interacciones se tradujera en una entalpía de unión más desfavorable, probablemente asociada a una contribución entrópica más favorable originada por una mayor movilidad de las moléculas de agua en posición 5. Estos efectos serían máximos si la mutación resultara en la desaparición de esta molécula de agua de la interfase de unión. El análisis termodinámico muestra que la interacción de p41 con el mutante N114A es significativamente menos exotérmica, aunque la magnitud del cambio entálpico ( $\Delta\Delta H = +6.3 \text{ kJ}\cdot\text{mol}^{-1}$ ) está en el límite inferior de los valores predichos por los estudios teóricos y de los datos experimentales descritos para otros sistemas. No obstante, es importante tener en cuenta que este valor refleja sólo el debilitamiento parcial de las interacciones establecidas por la molécula de agua en posición 5 y no su eliminación. En efecto, la estructura cristalográfica del complejo N114A/p41 muestra que este sitio de hidratación continúa ocupado por una molécula de agua que presenta un factor B sólo ligeramente superior al de la molécula equivalente en el complejo WT, lo que refleja que sigue firmemente unida en la interfase del complejo. En esta línea, las simulaciones de dinámica molecular indican que esta mutación resulta únicamente en una reducción del 25% en el tiempo de ocupación de esta posición de hidratación. Además, es interesante señalar, que en el complejo WT la interacción de la cadena lateral de Asn114 con las moléculas de agua, que de acuerdo con los datos cristalográficos aportaría tres enlaces de hidrógeno, se establece sólo durante menos del 30% y un 15% del tiempo de simulación para las posiciones 4 y 5 respectivamente, de modo que las moléculas de agua se mantienen en esta posición 5 fundamentalmente gracias a su interacción con el carbonilo de la prolina en posiciones 6 y 7 del ligando (presentes casi un 85% del tiempo) que presenta distancias y ángulos más favorables. Como consecuencia, la mutación N114A supone sólo una ligera perturbación del entorno de estas moléculas de agua, de modo que el debilitamiento de sus interacciones con el ligando y la proteína se ve compensada por el reforzamiento de la interacción entre ellas. Por lo tanto, nuestro

análisis indica claramente que, en efecto, la presencia de estas dos moléculas de agua en la interfase de unión contribuye muy significativamente a la elevada exotermicidad de la interacción. Aunque es aventurado especular, si extrapolamos los  $6.3 \text{ kJ}\cdot\text{mol}^{-1}$  asociados a la pérdida de un 25% de ocupancia al 100%, podemos considerar, como una estimación aproximada, que la contribución de la molécula de agua en posición 5 a la entalpía de unión podría estar en torno a los  $25 \text{ kJ}\cdot\text{mol}^{-1}$ .

La mutación N114Q se diseñó para investigar la posibilidad de que la mayor longitud de la cadena lateral de Gln frente a Asn permitiera la interacción directa de ésta con el péptido y la eliminación de alguna de las moléculas de agua en la interfase. Los resultados de nuestro análisis indican, sin embargo, que éste no es el caso. La molécula de agua en posición 5 sigue estando presente en la estructura cristalográfica y presenta un factor B similar o incluso ligeramente inferior al observado en el complejo WT. Esto podría atribuirse a un artefacto de cristalización ya que, como hemos comentado en la sección de resultados, la cadena lateral de Gln, que se dirige hacia el disolvente y fuera del sitio de unión, se encuentra estabilizada en esta conformación de forma artificial debido al establecimiento de contactos cristalinos con moléculas vecinas. Sería posible pensar que en disolución, la cadena lateral de Gln estaría dirigida hacia el interior del sitio de unión y podría desplazar la molécula de agua. No obstante, las simulaciones de dinámica molecular muestran que, aunque en efecto la Gln está implicada durante un elevado porcentaje del tiempo de simulación en interacciones en el sitio de unión, la ocupancia de la posición 5 aumenta significativamente con respecto al complejo WT hasta llegar a un 93%. El aumento de longitud de la cadena lateral asociado a la mutación N114Q se traduce, por tanto, en una reorganización de las interacciones mediadas por las moléculas de agua 4 y 5 que resultan entálpicamente más favorables. De este modo, aunque la interacción con las moléculas de agua en posición 4 desaparece, la frecuencia de formación del enlace de hidrógeno entre la cadena lateral de Gln y las moléculas de agua en posición 5 aumenta de un 15 a un 35%, debido fundamentalmente a que la interacción presenta una distancia de enlace mucho más favorable (unos  $2.75 \text{ \AA}$  para Gln114 frente a  $3.2 \text{ \AA}$  para Asn114). De igual modo, se observa que la interacción entre las moléculas de agua en posición 5 y los carbonilos de las prolinas 6 y 7 del ligando (establecida en total casi un 15% y un 70% del tiempo de simulación respectivamente en el complejo WT) se establece con

## **Discusión**

menos frecuencia y está distribuida entre ambas prolinas de forma equivalente (aproximadamente un 30% del tiempo de simulación con cada una), ya que, debido a la mayor longitud de la cadena la geometría de la interacción con la prolina 6 se hace más favorable. Al igual que ocurría con el mutante N114A, el efecto más notable es el fortalecimiento de la interacción entre las moléculas de agua 4 y 5. En resumen, esta reorganización de interacciones se traduce en que la molécula de agua 5, que en el complejo interaccionaba predominantemente con el ligando, en el mutante N114Q está más unida a la proteína. Esto se refleja en la frecuencia de intercambio de unas moléculas de agua por otras en esta posición, que es similar a la observada para el mutante N114A. De este modo, pasamos de 46 moléculas, algunas con tiempos de residencia superiores al nanosegundo, en el complejo WT a 168 moléculas con tiempos de residencia promedio de 66 ps, ninguna de las cuales supera los 380 ps de residencia.

En resumen, los resultados del análisis termodinámico, estructural y dinámica realizado con el complejo Abl-SH3/p41 y los mutantes en posición 114 indican claramente que las moléculas de agua en la interfase de unión, y en particular, las encontradas en las posición 4 y 5 de la región de la hélice  $3_{10}$ , juegan un papel determinante en la energética de unión y contribuyen significativamente a su marcado carácter exotérmico. No obstante, nuestros resultados son menos concluyentes en cuanto a su contribución a la afinidad de unión. Hasta el momento sólo podemos afirmar con seguridad que las mutaciones introducidas, que sólo distorsionan las interacciones establecidas en la interfase de unión, tienen efectos muy pequeños, dentro del error experimental, en la energía de Gibbs. No obstante, es posible que el desplazamiento completo de estas moléculas de agua de la interfase de unión pudiera tener mayores consecuencias energéticas.

### **4.1.4 La presencia de moléculas de agua en la interfase de unión es un fenómeno muy extendido entre los distintos dominios SH3 y puede ser en parte responsable de la especificidad de unión**

Además de aumentar la afinidad de unión, las moléculas de agua en las interfases biomoleculares pueden desempeñar funciones muy importantes en cuanto a la determinación de la especificidad de unión ya que constituyen elementos muy

versátiles: pueden desempeñar el papel de donador y aceptor en los enlaces de hidrógeno y debido a su pequeño tamaño imponen muy pocas restricciones estéricas de modo que, siendo menos voluminosas que cualquier cadena lateral polar, pueden participar en el establecimiento de múltiples enlaces de hidrógeno. Como consecuencia, en algunos casos las moléculas de agua proporcionan gran adaptabilidad a las superficies moleculares permitiendo un elevado grado de promiscuidad en las interacciones. Este es el caso de la proteína bacteriana de unión a oligopéptidos OppA que interacciona con péptidos de dos a cinco restos de longitud con independencia de su secuencia, lo que se consigue combinando dos tipos de interacciones: la interacción directa de la proteína con el esqueleto peptídico y la inserción de las cadenas laterales variables en bolsillos de unión hidrofóbicos que modifican su contenido en agua para optimizar las interacciones. De este modo, las moléculas de agua modulan la especificidad de la proteína ocupando sólo ciertas posiciones de entre un conjunto de posiciones conservadas que corresponden a sitios de hidratación bien definidos y altamente favorables en la superficie de la proteína (Tame *et al.* 1994; Tame *et al.* 1996). Por el contrario, en otras ocasiones, como es el caso del reconocimiento de ligandos fosfotirosina por parte de dominios SH2 (Henriques & Ladbury 2001), los enlaces de hidrógeno mediados por moléculas de agua parecen ser esenciales para mantener la especificidad por el resto glutámico dos posiciones posterior al resto fosfotirosina.

Dado que el trabajo que presentamos en esta memoria es la primera descripción de la presencia de moléculas de agua en la interfase de unión de dominios SH3, el papel que éstas desempeñan en la especificidad de unión en estos sistemas, que hasta el momento se han descrito en términos exclusivamente de interacciones hidrofóbicas, está por determinar. En primer lugar habría que establecer si este fenómeno es universal a todos los dominios SH3 o por el contrario es una característica distintiva de algunos de ellos. Esta cuestión es difícil de elucidar hoy por hoy ya que, debido a su pequeño tamaño, la técnica más utilizada para la determinación estructural de dominios SH3 y sus complejos es la resonancia magnética nuclear que no proporciona información acerca del disolvente. No obstante, se dispone de algunas estructuras cristalográficas de dominios SH3 distintos de Abl unidos a ligandos peptídicos (1bbz (Pisabarro *et al.* 1998), 1abo (Musacchio *et al.* 1994), 2o88 (Camara-Artigas *et al.* 2007)) o bien en el contexto de

## **Discusión**

la quinasa completa (1opk (Nagar *et al.* 2003)), que hasta el momento no habían sido analizadas en cuanto a la presencia de interacciones mediadas por moléculas de agua.

En el complejo Abl-SH3/p41 se observan moléculas de agua en dos regiones diferentes dentro del sitio de unión: la zona de la hélice 3<sub>10</sub> y la región del lazo n-Src. El análisis de dinámica molecular revela que las moléculas de agua presentes en ambas regiones forman dos grupos independientes con propiedades y comportamientos muy diferentes. Esto posiblemente está relacionado con las características específicas que cada una de estas regiones tiene en el contexto del propio dominio SH3. Estudios de intercambio hidrógeno-deuterio combinados con análisis termodinámico-estructurales han puesto de manifiesto que el sitio de unión de estos dominios está compuesto por algunas de las regiones más estables de la proteína, que incluyen las zonas centrales de las hebras beta y la región de la hélice 3<sub>10</sub> y por las zonas más inestables que coinciden con el lazo n-Src.

### **4.1.4.1. Moléculas de agua en la región de la hélice 3<sub>10</sub>**

En la región de la hélice 3<sub>10</sub> se observan dos sitios de hidratación claramente definidos en las estructuras cristalográficas (posiciones 4 y 5) que también lo están según los resultados de dinámica molecular de modo que todas las moléculas de agua que entran en esta región ocupan una de las dos posiciones y se ven implicadas en un conjunto de interacciones muy específicas. Las moléculas de agua que ocupan estos sitios de hidratación se encuentran mediando las interacciones entre la hélice 3<sub>10</sub>, que es la zona más estable estructuralmente y la más conservada en cuanto a la secuencia entre los distintos dominios SH3 y el motivo central poliprolina de p41 que es también la zona más estructurada y conservada entre los ligandos característicos de dominios SH3. Dado que las moléculas de agua que ocupan los dos sitios de hidratación en esta región presentan características bien diferenciadas y parecen desempeñar papeles diferentes en la interacción se han analizado individualmente.

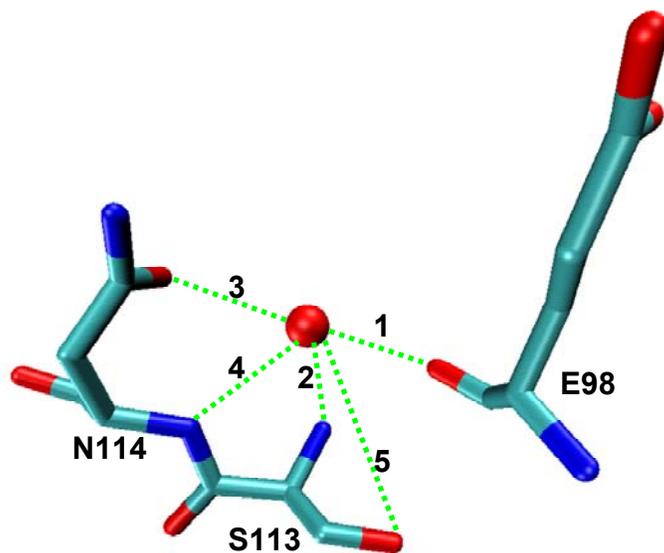
Las moléculas de agua en posición 4 están caracterizadas por tiempos de residencia muy altos que, con frecuencia, superan el ns, lo que es característico de aguas fuertemente unidas a la interfase de unión (Denisov & Halle 1995a; Denisov & Halle 1995b; Garcia & Hummer 2000; Levitt & Park 1993; Otting *et al.* 1991b). Estas moléculas raramente se intercambian con el disolvente y cuando esto ocurre son rápidamente reemplazadas por otras, de modo que esta posición está ocupada durante prácticamente el 100% del tiempo de simulación por un número pequeño de moléculas de agua, siempre inferior a 12, para cualquiera de las variantes del complejo Abl-SH3/p41. Los factores B que caracterizan a estas moléculas de agua en las estructuras cristalográficas son muy bajos, cercanos en muchos casos al valor mínimo de la estructura. Como ya hemos señalado, esta es la única posición que se encuentra ocupada en todas las estructuras cristalográficas de Abl, tanto en los complejos como en todas las estructuras de los dominios libres que tienen suficiente resolución como para permitir el análisis del disolvente. Como consecuencia, estamos ante moléculas de agua fuertemente unidas a la proteína que pueden considerarse como aguas estructurales que forman parte integral del sitio de unión del dominio.

Es importante señalar que las moléculas de agua en posición 4 interactúan fundamentalmente con átomos del esqueleto peptídico, con la excepción del puente de hidrógeno establecido con la cadena lateral del resto Asn114, que, como se ilustra en la Figura 3.15, página 161, está muy conservado entre los distintos dominios SH3. Dado que estas interacciones no estarían afectadas significativamente por variaciones locales en la secuencia, se podría pensar que existe una elevada probabilidad de que moléculas de agua en esta posición estén presentes en otros dominios SH3. De hecho, el análisis de estructuras cristalográficas de otros dominios SH3 libres muestra que, en casi todos los casos en los que la resolución de la estructura es buena, esta posición está ocupada por moléculas de agua caracterizadas por valores bajos de los factores B (ejemplos de esta situación son las estructuras 1jqq, 1oot, 1qkw, 1ruw o 2g6f). Existen, no obstante algunas excepciones que se pueden, en principio asociar, a diferencias en la secuencia de la región de la hélice  $3_{10}$  (como es el caso, por ejemplo, de la estructura 1wyx).

## Discusión

En cuanto a su presencia en los complejos, el análisis de las estructuras cristalográficas de dominios SH3 de Abl y de la familia Src (muy relacionados en la secuencia) disponibles hasta la fecha (dominios SH3 de las quinasas de Abl, Pdb: 1opk (Nagar *et al.* 2003), de las quinasas de Src, cuyos ficheros Pdb son 1fmk (Xu *et al.* 1997), 2ptk (Williams *et al.* 1997), 2src (Xu *et al.* 1999) y 1y57 (Cowan-Jacob *et al.* 2005), y de las quinasas de Hck, con Pdb: 2c0i, 2c0o y 2c0t (Burchat *et al.* 2006)) muestra que, en efecto, se observan moléculas de agua en todas ellas en una configuración muy similar a la observada para el complejo Abl-SH3/p41, que se muestra en la Figura 4.1, donde a cada interacción se le ha asignado un identificador. En la Tabla 4.1 se recogen las identidades de las moléculas de agua y las distancias de enlace correspondientes a cada una de las interacciones en las distintas estructuras que, como se puede observar están muy conservadas en todas ellas. Es interesante señalar que, al igual que ocurría con los dominios libres, en complejos SH3 más dispares en cuanto a secuencia, se observa también con mucha frecuencia la presencia de moléculas en esta posición, como es el caso de 1uzk, 1w70 o 1uti. En resumen, las moléculas de agua en posición 4 pueden considerarse como parte integrante del dominio Abl-SH3, son posiblemente importantes para su estabilidad y forman parte del sitio de unión. El hecho de que su presencia no dependa del ligando y que se encuentren en las regiones más estables del dominio tiene importantes implicaciones en cuanto al diseño.

Las moléculas de agua en posición 5 presentan tiempos de residencia promedio significativamente menores que los observados para la posición 4, que oscilan entre los 175 ps para el complejo WT y los 66 ps para el mutante N114Q. No obstante, estas moléculas aún entran dentro del intervalo de 0.1-1 ns característico de aguas fuertemente unidas al complejo (Denisov & Halle 1995a; Denisov & Halle 1995b; Garcia & Hummer 2000; Levitt & Park 1993; Otting *et al.* 1991b). De hecho, en la simulación para el WT hay varias moléculas de agua que superan los 0.5 ns. Como ya hemos visto, el análisis de las interacciones indica que estas moléculas de agua interaccionan fundamentalmente con el ligando, mientras que la única conexión con el sitio de unión de la proteína se establece mediante los enlaces de hidrógeno con el resto en posición 114 y a través de las moléculas de agua en posición 4.



**Figura 4.1.** Puentes de hidrógeno establecidos por la molécula de agua 4 en la interfase de unión de Abl-SH3/p41 (Pdb: 1bbz) con los restos E98, S113 y N114. Los átomos de cada resto se representan en código universal de colores y la molécula de agua 4 (1082 de la estructura a/b del complejo Abl-SH3/p41) como una esfera roja. Las distancias para cada puente de hidrógeno aparecen recogidas en la Tabla 4.1.

Posiblemente, los tiempos de residencia más cortos reflejan la mayor flexibilidad conformacional del ligando comparado con la hélice  $3_{10}$  del dominio. Esta molécula de agua está presente en la gran mayoría de las estructuras de los complejos de Abl y también en la estructura de la quinasa completa. En esta última estructura, la molécula de agua en posición 5 está coordinada por la cadena lateral de Asn114, la molécula de agua 4 y por la cadena lateral de un resto serina en la secuencia de conexión entre el dominio SH2 y el dominio catalítico que sustituye al resto prolina de p41. En las estructuras de otras quinasas, como es el caso de Src, la posición 5 está ocupada por la cadena lateral de un resto Gln en la secuencia de conexión SH2-quinasa que desplaza a la molécula de agua y establece interacciones con el agua en posición 4 directamente. El análisis de otros complejos SH3 con ligandos peptídicos proporciona situaciones muy variadas. En muchos casos, se observan moléculas de agua en posiciones equivalentes o similares, aunque existe una mayor variabilidad debida a diferencias en las secuencias peptídicas.

**Tabla 4.1.** Interacciones establecidas por las moléculas de agua 4 y 5 en el complejo Abl-SH3/p41 y otras moléculas de agua en los dominios SH3 de varias tirosinas quinazas

Proteína	Abl/p41	Abl/p41	Abl/p41	Abl/p41	Abl quinasa	Src quinasa	Src quinasa	Src quinasa	Src quinasa	Hck quinasa complejo A420983	Hck quinasa complejo A770041	Hck quinasa complejo A641359
Pdb	1bbz (a/b)	1bbz (c/d)	1bbz (e/f)	1bbz (g/h)	1opk	1fmk	2ptk	2src	1y57	2c0i	2c0o	2c0t
Resol (Å)	1.65	1.65	1.65	1.65	2.7	1.5	2.35	1.5	1.91	2.3	2.85	2.15
<b>AGUA POSICIÓN 4</b>												
Agua #	1082	1067	1097	1001	35	1013	13	1046	179	16	25	13
Interac.*	<b>Distancia (Å)</b>											
1	2.64	2.88	2.75	2.90	2.77	2.90	3.91	3.08	2.97	2.73	2.95	2.88
2	2.88	3.18	3.19	3.24	3.46	3.87	3.18	4.07	3.76	3.46	3.10	3.41
3	3.12	-	2.78	3.53	3.33	3.23	3.16	3.16	3.00	3.06	3.35	2.54
4	3.02	3.06	3.07	2.90	3.09	4.46	3.61	4.7	4.38	3.21	2.80	3.11
5	4.48	4.72	4.79	4.60	4.49	4.62	4.36	4.74	4.61	3.44	2.72	3.13
6	3.65	3.38	3.16	3.51	3.59	-	-	-	-	-	-	-
7	-	-	-	-	-	2.87	-	3.06	2.88 (3.03)	-	-	-

\* 1. Interacción con oxígeno de esqueleto peptídico de Asp117 en quinazas de Src, Glu117 en quinasa de Abl, y Glu98 de Abl-SH3.

2. Interacción con nitrógeno de esqueleto peptídico de Ser en hélice 3<sub>10</sub>.

3. Interacción con oxígeno delta de Asn en hélice 3<sub>10</sub>.

4. Interacción con nitrógeno de esqueleto peptídico de Asn en hélice 3<sub>10</sub>.

5. Interacción con oxígeno gamma de Ser en hélice 3<sub>10</sub>.

6. Interacción con agua en posición 5 para Abl/p41 y la quinasa de Abl (agua 44 en el Pdb: 1opk).

7. Interacción con nitrógeno de esqueleto peptídico de Gln253 en el lazo que conecta las quinazas de Src al dominio SH2. El valor entre paréntesis es para la interacción con oxígeno epsilon de Gln253.

Los valores que no aparecen en negrita corresponden a las distancias entre los átomos potenciales formadores del puente de hidrógeno que no llega a establecerse ya que superan el valor de 3.6 Å. Para mayor claridad en la Figura 4.1 se muestran solo las interacciones que se dan en todos los dominios SH3, interacciones 1-5.

Proteína	Abl/p41	Abl quinasa	Src quinasa	Src quinasa
Pdb	1bbz(a/b)	1opk	1fmk	2ptk
Resol. (Å)	1.65	2.7	1.5	2.35
<b>AGUA POSICIÓN 5</b>				
Agua #	1105	44		
Resto:átomo (Distancia Å)	Asn154:OD (2.64) Asn154:ND (2.70) P7:O Ligando (2.49)  Agua 1082 (agua posic. 4) (3.65)	Asn133:OD (2.72) Asn133:ND (3.54) T243:O lazo conector quinasa-SH2 (2.67) Agua 83 (2.78)		
<b>OTRAS MOLÉCULAS DE AGUA</b>				
Agua #			1046	69
Resto:átomo (Distancia Å)			Agua 1013 (agua posic. 4) (3.58)	Q253:N lazo conector quinasa-SH2 (2.79)  Agua 13 (3.08) N135:ND (3.1) D117:O (3.35)

## **Discusión**

Es interesante señalar que, sobre todo en ligandos en los que los restos Pro que interaccionan con la molécula de agua 5 en el complejo Abl-SH3/p41 están sustituidos por otros aminoácidos, ésta desaparece, de modo que el péptido adopta una conformación ligeramente diferente que permite el establecimiento de una interacción directa entre la cadena lateral del resto equivalente a Asn114 y el ligando.

### **4.1.4.2. Moléculas de agua en el bolsillo de especificidad (lazos n-Src y -RT)**

La región del lazo n-Src es de las más variables en secuencia en los distintos dominios SH3 y se ha identificado, junto con el lazo RT, como responsable de la especificidad de unión en estos sistemas. Las moléculas de agua en esta región del complejo Abl-SH3/p41 están menos localizadas y presentan, en general, tiempos de residencia bajos, lo que está de acuerdo con la elevada flexibilidad conformacional de los aminoácidos que las coordinan. Además, las posiciones de hidratación en esta zona están menos definidas, de modo que además de las posiciones 1 y 2, predominantes, se encuentran moléculas de agua que ocupan otras posiciones de forma transitoria.

El análisis de las distintas estructuras muestra una elevada variabilidad tanto en el número como en la configuración de las moléculas de agua encontradas en esta zona de especificidad, en muchos casos asociada a diferencias en la conformación que adopta el lazo n-Src en el complejo. En este sentido, recientemente se ha resuelto la estructura en disolución del complejo de p41 con el dominio SH3 de espectrina en el que se observa que, en efecto, el lazo n-Src presenta una conformación más abierta que implicaría la pérdida de las interacciones con las moléculas de agua en esta región del complejo con Abl-SH3 (Casares *et al.* 2007a).

Es interesante señalar que en muchos casos, como por ejemplo los complejos 2bz8, 1uti, 1ssh o 1uzk) se observan moléculas de agua enterradas en la interfase de unión también en la región del lazo RT, generalmente asociadas a la presencia de restos Arg o Lys en la zona de especificidad del ligando. Estas moléculas de agua no se observan en ninguno de los complejos de Abl y los ligandos de este

dominio se caracterizan precisamente por la ausencia de restos básicos en sus secuencias, muy comunes e importantes para la interacción en otros complejos.

Todas estas observaciones demuestran que el doble mecanismo propuesto en esta Memoria para el complejo Abl-SH3/p41 consistente en la combinación de las interacciones hidrofóbicas con enlaces de hidrógeno mediados por moléculas de agua es una característica universal del reconocimiento de secuencias ricas en prolina por dominios SH3, de modo que, en todas y cada una de las estructuras cristalográficas de resolución mayor de 2 Å analizadas, se observan moléculas de agua ocluidas en la interfase de unión. Es más, como hemos visto, la disposición de estas moléculas de agua y de las interacciones que establecen varía según la secuencia del ligando y la del propio dominio, lo que indica claramente que desempeñan un papel importante en la especificidad de unión.

#### **4.1.5. La incorporación de las interacciones mediadas por moléculas de agua a estrategias de diseño racional permitiría optimizar la afinidad y especificidad de unión en los complejos SH3**

Hasta hace algunos años, la práctica más común en los estudios computacionales sobre reconocimiento molecular en proteínas era ignorar por completo las moléculas de agua en las interfases de unión. No obstante, aunque tradicionalmente el proceso de diseño se ha centrado fundamentalmente en la optimización de las interacciones directas entre los ligandos y el sitio de unión, hoy en día está bien establecido que la eficiencia de las estrategias de diseño racional puede incrementarse notablemente si se consideran los efectos relacionados con el disolvente, más allá del consabido efecto hidrofóbico; es decir, si se posee una comprensión adecuada del papel y la importancia de las distintas moléculas de agua en el sitio de unión (Ben-Naim 2002; Wang *et al.* 1996a). En este sentido, recientemente se han publicado algunos trabajos que describen procedimientos de docking (Minke *et al.* 1999) (Rarey *et al.* 1999; Verdonk *et al.* 1999) y diseño racional (García-Sosa *et al.* 2005; García-Sosa & Mancera 2006; Jiang *et al.* 2005; Mancera 2002) que permiten incluir en los cálculos las moléculas de agua

## **Discusión**

interfaciales y que demuestran que los resultados obtenidos son mucho más precisos cuando éstas se consideran.

Las moléculas de agua pueden participar en las interacciones entre proteína y ligando de muy diversas formas: pueden estar unidas fundamentalmente a la proteína, al ligando o a ambos por igual; pueden encontrarse en el sitio de unión o bien en su periferia; pueden estar firmemente unidas a la proteína libre o incorporarse a la interfase de unión sólo cuando el ligando está presente; pueden mediar interacciones directamente entre el ligando y la proteína o formar parte de una red de moléculas de agua contribuyendo a la estabilización de otras moléculas de agua. La estrategia de diseño que nos permita obtener la mayor ventaja energética de cada uno de estos tipos de moléculas de agua será diferente.

En general, son dos los planteamientos más habituales en diseño racional con respecto a las moléculas de agua: a) diseño de un ligando que optimice sus interacciones con moléculas de agua en la interfase de unión y b) diseño de ligandos que sustituyan o desplacen moléculas de agua en la interfase de unión presentes en la proteína libre o en complejos con ligandos relacionados. En cualquiera de las dos situaciones es muy importante seleccionar adecuadamente las moléculas de agua diana. En el primer caso, sólo se obtendría un beneficio energético significativo si las interacciones adicionales establecidas entre el ligando y la molécula de agua superan cualquier penalización entrópica asociada a la restricción adicional de los grados de libertad configuracionales de ésta. Esto se conseguiría dirigiendo el diseño del ligando hacia la interacción con moléculas de agua firmemente unidas a la proteína, preferentemente en zonas de alta estabilidad estructural (Lafont *et al.* 2007). Las moléculas de agua en posición 4 constituyen un claro ejemplo de la diana ideal para este tipo de diseño. En el segundo caso, el planteamiento debe ser el opuesto y es muy importante seleccionar que moléculas de agua es posible desplazar del sitio de unión. Discriminar qué moléculas de agua sería beneficioso mantener como mediadores de la interacción y cuáles puede ser ventajoso desplazar del sitio de unión a partir únicamente de la estructura cristalográfica no es sencillo. En este sentido, durante los últimos años se han desarrollado varios algoritmos para la clasificación de las moléculas de agua en la interfase de unión (Barillari *et al.* 2007; Poornima & Dean 1995a; Poornima & Dean 1995b; Teyra *et al.*

2006) y para la predicción de la facilidad con cada una de ellas sería desplazadas por grupos polares en el ligando, generalmente basados en estudios geométricos de las superficies moleculares, de los enlaces de hidrógeno establecidos, de la movilidad de las moléculas (factores B cristalográficos) y de la proximidad de restos polares, entre los que cabe destacar Consolv (Raymer *et al.* 1997), Waterscore (García-Sosa *et al.* 2003) o HINT/RANK (Chen & Kellogg 2005; Kellogg *et al.* 1991). Estos estudios indican, en general, que la energía libre de cada molécula de agua depende de la naturaleza del ambiente en que está localizada. Así, las moléculas de agua que están fuertemente unidas se encuentran generalmente en cavidades polares y establecen al menos tres enlaces de hidrógeno con la proteína y el ligando. Por otro lado, las moléculas de agua unidas más débilmente generalmente establecen menos de tres enlaces de hidrógeno y se encuentran situadas en ambientes parcialmente apolares. No obstante la capacidad predictiva de estas metodologías es aún bastante limitada.

Por el contrario, aunque obviamente requiere una inversión de trabajo muy superior al mero análisis estructural, el tipo de estudio presentado en esta Memoria en el que se combinan el análisis termodinámico, estructural y dinámico del sistema nos ha proporcionado una descripción detallada de las características y propiedades de las distintas moléculas de agua en la interfase de unión del complejo Abl-SH3/p41 que permiten la racionalización de las estrategias de diseño en este sistema. Así, nuestros resultados muestran claramente que la molécula de agua 4 es parte integral del dominio, ya que se encuentra interaccionando fuertemente con las zonas más estables del dominio, independientemente de la presencia del ligando. Como consecuencia, puede considerarse como una extensión del sitio de unión con el que el ligando podría establecer interacciones entálpicamente favorables con poca penalización entrópica. El papel de las moléculas de agua en posición 5, que interaccionan predominantemente con el ligando, es muy diferente ya que su presencia está condicionada a la formación del complejo y tienen mayor flexibilidad conformacional. En este sentido, la introducción de modificaciones en el ligando que las desplacen del sitio de unión y que establezcan interacciones directas con las moléculas de agua en posición 4 podría ser favorable energéticamente y conducir a una mayor afinidad de unión. De hecho, como hemos visto en el apartado anterior, esta situación se observa en algunos ligandos naturales, como la quinasa de Src, en la que una cadena de Gln en el ligando natural sustituye a la molécula de agua.

## **Discusión**

No obstante, es importante seleccionar cuidadosamente las modificaciones a introducir en el ligando de modo que la ventaja entrópica que supone la liberación de la molécula de agua no se vea completamente anulada por otros efectos. Así, conviene tener en cuenta que el enlace de hidrógeno entre la molécula de agua en posición 5 y los oxígenos carbonilos de las prolinas 6 y 7 del ligando están posiblemente muy optimizados entálpicamente debido tanto a las características electrónicas específicas de éstos que los hacen muy buenos grupos aceptores como a las mayores posibilidades conformacionales de la molécula de agua que permiten posiblemente la adopción de distancias y ángulos de enlace óptimos, hecho que sin duda se observa en la estructura y en las dinámicas. Por tanto, es importante que los nuevos grupos en el ligando estén diseñados para establecer enlaces de hidrógeno óptimos. Por otro lado, hay que considerar que la sustitución del resto prolina7 por otra cadena lateral de mayor longitud, como una Gln, llevará asociado una penalización entrópica adicional originada por la pérdida de grados de libertad de ésta y por la posible desestabilización de la conformación PPII en el ligando libre. En este sentido, la utilización de estructuras proteicas como el neuropéptido aviar en el que la hélice PPII está estabilizada por interacciones intramoleculares con otros elementos de la molécula puede resultar interesante para minimizar estos efectos. Estas estrategias serían aplicables para aumentar la afinidad de unión no sólo en el dominio Abl-SH3 sino en otros dominios SH3.

La presencia de moléculas de agua en la interfase de unión de dominios SH3 abre también nuevas posibilidades de diseño en cuanto a la optimización de la especificidad de unión. Las interacciones de las moléculas de agua en la región del lazo n-Src están mucho menos conservadas entre los distintos complejos SH3. De hecho, recientemente se ha propuesto que las diferencias en la termodinámica de unión del ligando p41 a los dominios SH3 de Abl y espectrina (Spc) pueden estar relacionados con diferencias en la secuencia en el lazo n-Src, de modo que las tres moléculas de agua que se encuentran enterradas en la estructura cristalográfica de Abl-SH3/p41 no lo estarían en el complejo Spc-SH3/p41 según la estructura de este complejo en disolución (Casares *et al.* 2007a). El análisis de las estructuras cristalográficas de complejos de otros dominios SH3 revela la presencia sistemática de moléculas de agua en la zona de especificidad aunque con configuraciones muy variables, de modo que, en algunos casos éstas se encuentran mediando las

interacciones entre el lazo n-Src y el ligando, en otros entre el lazo -RT y en otros en ambos. Es muy posible que estas interacciones, que no se han considerado hasta el momento en ninguna descripción de las interacciones de dominios SH3, sean en parte responsables de la especificidad de unión en estos sistemas. De este modo, aunque la información que tenemos hasta el momento sobre las moléculas de agua en estas regiones, especialmente sobre su contribución energética a la interacción, es más limitada, su presencia abre un enorme campo de posibilidades de diseño para la optimización de los ligandos en cuanto a especificidad de unión.

**4.1.6. Efectos inesperados en la energética de unión inducidos por modificaciones en la secuencia del ligando ponen de manifiesto la elevada complejidad del reconocimiento de secuencias ricas en prolina por dominios SH3**

El estudio de las interacciones del ligando p41 con el dominio Abl-SH3 presentado en esta Memoria ha puesto de relevancia que el reconocimiento de secuencias ricas en prolina por parte de dominios SH3 no se corresponde con la visión simplista de una interacción basada fundamentalmente en la interacción directa entre interfases hidrofóbicas. Por el contrario, la influencia de factores como los equilibrios conformacionales del ligando y del propio dominio así como las interacciones mediadas por moléculas de agua confieren un mayor nivel de complejidad. En este contexto, los resultados obtenidos del análisis de la energética de unión de un conjunto de péptidos derivados de p41 resultan un ejemplo claro de que la descripción de estas interacciones de acuerdo al mecanismo de interacción tradicionalmente aceptado, basado en exclusivamente en la intercalación de restos prolina en los bolsillos hidrofóbicos del dominio, resulta claramente insuficiente para proporcionar una base adecuada al diseño racional.

El ligando p41 se desarrolló mediante un proceso de diseño racional a partir del ligando natural de Abl-SH3 denominado 3BP1 (APTMPPPLPP) con el propósito doble de aumentar su afinidad de unión a Abl-SH3 a la vez que optimizar la especificidad de la interacción con respecto al dominio SH3 de Fyn (Pisabarro & Serrano 1996). Los autores observaron que en la estructura del complejo Abl-

## Discusión

SH3/3BP1 existen algunos restos diferentes a prolina en el ligando, como es el caso de la Leu en posición 8, que se encuentran en conformación PPII en el complejo y que además están dirigidos hacia el disolvente, de modo que sus cadenas laterales no establecen ningún tipo de interacciones con el dominio. De acuerdo con la descripción tradicional de los complejos SH3 y aplicando principios biofísicos básicos los autores fundamentaron su estrategia de diseño en la idea de que la sustitución de este resto de Leu por Pro debería suponer un incremento significativo de la afinidad de unión, ya que la mayor propensión del aminoácido Pro a la formación de la hélice PPII (Creamer 1998; Creamer & Campbell 2002; Kay *et al.* 2000) desplazaría el equilibrio del péptido libre hacia esta conformación y minimizaría el coste entrópico de la interacción. De hecho, experimentalmente se pudo comprobar que, efectivamente, el péptido modificado presenta un aumento significativo en la señal PPII en el espectro de dicroísmo circular e interacciona con mayor afinidad con el dominio Abl-SH3.

No obstante, el análisis termodinámico de las interacciones de p41 y un conjunto de péptidos relacionados procedentes del mismo proceso de diseño, han puesto de manifiesto que, inesperadamente, la sustitución de restos Leu por Pro en posición 8 resulta en entalpías de unión considerablemente más negativas y en contribuciones entrópicas más desfavorables, lo que indica claramente que el incremento en afinidad de unión observado tiene su origen en otros factores que enmascaran los efectos entrópicos favorables esperados. Es interesante señalar que, no sólo es sorprendente el origen entálpico del aumento en afinidad de unión, sino que los cambios entálpicos y entrópicos asociados a estas sustituciones son muy notables (en algunos casos superan los  $15 \text{ kJ}\cdot\text{mol}^{-1}$ ) sobre todo considerando que se trata de modificaciones de las cadenas laterales que no están en contacto con el dominio. Además, la magnitud de los efectos de la sustitución en posición 8 dependen de la naturaleza del resto en posición 5 y viceversa, lo que indica la existencia de algún tipo de cooperatividad entre ambas posiciones en el ligando. Con la información disponible hoy en día es difícil establecer claramente el origen de estos efectos, aunque nuestra hipótesis es que están relacionados con las características específicas del aminoácido prolina. Así, aunque la cadena lateral del resto en posición 8 en el ligando se encuentran expuesta al disolvente en la estructura cristalográfica del complejo Abl-SH3/p41, su grupo carbonílico está

dirigido hacia el interior del sitio de unión y establece un puente de hidrógeno muy conservado en todos los complejos SH3 con el resto Tyr115 en la hélice 3<sub>10</sub>. De acuerdo con nuestros resultados de dinámica molecular, ésta es la interacción directa más importante entre la proteína y el ligando, ya que presenta ángulos y distancias de enlace muy optimizados y se establece durante más del 60% del tiempo de simulación. Debido a la ausencia del protón amida, el grupo carbonilo del aminoácido Pro posee una elevada densidad electrónica y además, a causa de las restricciones conformacionales impuestas por su cadena lateral cíclica, está pobremente solvatado, por lo que es mejor formador de enlaces de hidrógeno que los carbonilos de otros aminoácidos (Ball *et al.* 2005; Creamer 1998; Creamer & Campbell 2002; Kay *et al.* 2000; Kelly *et al.* 2001). Por tanto, es de esperar que la sustitución de Leu por Pro en posición 8 resulte en un fortalecimiento de este enlace de hidrógeno así como la intensificación de los efectos que la unión del ligando tiene sobre el propio dominio SH3 ya que, como hemos comentado en secciones anteriores, se ha descrito en estos sistemas una propagación eficiente de las interacciones en el sitio de unión a otras regiones de la proteína que se refleja en alteraciones de las distancias de enlace en la red de puentes de hidrógeno a lo largo de todo el dominio (Cordier *et al.* 2000; Wang *et al.* 2001). Es interesante señalar, también, que, como ya hemos comentado en secciones anteriores, cuando las posiciones centrales del motivo PxxP no son prolina, se observa con frecuencia que la molécula de agua 5 no está presente y se establecen interacciones directas entre el resto Asn114 y el ligando. Un ejemplo claro de esta situación, es el complejo del dominio SH3 de MONA/GADS con un ligando peptídico de secuencia GNPPLVPPRLELMRGL, donde los restos Leu y Val subrayados corresponden a las prolinas 7 y 8 de p41. En este complejo, la estructura PPII del péptido está algo distorsionada y permite el establecimiento de un enlace de hidrógeno directo entre el carbonilo del resto en posición 7 y la cadena lateral de Asn114 (a 3.2 Å de distancia) que en Abl-SH3/p41 está a más de 5 Å. Por lo tanto, parece que la presencia de la molécula de agua en posición 5 está asociada a la adopción de una conformación PPII ideal y rígida por parte del péptido. En este sentido, es posible que los efectos termodinámicos asociados a la introducción de Pro en lugar de Leu en posición 8 estén asociados a la entrada de la molécula de agua en posición 5 que conllevaría efectos entálpicos favorables y entrópicos desfavorables que compensarían parcialmente los beneficios entrópicos de la sustitución.

## Discusión

Por otro lado, la sustitución de Ser en posición 5 por Pro resulta, como era de esperar, en un beneficio entrópico asociado a la reducción del espacio conformacional accesible al péptido libre que, sin embargo, está completamente compensado por contribuciones entálpicas desfavorables, por lo que no tiene ningún impacto en la afinidad de unión. La cooperatividad entre ambas posiciones se refleja claramente en que, aunque el efecto entrópico resultante de la introducción de Pro en posición 5 cuando la posición 8 está ocupada por Leu ( $-2.1 \text{ kJ}\cdot\text{mol}^{-1}$ ) es comparable a los valores descritos en la literatura para sustituciones Ala/Pro (Ferreon & Hilser 2003a), el efecto de la sustitución se amplifica notablemente ( $-11.3 \text{ kJ}\cdot\text{mol}^{-1}$ ) cuando tenemos Pro en posición 8. Lo mismo ocurre con las contribuciones entálpicas. Según la estructura cristalográfica del complejo Abl-SH3/p41, la sustitución del resto Ser en posición 5 por Pro resultaría en la pérdida de un enlace de hidrógeno intramolecular establecido entre el grupo hidroxilo de la cadena lateral de S5 y el oxígeno carbonílico del resto Pro en posición 6. Este tipo de enlaces de hidrógenos intramoleculares entre cadenas laterales cortas y el oxígeno carbonílico del resto contiguo son muy comunes (Vijayakumar *et al.* 1999) y se ha postulado que son en parte responsables de la elevada propensidad para formar hélices PPII que presentan este tipo de aminoácidos (Kelly *et al.* 2001). Además, según nuestros resultados de dinámica molecular, la cadena lateral del resto S5 se encuentran implicada en un enlace de hidrógeno con la cadena lateral del resto Glu98 situado en el lazo n-Src del dominio. Esta interacción, que no se observa en la estructura cristalográfica, está presente durante un porcentaje muy considerable del tiempo de simulación. Es muy probable que la eliminación de estas interacciones con la sustitución de Ser por Pro esté relacionada con la pérdida entálpica. Es asimismo interesante señalar que, tanto el carbonilo del resto P6 en el ligando como la cadena lateral de Glu98 están implicadas también en las interacciones mediadas por moléculas de agua, el primero coordinando a las moléculas de agua en posición 5 en la zona de la hélice  $3_{10}$  y el segundo asistiendo a la entrada de moléculas de agua en la región de especificidad.

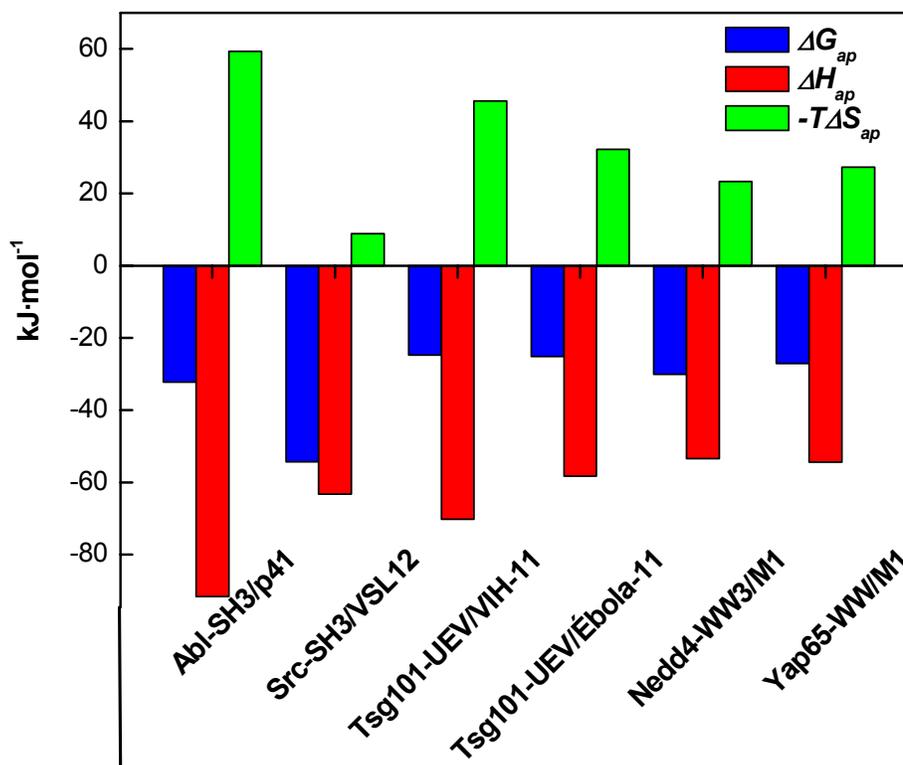
## **4.2. RECONOCIMIENTO DE SECUENCIAS RICAS EN PROLINA POR OTROS DOMINIOS MODULARES**

Ante la complejidad asociada a las interacciones establecidas por los dominios SH3, hemos extendido nuestro estudio a otros dominios modulares que también reconocen secuencias ricas en prolina, para ver si esta complejidad está también presente. Hemos estudiado los dominios WW, que junto con los dominios SH3 representan los dominios modulares más ampliamente representados en el proteoma humano y para los que existen algunos estudios (Macias *et al.* 2002; Mayer 2001; Zarrinpar & Lim 2000); y los dominios UEV de los que hasta la fecha se conoce muy poco. Concretamente, hemos caracterizado la unión de algunos ligandos víricos naturales al dominio WW de Yap65 (Macias *et al.* 1996) y WW3 de Nedd4 (Kanelis *et al.* 2006; Timmins *et al.* 2003; Yasuda *et al.* 2003) que reconocen secuencias con especificidad tipo PPxY (Otte *et al.* 2003); y al dominio UEV de Tsg101 (Demirov *et al.* 2002; O. Pornillos *et al.* 2002), que reconoce secuencias tipo PTAP. En el caso de los dominios WW hemos utilizado además el ligando sintético denominado como **M1** (EYPPYP**PPY**PSG) (Pires *et al.* 2001).

### **4.2.1. El reconocimiento de ligandos ricos en prolina por parte de dominios WW y UEV presenta el mismo patrón que los dominios SH3, incongruente con la elevada hidrofobicidad de las interacciones**

Al igual que ocurre con dominios SH3 la unión de los ligandos a estos dominios WW y UEV muestra un patrón termodinámico que es incongruente con la naturaleza hidrofóbica de la interfase de unión de los complejos y es difícil de explicar teniendo en cuenta la ausencia o escasez de interacciones polares. Este patrón termodinámico coincide con el observado en dos experimentos puntuales de calorimetría isotérmica de titulación llevados a cabo muy recientemente con dos dominios WW diferentes, en los que, aunque no se dan los valores de las contribuciones entálpicas y entrópicas, de acuerdo con las titulaciones calorimétricas mostradas puede observarse que la interacción presenta contribuciones entálpicas favorables (Meiyappan *et al.* 2007; Morales *et al.* 2007).

## Discusión



**Figura 4.2.** Representación en forma de barras de la energía libre de Gibbs, y sus contribuciones entálpica y entrópica para la interacción de dominios modulares SH3, UEV y WW con ligandos conteniendo secuencias ricas en prolina. Los experimentos con los dominios SH3 y WW se llevaron a cabo en fosfato sódico 20 mM pH 7.0, y con Tsg101-UEV en fosfato sódico 20 mM, BME 5 mM, pH 7.2, todos a 25 °C.

En la Figura 4.2, se muestran los parámetros termodinámicos obtenidos para la unión de algunos de los ligandos ricos en prolina a los dominios WW y UEV estudiados en esta Memoria y a los dominios SH3 de Abl y Src. Los valores de entalpía obtenidos para la unión de los ligandos a los dominios WW estudiados van desde los -40 hasta -70 kJ·mol<sup>-1</sup>, y en el caso del dominio UEV desde -55 hasta -72 kJ·mol<sup>-1</sup>, rangos que son muy similares a los mostrados por los dominios SH3 estudiados hasta la fecha (-35 a -92 kJ·mol<sup>-1</sup>) (Arold *et al.* 1998; Ferreón & Hilser 2004; Palencia *et al.* 2004; Renzoni *et al.* 1996). Del mismo modo, las contribuciones entrópicas y las constantes de asociación oscilan en rangos del mismo orden.

En el caso de los dominios SH3, hemos comprobado que algunos factores como la presencia de moléculas de agua en la interfase de unión de los complejos tienen un papel importante para la interacción (Palencia *et al.* 2004). Además, otros

factores como el equilibrio conformacional del ligando (Ferreon & Hilser 2004) y de la proteína (Wang *et al.* 2001) también parecen ser importantes en las interacciones establecidas por estos dominios. Aunque en el caso de los dominios WW y UEV no existen trabajos en solución ni apenas estructuras cristalográficas de complejos donde se haya manifestado la importancia de estos factores, si tenemos en cuenta la similitud estructural y el modo en que reconocen secuencias ricas en prolina, caracterizado por la inserción de prolinas en bolsillos hidrofóbicos y especificidad determinada por lazos flexibles variables en cada dominio (Ball *et al.* 2005), y el hecho de que comparten el mismo patrón termodinámico que los dominios SH3, podría pensarse que algunos de estos factores pueden estar presentes para estos dominios.

En el caso de los dominios WW, un trabajo muy reciente con el dominio WW de FE65 (Meiyappan *et al.* 2007), que ha permitido obtener la estructura cristalográfica del complejo de esta proteína con el ligando rico en prolina Mena9 (**PPPPPPLPP**), ha revelado la presencia de dos molécula de agua (aguas 11 y 310 del fichero Pdb: 2oei) en la interfase de unión que se encuentran mediando interacciones entre la proteína y ligando al igual que ocurre con Abl-SH3 y el ligando p41. Concretamente, la molécula de agua 11 establece un puente de hidrógeno con el resto Thr37 en la proteína y otro con la Pro3 del ligando; y la molécula de agua 310 establece un puente de hidrógeno con Tyr28 en la proteína y otro con la Pro2 del ligando. Por tanto, las interacciones establecidas por estas moléculas de agua afectarán favorablemente al valor de entalpía de unión, hecho que puede ayudar a explicar el patrón termodinámico inesperado que se obtiene en dicho trabajo (Meiyappan *et al.* 2007). Al igual que en el caso de las moléculas de agua 4 y 5 en el complejo Abl-SH3/p41 el valor del factor B de la molécula de agua 11 en este nuevo complejo es parecido al de los restos con los que interacciona, P3 y Thr37, y menor al observado para otras moléculas de agua en la estructura, lo que, como hemos visto en secciones anteriores, es característico de aguas importantes para la estructura y función de las proteínas (Damjanovic *et al.* 2005; Park & Saven 2005; Petrone & Garcia 2004). Conviene señalar, que esta molécula de agua interacciona con el hidroxilo del resto en la posición 37 del dominio que, bien por Ser o Thr, está muy conservado entre los diferentes tipos de dominios WW (90% del total que se conocen) (Morales *et al.* 2007), incluidos los dominios Yap65-WW y Nedd4-WW3

## **Discusión**

estudiados en esta Memoria, y según los estudios realizados con otros complejos con dominios WW tipo I, de los cuales se dispone de información estructural en disolución (Huang *et al.* 2000; Kanelis *et al.* 2006), tiene un papel fundamental para la interacción con los ligandos. De hecho, en un trabajo reciente llevado a cabo en el grupo de la Dra. Macias (Instituto de Investigación en Biomedicina, Parc Cientific de Barcelona, España) han observado que la fosforilación de la Thr37 en el dominio Itch-WW3 inhibe la interacción con un péptido tipo PPxY (Morales *et al.* 2007). Por tanto, aunque apenas se dispone de estructuras cristalográficas que permitan obtener información estructural del disolvente, puede que existan más moléculas de aguas en otros dominios WW que ocupen una posición similar a ésta molécula de agua en el complejo FE65-WW/MENA y que sean importantes para mediar interacciones indirectas entre los ligandos y el resto Thr/Ser en posición 37.

En el caso del dominio UEV, aunque todavía no disponemos de estructuras cristalográficas de complejos que nos permitan evaluar la importancia del disolvente para la unión a ligandos ricos en Pro, la obtención de la estructura cristalográfica de la proteína libre (Palencia *et al.* 2006) ha revelado la presencia de dos moléculas de agua en el sitio de unión que podrían ser importantes para la interacción con el ligando vírico **VIH-9** (PE**PTAP**PEE). Una de estas moléculas de agua (agua 195 en el fichero Pdb: 2f0r) ocupa una posición ideal para el establecimiento de una interacción indirecta con este ligando (Figura 3.43 página 225), del mismo modo que hemos observado para el complejo Abl-SH3/p41, lo que contribuiría favorablemente al valor de de entalpía de unión obtenido.

La obtención de nuevas estructuras cristalográficas así como estudios computacionales (p.ej., simulaciones de dinámica molecular) de nuevos complejos con algunos ligandos, tareas en la que nuestro grupo ya trabaja actualmente, nos permitirán estudiar más detalladamente la importancia del disolvente para las interacciones establecidas por estos dominios modulares. Igualmente, estudios de tipo estructural en disolución permitirán analizar la importancia de otros factores como el equilibrio conformacional de la proteína o el ligando. Una comprensión detallada de la influencia de estos factores para la unión a los ligandos será de gran ayuda para abordar el diseño de nuevas moléculas que inhiban las interacciones establecidas por estos dominios modulares que reconocen poliprolina, que, aunque

todavía pueda ser muy difícil afirmarlo, de acuerdo con los resultados obtenidos en esta Memoria podrían presentar cierta universalidad.

#### **4.2.2. Reconocimiento de dominios L víricos tipo PPxY por dominios WW**

Tanto los experimentos de fluorescencia como de ITC muestran que la afinidad del ligando **M1** por ambos dominios WW es mucho mayor que la mostrada por el resto de ligando naturales. Los experimentos de calorimetría isotérmica de titulación indican que el aumento de afinidad obtenido para este ligando proviene de un coste entrópico menos desfavorable, probablemente debido a que el ligando **M1** posee un mayor número de prolinas (EYPPY**PPY**PSG) lo que favorece el desplazamiento hacia la conformación PPII necesaria para la unión al dominio WW. Del mismo modo, el incremento de afinidad obtenido para Nedd4-WW3 respecto a Yap65-WW, también proviene de una contribución entrópica menos desfavorable. Si comparamos el sitio de unión para ambos dominios WW, encontramos que el de Nedd4-WW3 posee mayor apolaridad que el de Yap65-WW lo que podría favorecer la inserción de las prolinas del ligando y tener una mayor desolvatación de las superficies de interacción de proteína y ligando cuando se unen. Concretamente, los restos del sitio de unión en posiciones 22 y 28, que son muy importantes para la interacción con los ligando (Kanelis *et al.* 2006; Pires *et al.* 2001), son Phe e Ile en Nedd4-WW3, mientras que en Yap65-WW son Thr y Tyr. En este sentido, los valores del cambio de capacidad calorífica que hemos obtenido para la unión de **M1** a Nedd4 y Yap65,  $-1.6$  y  $-0.8$   $\text{kJ}\cdot(\text{K}\cdot\text{mol}^{-1})$  respectivamente, indican que en el caso de Nedd4-WW3 el cambio en la superficie de hidratación podría ser mucho mayor que para Yap65-WW. Este resultado concuerda con los estudios realizados con el dominio homólogo de WW3 de Nedd4 en *Drosófila*, muy similar al de humanos estudiado en esta Memoria, y un péptido rico en prolina de la proteína *Comm*, donde se obtiene un cambio de área accesible al disolvente tras la unión del ligando que es considerablemente mayor ( $895 \text{ \AA}^2$ ) que el de otros dominios WW (Kanelis *et al.* 2006) como Yap65, para el que se obtienen valores desde  $620$  hasta  $730 \text{ \AA}^2$  para un amplio juego de ligandos ricos en prolina (Macias *et al.* 1996; Pires *et al.* 2001).

## Discusión

A pesar de estas diferencias, el hecho de que el ligando **M1** sea el que posee mayor afinidad nos indica que la estrategia de diseño empleada para Yap65-WW, para el que se sintetizó originalmente el ligando **M1** (Macias *et al.* 1996; Pires *et al.* 2001), podría ser empleada del mismo modo en el caso de Nedd4-WW3.

En cuanto al estudio llevado a cabo con los ligandos víricos naturales, hemos obtenido una afinidad baja en general, aunque ligeramente mayores en el caso de Nedd4-WW3. El ligando del virus de leucemia humana, denominado **Leucemia-b** (SDPQI**PPY**VEP), es el que presenta mas afinidad después del ligando **M1**, lo que coincide para Nedd4-WW3 y Yap65-WW. Curiosamente, el valor de entalpía de unión al dominio Nedd4-WW3 que se obtiene para la unión del ligando **Leucemia-b** y del otro ligando de leucemia humana, denominado **Leucemia-a** (**PPY**VE**PTAP**), es el mayor de todos y es independiente de la posición que ocupe la secuencia PPPY en el ligando (Figura 4.2). La presencia de algunos restos en el extremo C-terminal contiguos a la secuencia PPPY podría ser la causa del valor de entalpía de unión tan elevado que se obtiene. Concretamente, la presencia del resto glutámico (**E**) en los ligandos de leucemia humana, cargado negativamente, podría establecer alguna interacción electrostática que contribuiría favorablemente a la entalpía. Aunque todavía no disponemos de información estructural que permita analizar las interacciones que establecería este resto, la estructura del complejo de Nedd4-WW3 en *Drosófila* con el péptido *Comm* (TGLPSYDNEALH) (Kanelis *et al.* 2006) presenta una interacción de tipo electrostático entre Arg35 y el glutámico (**E**) del ligando. En el caso del dominio Nedd4-WW3 humano estudiado en esta Memoria la Arg está conservada en esta posición, por lo que es probable que se establezca la misma interacción con el glutámico equivalente en los ligandos de leucemia humana. Por el contrario, en Yap65-WW, la ausencia de un resto cargado negativamente en tal posición, en la que hay una glutamina, impediría su formación.

De acuerdo con estos resultados, podemos decir que las interacciones establecidas por restos contiguos al motivo PPxY en el ligando vírico **Leucemia-b** son importantes para la unión al dominio WW3 de Nedd4, hecho que no es muy común para otros dominios WW tipo I (Huang *et al.* 2000; Macias *et al.* 1996; Pires *et al.* 2001), donde la interacción con los ligandos parece fundamentarse en las interacciones con la secuencia PPxY.

En cuanto a los otros ligandos naturales estudiados en esta Memoria, **Ébola-11** y **PPPY**, hay que decir que la afinidad es muy similar para ambos dominios WW, y que ésta viene determinada fundamentalmente por la secuencia PPPY, ya que la presencia de restos adyacentes no influye significativamente en la afinidad de unión. Por último, el dominio vírico L tipo PTAP que contienen el ligando **Leucemia-a** (**PPPYVEPTAP**) y **Ebola-11** (**ILPTAPPEYME**) tampoco parece contribuir a la afinidad por ninguno de estos dominios WW.

En definitiva, podemos decir que las interacciones establecidas por estos dominios WW presentan un elevado grado de especificidad, ya que pequeñas diferencias en las secuencias de los ligandos provoca diferencias significativas en la afinidad de unión, y también en los orígenes de las fuerzas que contribuyen a ésta. Además, el hecho de que la interacción con el ligando **Leucemia-b** (**SDPQIPPPYVEP**) esté más optimizada con Nedd4-WW3 que con Yap65-WW podría suponer un punto de partida para el diseño de ligandos de alta especificidad dentro de esta clase de dominios WW.

#### 4.2.3. Reconocimiento de dominios L víricos tipo PTAP por el dominio UEV de Tsg101

Aunque el patrón termodinámico obtenido es similar para todos los ligandos víricos que hemos estudiado, existen diferencias significativas en cuanto a afinidad por el dominio UEV de Tsg101, así como en los orígenes de las fuerzas que contribuyen a ésta. Cabe señalar que mientras que los ligandos de los virus de ébola y VIH, tanto en su versión corta de 9 aminoácidos como la de mayor longitud de 11, poseen afinidades del mismo orden ( $K_a$  en torno a  $10^4 \text{ M}^{-1}$ ), el ligando de leucemia humana tipo I, denominado **Leucemia-c** (**YVEPTAPQVL**), cuya secuencia es muy parecida a la de los anteriores, posee una constante de asociación con un orden de magnitud inferior a la de los otros ligandos víricos, lo que nos indica la elevada especificidad mostrada en el reconocimiento de estas secuencias.

Aunque la afinidad mostrada por los ligandos de ébola y VIH es muy parecida, observamos que los ligandos de VIH presentan una entalpía de unión que es más favorable ( $13.4 \text{ kJ}\cdot\text{mol}^{-1}$  más negativa), lo que podría estar relacionado con las

## Discusión

interacciones establecidas por determinados restos acídicos presentes sólo en el caso de los ligandos de VIH. Como se ha observado en los experimentos de titulación de RMN del ligando **VIH-9** ( $^5\text{-PEPTAPPEE}^{-13}$ ) con Tsg101-UEV, los restos acídicos **E** que flanquean la secuencia consenso PTAP en posiciones 6 y 13 son muy importantes para la interacción con el dominio UEV (O. Pornillos *et al.* 2002). Concretamente, el resto **E6** establece un puente de hidrógeno con la cadena lateral de Asn69 del dominio Tsg101-UEV que parece estar muy optimizado, tal y como indica su corta longitud de enlace (2.53 Å), Figura 1.10, página 36, y es de esperar que tenga un efecto favorable sobre las contribuciones entálpicas. El resto **E13** parece ser importante para la orientación adecuada del péptido frente al dominio UEV mediante una interacción de tipo electrostático con el resto Arg144. Por el contrario, los ligandos del virus del ébola no poseen **E** en estas posiciones, ya que en la posición 6 hay **L** y en la 13 hay **Y**, los cuales no pueden establecer las mismas interacciones se observan con el ligando **VIH-9**, lo que concuerda con que su valor de entalpía sea menos favorable.

Por otra parte, el ligando **Ébola-11** contiene además un dominio L tipo PPxY (IL**PTAP****PEY**ME) que no parece ser importante para la interacción con Tsg101-UEV ya que su afinidad no difiere significativamente con respecto al ligando **VIH-11** que no posee este dominio L (PE**PTAP**PEESF). Este resultado está de acuerdo con estudios anteriores que indican que mutaciones a alanina en el motivo **PPEY**, y su delección en la secuencia del dominio L de ébola ( $^7\text{-PTAP**PEY}^{-13}**$ ) no influyen en la interacción con Tsg101-UEV y originan una liberación de partículas virales similar o muy ligeramente inferior a la que ocurre normalmente (Irie *et al.* 2004), mientras que la delección simultánea de las secuencias PTAP y PPxY impide totalmente la liberación de partículas virales de la célula infectada (Licata *et al.* 2003). Estos trabajos indican que el dominio L del virus del ébola podría utilizar tanto la secuencia PTAP como PPxY para la liberación de partículas virales de la célula infectada. En el primer caso utilizaría las interacciones con el dominio UEV de la proteína celular Tsg101, y en el segundo con el dominio WW3 de la ubiquitina ligasa humana Nedd4. Este hecho supone una diferencia importante con respecto a otros virus que también contienen ambos dominios L, como por ejemplo el de Mason-Pfizer, en el cual ambos son necesarios para la correcta liberación de las partículas virales, tal y como indican bien las mutaciones en la secuencia PTAP, que impiden en gran medida la

liberación del virus, o bien las mutaciones en la secuencia PPxY, que impiden totalmente la liberación de partículas virales (Gottwein *et al.* 2003).

Los resultados que hemos obtenidos para los ligandos **Leucemia-a** (PPPYVEPTAP) y **Leucemia-c** (YVEPTAPQVL) indican que el ligando **Leucemia-a**, que contiene el motivo PPPY, presenta una afinidad muy baja por Tsg101-UEV y difícil de determinar con fiabilidad mediante titulaciones calorimétricas, lo que nuevamente indica que la secuencia PPPY es independiente para la interacción Tsg101-UEV/PTAP al igual que se ha observado en estudios previos realizados con el dominio L de este virus (Bouamr *et al.* 2003; Sakurai *et al.* 2004). El hecho de que el ligando **Leucemia-c**, que no contiene la secuencia PPPY y posee la secuencia PTAP en posición diferente de la secuencia del ligando, presente mayor afinidad por Tsg101-UEV que el ligando **Leucemia-a** indica que la presencia de restos adyacentes a PTAP es importante para la interacción con Tsg101-UEV. Finalmente, y dado que en este virus la presencia de PPPY es suficiente para la liberación de las partículas virales de la célula infectada (Sakurai *et al.* 2004), no parece estar clara la función desempeñada por la interacción entre Tsg101-UEV y PTAP.

En definitiva, podemos decir que los virus que contienen estos dominios tipo L han evolucionado de un modo divergente, que les ha permitido desarrollar la capacidad de escapar de la célula infectada bien mediante la interacción entre PPxY y el dominio WW3 de Nedd4 o mediante la interacción entre PTAP y el dominio UEV de Tsg101, o bien mediante las dos. Así, mientras que para los virus de leucemia humana tipo I y VIH la interacción con PPxY y PTAP, respectivamente, es suficiente por sí sola para la liberación de sus partículas virales de la célula infectada, otros virus como el de Mason-Pfizer necesitan establecer interacciones tanto con PPxY como con PTAP. Curiosamente, el virus del ébola ha desarrollado la habilidad de usar las interacciones bien con PTAP o bien con PPxY para el mismo propósito.



## **RESUMEN Y CONCLUSIONES**



## **RESUMEN Y CONCLUSIONES**

En esta Memoria se describe el estudio de las interacciones establecidas por dominios modulares que reconocen secuencias ricas en prolina (MRP), como son el SH3, WW y UEV. El estudio con dominios SH3 ha consistido en una caracterización en profundidad de las interacciones que establece el dominio Abl-SH3 con el ligando rico en prolina p41 y otros derivados de ambos, utilizando tanto técnicas experimentales (calorimetría de titulación, espectroscopías de dicroísmo circular, de fluorescencia y difracción de rayos X principalmente) como computacionales (simulaciones de dinámica molecular). La caracterización termodinámica completa de la interacción ha revelado un patrón termodinámico que es incongruente con la naturaleza hidrofóbica de la interfase de unión de los complejos y que no puede ser justificando atendiendo exclusivamente a información de tipo estructural. La integración de la información obtenida mediante las diferentes técnicas indica que la presencia de moléculas de agua en la interfase de unión del complejo Abl-SH3/p41 es la responsable del patrón termodinámico observado. La caracterización termodinámica de las interacciones establecidas por los dominios modulares WW de Yap65 y Nedd4, y por el UEV de Tsg101, con ligandos ricos en prolina que contienen los dominios L de algunos virus como el VIH, ébola y la leucemia humana, además del ligando M1, ha revelado un patrón termodinámico similar al mostrado por los dominios SH3. Este hecho indica que algunos de los factores presentes en las interacciones establecidas por estos últimos podrían ser también importantes para las interacciones establecidas por dominios UEV y WW.

Como colofón del trabajo descrito en esta Memoria se proponen las siguientes conclusiones:

1. El patrón termodinámico mostrado por la unión del dominio SH3 de Abl con el ligando rico en prolina p41 está gobernado por una entalpía de unión favorable a la que se opone parcialmente una contribución entrópica desfavorable, lo que es incongruente con la naturaleza hidrofóbica de la interfase de unión del complejo. Sin embargo, coincide con lo obtenido para otros dominios SH3. Aún así, la interacción Abl-SH3/p41 posee un carácter mucho más exotérmico, ya que su valor,  $-92 \text{ kJ}\cdot\text{mol}^{-1}$ , es el mayor obtenido hasta ahora para dominios SH3.

## **Resumen y Conclusiones**

2. En la estructura cristalográfica del complejo Abl-SH3/p41 se han identificado 5 moléculas de agua enterradas en la interfase de unión que podrían justificar adecuadamente el patrón termodinámico observado. Las moléculas de agua se localizan en dos zonas, una estaría localizada en el bolsillo de especificidad y contiene tres moléculas de agua que establecen interacciones con los restos N94 a E98 del lazo n-Src y con la región de especificidad del ligando p41; mientras que la segunda, localizada en la región más conservada del dominio, contiene 2 moléculas de agua que establecen interacciones con los restos de la hélice 3<sub>10</sub> y P7 y P8 del ligando.

3. El análisis calorimétrico y estructural indica que las mutaciones N94A/T/Q, N96A y N114A/T/Q del dominio Abl-SH3 silvestre, cuyas cadenas laterales establecen interacciones con las moléculas de agua en la interfase de unión, no parecen producir *per se* cambios estructurales significativos ni en las propiedades conformacionales con respecto al dominio silvestre, a excepción de los mutantes N94A/Q.

4. Las mutaciones diseñadas en Abl-SH3 producen un impacto significativo en la energética de unión al ligando p41 que podría deberse a la alteración de las interacciones mediadas por las moléculas de agua en la interfase de unión. Los mutantes de la posición 114, que poseen una estabilidad y propiedades conformacionales más similares a la proteína silvestre son, precisamente, los que muestran las diferencias más notables en la energética de unión al ligando.

5. La caracterización estructural por difracción de rayos X de los mutantes N114A y N114Q y de sus complejos con el ligando p41 indica que no existen cambios conformacionales significativos con respecto a la proteína silvestre como consecuencia del proceso de unión. Así mismo, la alta resolución obtenida en las estructuras nos ha permitido comprobar que las aguas 1, 2, 4 y 5 de las cinco moléculas de agua presentes en Abl-SH3/p41 se conservan en los complejos N114A/p41 y N114Q/p41.

6. El análisis dinámico realizado con las estructuras Abl-SH3/p41, N114A/p41 y N114Q/p41 identifica claramente las dos regiones de hidratación descritas y nos

muestra que las moléculas de agua cristalográficas se encuentran en un intercambio continuo y espontáneo con el disolvente a lo largo de toda la trayectoria de dinámica molecular, ocupando posiciones equivalentes a las anteriores. Así, aunque con distintas frecuencias de intercambio, las aguas de ambas zonas se comportan independientemente durante la simulación y además muestran diferentes mecanismos de entrada y salida de las superficies de unión.

7. Los cambios en las propiedades dinámicas de las moléculas de agua en N114A/p41 y N114Q/p41 con respecto a Abl-SH3/p41 provocan alteraciones significativas, tanto de las interacciones indirectas que median las moléculas de agua como de las directas entre proteína y ligando, que justifican las diferencias en los valores de entalpía de unión al ligando p41.

8. Las diferencias existentes en los parámetros termodinámicos de unión a Abl-SH3 entre una serie de derivados monosustituídos de p41, que se diferencian en los restos 5 y 8, que están expuestos al disolvente y no establecen interacciones en la interfase de unión, son inesperadas para una interacción puramente hidrofóbica y sugieren que existe algún tipo de cooperatividad entre ambas posiciones.

9. Estos estudios confirman la importancia de las moléculas de disolvente para la interacción de dominios SH3 con ligandos ricos en prolina y proporciona una nueva visión de lo que hasta ahora se había considerado una interacción típicamente hidrofóbica. En particular, las aguas en posición 4, que establecen interacciones con los átomos del esqueleto peptídico de los restos E98, S113 y N114 muy conservados entre los dominios SH3, poseen unas propiedades dinámicas características de aguas que están fuertemente unidas a las proteínas. Esta posición es, de hecho, la única ocupada en Abl-SH3 e incluso aparecen aguas en ese lugar en la quinasa completa de Abl y en otras como la de Src y Hck, por lo que podría ser determinante para la funcionalidad de estos dominios *in vivo* en el contexto de las proteínas que los contienen.

11. El patrón termodinámico mostrado por las interacciones establecidas por otros dominios modulares que reconocen secuencias ricas en prolina, como Nedd4-WW3, Yap65-WW y Tsg101-UEV es igualmente incongruente con la elevada

## Resumen y Conclusiones

hidrofobicidad de sus sitios de unión y ligandos, y de igual naturaleza que en los dominios SH3.

12. La determinación estructural por difracción de rayos X del dominio Tsg101-UEV libre ha revelado la presencia de una molécula de agua que ocupa una posición ideal para mediar un puente de hidrógeno entre el resto S143 del dominio y el resto T4 de los ligandos derivados del dominio L del virus VIH, hecho que contribuiría a explicar en parte la elevada exotermicidad mostrada por la interacción. Parece, por tanto, que la complejidad mostrada en el reconocimiento de ligandos por dominios SH3 podría ser universal para todos los dominios modulares que reconocen prolina.

13. En el caso de los ligandos que contienen dominios víricos L tipo PPxY y PTAP, hemos comprobado que la energética de la interacción Tsg101-UEV/PTAP no está afectada por la presencia de PPxY, mientras que la de la interacción Nedd4-WW3/PPxY tampoco se ve afectada por la presencia de PTAP.

14. La presencia de determinados restos adyacentes a la secuencia PTAP origina diferencias importantes en la afinidad de los ligandos víricos naturales por Tsg101-UEV. Mientras los ligandos **VIH-11** (PE**PTAP**PEESF) y **Ébola-11** (IL**PTAP**PEYME) poseen afinidades muy parecidas (del orden de  $10^4 \text{ M}^{-1}$ ), el ligando del dominio L de la leucemia humana (YVE**PTAP**QVL) posee una afinidad un orden de magnitud inferior ( $\simeq 10^3 \text{ M}^{-1}$ ).

15. El ligando sintético M1 (EYPPY**PPY**PSG), diseñado originalmente para Yap65-WW, es el de más afinidad para Nedd4-WW3 ( $10^5 \text{ M}^{-1}$ ) debido a una contribución entrópica mucho menos desfavorable, probablemente debido al mayor contenido de Pro, que favorece la conformación de PPII necesaria para la unión al dominio. Los ligandos víricos naturales de ébola y de la leucemia humana poseen afinidades muy bajas (en el rango de  $10^3 \text{ M}^{-1}$ ) que vienen dictadas fundamentalmente por la presencia de la secuencia PPPY. Excepcionalmente, la interacción con el ligando **Leucemia-b** (SDPQI**PPY**VEP) con Nedd4-WW3 presenta un ligero incremento de afinidad que parece ser debido a la presencia de determinados restos adyacentes a PPPY. Esta interacción está más optimizada que

con Yap65-WW, lo que podría suponer un punto de partida para el diseño de ligandos de alta especificidad dentro de esta clase de dominios.



## **APÉNDICES**



**APÉNDICES****Apéndice I. Modelo de DSC con dependencia lineal frente a la temperatura para las funciones capacidad calorífica molar parcial de estado nativo y desnaturalizado**

$$T = x + 273.15;$$

$$R = 8.314E - 3;$$

$$T_m = x_m + 273.15 ;$$

$$a_m = a_0 + b_0 \cdot T_m ;$$

$$d_m = a_m - \Delta C_{P,m};$$

$$d_0 = d_m + e_0 \cdot T_m ;$$

$$C_{P,N} = d_0 + e_0 \cdot T ;$$

$$C_{P,D} = a_0 + b_0 \cdot T ;$$

$$\Delta C_P = C_{P,D} - C_{P,N} ;$$

$$\Delta H = \Delta H_m + (a_0 - d_0) \cdot (T - T_m) + \frac{(b_0 - e_0) \cdot (T^2 - T_m^2)}{2} ;$$

$$\Delta S_m = \frac{\Delta H_m}{T_m} ;$$

$$\Delta S = \Delta S_m + (a_0 - d_0) \cdot \ln\left(\frac{T}{T_m}\right) + (b_0 - e_0) \cdot (T - T_m);$$

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

$$K = \exp\left(-\frac{\Delta G}{R \cdot T}\right);$$

$$y = C_{P,N} + \frac{\Delta C_P \cdot K}{1 + K} + \frac{\Delta H^2 \cdot K}{(R \cdot T^2) \cdot (1 + K)^2} ;$$



**Apéndice II. Condiciones de cristalización obtenidas para los mutantes de Abl-SH3**

<b>Proteína/ Mutante</b>	<b>Tampón pH</b>	<b>Precipitante</b>	<b>Precipitante</b>	<b>Temperatura °C</b>	<b>Técnica</b>	<b>Tiempo aparición</b>
N114A	HEPES 0.1M pH 7.0	Sulfato amónico 2 M	PEG300 5%	15	Gota sentada	5-10 días
N114A	Citr. sódico 0.1 M pH 5.0	Sulfato amónico 2 M	PEG300 5%	15	Gota sentada	15-30 días
N114A	Citr. sódico 0.1 M pH 4.0	Sulfato amónico 2 M	PEG300 5%	15	Gota sentada	15-30 días
N114A	Citr. sódico 0.1 M pH 3.0	Sulfato amónico 2 M	PEG300 5%	15	Gota sentada	15-30 días
N114T	HEPES 0.1M pH 7.0	Sulfato amónico 2 M	PEG300 5%	15	Gota sentada	15-30 días
N114Q	Citr. sódico 0.1 M pH 3.0	Sulfato amónico 2 M	PEG300 5%	15	Gota sentada	15-30 días



**Apéndice III. Condiciones de cristalización obtenidas para los complejos de los mutantes de Abi-SH3 con p41**

<b>Complejo</b>	<b>Tampón pH</b>	<b>Precipitante</b>	<b>Precipitante</b>	<b>Temperatura °C</b>	<b>Técnica</b>	<b>Tiempo de aparición</b>
N114A/p41	MES 0.1M pH 6.0	Sulfato amónico 3 M	-	20	Contra-difusión	5-10 días
N114A/p41	HEPES 0.1 M pH 7.0	Sulfato amónico 3 M	-	20	Contra-difusión	5-10 días
N114A/p41	TRIS 0.1 M pH 8.0	Sulfato amónico 3 M	-	20	Contra-difusión	5-10 días
N114Q/p41	Citr. sódico 0.1 M pH 3.5	Sulfato amónico 2 M	Cloruro sódico 0.3 M	15	Gota sentada	15-30 días
N114Q/p41	Citr. sódico 0.1 M pH 3.5	Sulfato amónico 2 M	Cloruro sódico 0.4 M	15	Gota sentada	15-30 días
N114Q/p41	Citr. sódico 0.1 M pH 3.5	Sulfato amónico 2 M	Cloruro sódico 0.5 M -	15	Gota sentada	15-30 días
N114Q/p41	Citr. sódico 0.1 M pH 3.5	Sulfato amónico 2 M	Cloruro sódico 0.6 M -	15	Gota sentada	15-30 días
N114Q/p41	Citr. sódico 0.1M pH 3.5	Sulfato amónico 2 M	Cloruro sódico 0.7 M -	15	Gota sentada	15-30 días

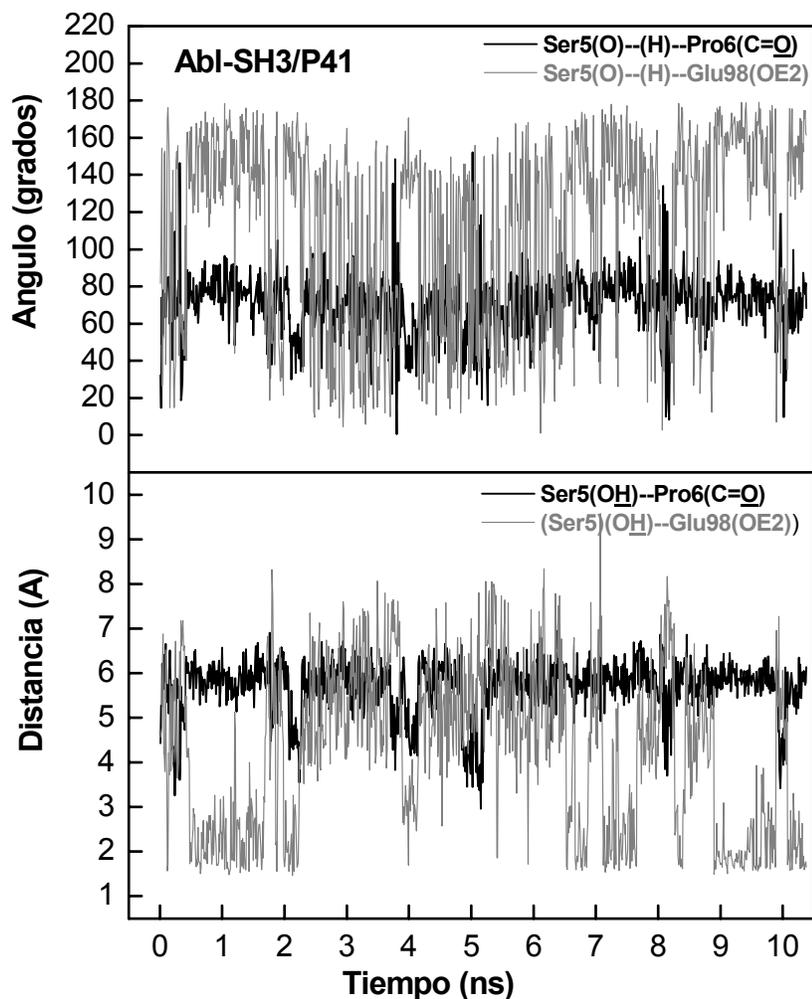


#### **Apéndice IV. Estudio de la interacción intramolecular S5-P6 en el ligando p41**

Hemos analizado el puente de hidrógeno intramolecular entre los restos S5 y P6 del ligando p41, observado en las estructuras cristalográficas Abl-SH3/p41, N114A/p41 y N114Q/p41 (Figura 3.5, página 161) y hemos observado que no llega a formarse en ninguna de nuestras simulaciones. Este puente de hidrógeno se ha considerado importante para el ligando al dotar de adaptabilidad al péptido p41, ya que conecta la parte encargada de la especificidad (APSYS) y la que adopta la estructura tipo poliprolina II (PPPPP) que se acomoda en los bolsillos hidrofóbicos de la proteína (Pisabarro *et al.* 1994; Pisabarro *et al.* 1998). En su lugar, la S5 del ligando establece un puente de hidrógeno con el resto Glu98 de Abl-SH3, que como hemos visto en la sección de dinámica molecular (3.1.1.2.5.2.1) tiene un frecuencia de formación considerable en todos los casos.

Para comprobar este resultado, hemos llevado a cabo un estudio del ángulo y la distancia de la interacción entre los átomos que formarían parte del puente de hidrógeno entre S5 y P6, así como del puente de hidrógeno que S5 establece con Glu98 (Figura A.2). Para el enlace intramolecular en el ligando hemos analizado el ángulo entre oxígeno-hidrógeno del hidroxilo de S5 con el oxígeno del carbonilo de P6 y la distancia entre el hidrógeno del hidroxilo de S5 y el oxígeno del carbonilo de P6. Durante la totalidad de la simulación hemos visto que la distancia entre éstos oscila entre 4 y 6 Å y el ángulo está por encima de 80 grados, por lo que está lejos de la geometría ideal para la formación del puente de hidrógeno (2.8 Å, y 50 grados) (Jiang *et al.* 2005)). Para el puente de hidrógeno intermolecular entre S5 y Glu98 la distancia durante la simulación frecuente valores entre 2 y 3 Å y del ángulo en torno a 20 grados, los cuales están más cerca de los valores óptimos para este puente de hidrógeno, y propicia la interacción durante un tiempo considerable.

El hecho de que esta interacción no se observe en ninguna de las estructuras cristalográficas podría estar relacionado con el contacto cristalográfico que establece el E98 con el resto en la posición 114 de otra molécula del dominio en la unidad asimétrica.



**Figura A.2.** Análisis del ángulo y distancia entre los átomos formadores del puente de hidrógeno observado entre S5 y P6 del ligando p41 en la estructura cristalográfica Abl-SH3/p41; y del puente de hidrógeno observado en la simulaciones de dinámica molecular que se establece entre los restos S5 del ligando y Glu98 de la proteína.

Es posible que este contacto cristalográfico pueda forzar una orientación de la cadena lateral de Glu98 que no es la adecuada para formar la interacción con S5 del ligando.

## **BIBLIOGRAFÍA**



**BIBLIOGRAFÍA**

1. Aghazadeh, B. & Rosen, M. K. (1999). Ligand recognition by SH3 and WW domains: the role of N-alkylation in PPII helices. *Chem Biol* **6**, R241-6.
2. Arold, S., O'Brien, R., Franken, P., Strub, M. P., Hoh, F., Dumas, C. & Ladbury, J. E. (1998). RT loop flexibility enhances the specificity of Src family SH3 domains for HIV-1 Nef. *Biochemistry* **37**, 14683-91.
3. Arold, S. T., Ulmer, T. S., Mulhern, T. D., Werner, J. M., Ladbury, J. E., Campbell, I. D. & Noble, M. E. (2001). The role of the Src homology 3-Src homology 2 interface in the regulation of Src kinases. *J Biol Chem* **276**, 17199-205.
4. Azam, M., Latek, R. R. & Daley, G. Q. (2003). Mechanisms of autoinhibition and STI-571/Imatinib resistance revealed by mutagenesis of BCR-ABL. *Cell* **112**, 831-43.
5. Bagchi, B. (2005). Water dynamics in the hydration layer around proteins and micelles. *Chem Rev* **105**, 3197-219.
6. Baker, B. M. & Murphy, K. P. (1996). Evaluation of linked protonation effects in protein binding reactions using isothermal titration calorimetry. *Biophys J* **71**, 2049-55.
7. Ball, L. J., Jarchau, T., Oschkinat, H. & Walter, U. (2002). EVH1 domains: structure, function and interactions. *FEBS Lett* **513**, 45-52.
8. Ball, L. J., 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.
9. Barila, D. & Superti-Furga, G. (1998). An intramolecular SH3-domain interaction regulates c-Abl activity. *Nat Genet* **18**, 280-2.
10. Barillari, C., Taylor, J., Viner, R. & Essex, J. W. (2007). Classification of water molecules in protein binding sites. *J Am Chem Soc* **129**, 2577-87.
11. Bauer, F., Schweimer, K., Meiselbach, H., Hoffmann, S., Rosch, P. & Sticht, H. (2005). Structural characterization of Lyn-SH3 domain in complex with a herpesviral protein reveals an extended recognition motif that enhances binding affinity. *Protein Sci* **14**, 2487-98.
12. Ben-Naim, A. (2002). Molecular recognition--viewed through the eyes of the solvent. *Biophys Chem* **101-102**, 309-19.
13. Ben-Naim, A. (2003). Hydrophobic hydrophilic phenomena in biochemical processes. *Biophys Chem* **105**, 183-93.
14. Berthoux, L., Pechoux, C., Ottmann, M., Morel, G. & Darlix, J. L. (1997). Mutations in the N-terminal domain of human immunodeficiency virus type 1 nucleocapsid protein affect virion core structure and proviral DNA synthesis. *J Virol* **71**, 6973-81.
15. Bhat, T. N., Bentley, G. A., Boulot, G., Greene, M. I., Tello, D., Dall'Acqua, W., Souchon, H., Schwarz, F. P., Mariuzza, R. A. & Poljak, R. J. (1994). Bound water molecules and conformational stabilization help mediate an antigen-antibody association. *Proc Natl Acad Sci U S A* **91**, 1089-93.
16. Blanco, F. J., Ortiz, A. R. & Serrano, L. (1997). <sup>1</sup>H and <sup>15</sup>N NMR assignment and solution structure of the SH3 domain of spectrin: comparison of unrefined and refined structure sets with the crystal structure. *J Biomol NMR* **9**, 347-57.
17. Borchert, T. V., Mathieu, M., Zeelen, J. P., Courtneidge, S. A. & Wierenga, R. K. (1994). The crystal structure of human CskSH3: structural diversity near the RT-Src and n-Src loop. *FEBS Lett* **341**, 79-85.

## Bibliografía

18. Bork, P. & Sudol, M. (1994). The WW domain: a signalling site in dystrophin? *Trends Biochem Sci* **19**, 531-3.
19. Bouamr, F., Melillo, J. A., Wang, M. Q., Nagashima, K., de Los Santos, M., Rein, A. & Goff, S. P. (2003). PPPYEPTAP 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. *J. Virol.* **77**, 11882-11895.
20. 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**, 664-70.
21. Brannetti, B. & Helmer-Citterich, M. (2003). iSPOT: A web tool to infer the interaction specificity of families of protein modules. *Nucleic Acids Res* **31**, 3709-11.
22. Brasher, B. B., Roumiantsev, S. & Van Etten, R. A. (2001). Mutational analysis of the regulatory function of the c-Abl Src homology 3 domain. *Oncogene* **20**, 7744-52.
23. Breslauer, K. J., Freire, E. & Straume, M. (1992). Calorimetry: a tool for DNA and ligand-DNA studies. *Methods Enzymol* **211**, 533-67.
24. Burchat, A., Borhani, D. W., Calderwood, D. J., Hirst, G. C., Li, B. & Stachlewitz, R. F. (2006). Discovery of A-770041, a src-family selective orally active Ick inhibitor that prevents organ allograft rejection. *Bioorg Med Chem Lett* **16**, 118-22.
25. Camara-Artigas, A., Palencia, A., Martinez, J. C., Luque, I., Gavira, J. A. & Garcia-Ruiz, J. M. (2007). Crystallization by capillary counter-diffusion and structure determination of the N114A mutant of the SH3 domain of Abl tyrosine kinase complexed with a high-affinity peptide ligand. *Acta Crystallogr D Biol Crystallogr* **63**, 646-52.
26. Candel, A. M., Conejero-Lara, F., Martinez, J. C., van Nuland, N. A. & Bruix, M. (2007). The high-resolution NMR structure of a single-chain chimeric protein mimicking a SH3-peptide complex. *FEBS Lett* **581**, 687-92.
27. Casares, S., Ab, E., Eshuis, H., Lopez-Mayorga, O., van Nuland, N. A. & Conejero-Lara, F. (2007a). The high-resolution NMR structure of the R21A Spc-SH3:P41 complex: understanding the determinants of binding affinity by comparison with Abl-SH3. *BMC Struct Biol* **7**, 22.
28. Casares, S., Lopez-Mayorga, O., Vega, M. C., Camara-Artigas, A. & Conejero-Lara, F. (2007b). Cooperative propagation of local stability changes from low-stability and high-stability regions in a SH3 domain. *Proteins* **67**, 531-47.
29. Casares, S., Sadqi, M., Lopez-Mayorga, O., Conejero-Lara, F. & van Nuland, N. A. (2004). Detection and characterization of partially unfolded oligomers of the SH3 domain of alpha-spectrin. *Biophys J* **86**, 2403-13.
30. Casares, S., Sadqi, M., Lopez-Mayorga, O., Martinez, J. C. & Conejero-Lara, F. (2003). Structural cooperativity in the SH3 domain studied by site-directed mutagenesis and amide hydrogen exchange. *FEBS Lett* **539**, 125-30.
31. Case, D. A., Darden, T. E., Cheatham, I. T. E., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Merz, K. M., Wang, D. A., Pearlman, M., Crowley, M., Brozell, S., Tsui, V., Gohlke, H., Mongan, J., Hornak, V., Cui, G., Beroza, P., Schafmeister, C., Cladwell, J. W., Ross, W. S. & Kollman, P. A. (2004). University of California, San Francisco.
32. Cesareni, G., Panni, S., Nardelli, G. & Castagnoli, L. (2002). Can we infer peptide recognition specificity mediated by SH3 domains? *FEBS Lett* **513**, 38-44.

33. Chaires, J. B. (1997). Possible origin of differences between van't Hoff and calorimetric enthalpy estimates. *Biophys Chem* **64**, 15-23.
34. Cheatham, T. E. I., Miller, J. L., Fox, T., Darden, T. A. & Kollman, P. A. (1995). Molecular Dynamics Simulations on Solvated Biomolecular Systems: The Particle Mesh Ewald Method Leads to Stable Trajectories of DNA, RNA, and Proteins. *J. Am. Chem. Soc.* **117**, 4193-4194.
35. Chen, D. L. & Kellogg, G. E. (2005). A computational tool to optimize ligand selectivity between two similar biomacromolecular targets. *J Comput Aided Mol Des* **19**, 69-82.
36. Chen, H. I. & Sudol, M. (1995). The WW domain of Yes-associated protein binds a proline-rich ligand that differs from the consensus established for Src homology 3-binding modules. *Proc Natl Acad Sci U S A* **92**, 7819-23.
37. Chen, J. M., Xu, S. L., Wawrzak, Z., Basarab, G. S. & Jordan, D. B. (1998). Structure-based design of potent inhibitors of scytalone dehydratase: displacement of a water molecule from the active site. *Biochemistry* **37**, 17735-44.
38. Chong, P. A., Lin, H., Wrana, J. L. & Forman-Kay, J. D. (2006). An expanded WW domain recognition motif revealed by the interaction between Smad7 and the E3 ubiquitin ligase Smurf2. *J Biol Chem* **281**, 17069-75.
39. Christensen, J. J., Hansen, L. D. & Izatt, R. M. (1976). Handbook of Proton Ionization Heats and Related Thermodynamic Quantities. *Wiley, New York*.
40. Cicchetti, P., Mayer, B. J., Thiel, G. & Baltimore, D. (1992). Identification of a protein that binds to the SH3 region of Abl and is similar to Bcr and GAP-rho. *Science* **257**, 803-6.
41. Cirinna, M., Trotta, R., Salomoni, P., Kossev, P., Wasik, M., Perrotti, D. & Calabretta, B. (2000). Bcl-2 expression restores the leukemogenic potential of a BCR/ABL mutant defective in transformation. *Blood* **96**, 3915-21.
42. Clarke, C., Woods, R. J., Gluska, J., Cooper, A., Nutley, M. A. & Boons, G. J. (2001). Involvement of water in carbohydrate-protein binding. *J Am Chem Soc* **123**, 12238-47.
43. Combs, A. P., Kapoor, T. M., Feng, S., Chen, J. K., Daude-Snow, L. F. & Schreiber, S. L. (1996). Protein Structure-Based Combinatorial Chemistry: Discovery of Non-Peptide Binding Elements to Src SH3 Domain. *J. Am. Chem. Soc.* **118**, 287-288.
44. Conejero-Lara, F. & Mateo, P. L. (1996). Presence of a slow dimerization equilibrium on the thermal unfolding of the 205-316 thermolysin fragment at neutral pH. *Biochemistry* **35**, 3477-86.
45. Connelly, P. R., Aldape, R. A., Bruzzese, F. J., Chambers, S. P., Fitzgibbon, M. J., Fleming, M. A., Itoh, S., Livingston, D. J., Navia, M. A., Thomson, J. A. & et al. (1994). Enthalpy of hydrogen bond formation in a protein-ligand binding reaction. *Proc Natl Acad Sci U S A* **91**, 1964-8.
46. 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.
47. Cordier, F., Wang, C., Grzesiek, S. & Nicholson, L. K. (2000). Ligand-induced strain in hydrogen bonds of the c-Src SH3 domain detected by NMR. *J Mol Biol* **304**, 497-505.
48. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W. & Kollman, P. A. (1995). A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **117**, 5179-5197.

## Bibliografía

49. Covell, D. G. & Wallqvist, A. (1997). Analysis of protein-protein interactions and the effects of amino acid mutations on their energetics. The importance of water molecules in the binding epitope. *J Mol Biol* **269**, 281-97.
50. Cowan-Jacob, S. W., Fendrich, G., Manley, P. W., Jahnke, W., Fabbro, D., Liebetanz, J. & Meyer, T. (2005). The crystal structure of a c-Src complex in an active conformation suggests possible steps in c-Src activation. *Structure* **13**, 861-71.
51. Crane, J. C., Koepf, E. K., Kelly, J. W. & Gruebele, M. (2000). Mapping the transition state of the WW domain beta-sheet. *J Mol Biol* **298**, 283-92.
52. Creamer, T. P. (1998). Left-handed polyproline II helix formation is (very) locally driven. *Proteins* **33**, 218-26.
53. Creamer, T. P. & Campbell, M. N. (2002). Determinants of the polyproline II helix from modeling studies. *Adv Protein Chem* **62**, 263-82.
54. Creighton, T. E. (1989). *Protein Structure a practical approach*. Practical Approach Series (Rickwood, D., Hames, B.D., Ed.), Oxford (England).
55. Dalgarno, D. C., Botfield, M. C. & Rickles, R. J. (1997). SH3 domains and drug design: ligands, structure, and biological function. *Biopolymers* **43**, 383-400.
56. Dall'Acqua, W., Goldman, E. R., Lin, W., Teng, C., Tsuchiya, D., Li, H., Ysern, X., Braden, B. C., Li, Y., Smith-Gill, S. J. & Mariuzza, R. A. (1998). A mutational analysis of binding interactions in an antigen-antibody protein-protein complex. *Biochemistry* **37**, 7981-91.
57. Damjanovic, A., Garcia-Moreno, B., Lattman, E. E. & Garcia, A. E. (2005). Molecular dynamics study of water penetration in staphylococcal nuclease. *Proteins* **60**, 433-49.
58. DeLeeuw, S. W., Perran, J. W. & Smith, E. R. (1980). Proc. R. Soc. London Ser. A **373**, 27.
59. Demchenko, A. P. (1981). *Ultraviolet Spectroscopy of Proteins*, Springer-Verlag, New York (USA).
60. Demirov, D. G. & Freed, E. O. (2004). Retrovirus budding. *Virus Res* **106**, 87-102.
61. Demirov, D. G., Ono, A., Orenstein, J. M. & Freed, E. O. (2002). Overexpression of the N-terminal domain of TSG101 inhibits HIV-1 budding by blocking late domain function. *PNAS* **99**, 955-960.
62. Denisov, V. P. & Halle, B. (1995a). Hydrogen exchange and protein hydration: the deuteron spin relaxation dispersions of bovine pancreatic trypsin inhibitor and ubiquitin. *J Mol Biol* **245**, 698-709.
63. Denisov, V. P. & Halle, B. (1995b). Protein hydration dynamics in aqueous solution: a comparison of bovine pancreatic trypsin inhibitor and ubiquitin by oxygen-17 spin relaxation dispersion. *J Mol Biol* **245**, 682-97.
64. Denisov, V. P. & Halle, B. (1996). Protein hydration dynamics in aqueous solution. *Faraday Discuss*, 227-44.
65. Denisov, V. P., Peters, J., Horlein, H. D. & Halle, B. (1996). Using buried water molecules to explore the energy landscape of proteins. *Nat Struct Biol* **3**, 505-9.
66. Ducruix, A. & Giegé, R. (1999). *Oxford, Oxford University Press*.
67. Dunitz, J. D. (1994). The Entropic Cost of Bound Water in Crystals and Biomolecules. *Science* **264**, 670.
68. Dunitz, J. D. (1995). Win some, lose some: enthalpy-entropy compensation in weak intermolecular interactions. *Chem Biol* **2**, 709-12.
69. Edelhoch, H. (1967). Spectroscopic determination of tryptophan and tyrosine in proteins. *Biochemistry* **6**, 1948-54.

70. Eftink, M. R., Anusiem, A. C. & Biltonen, R. L. (1983). Enthalpy-entropy compensation and heat capacity changes for protein-ligand interactions: general thermodynamic models and data for the binding of nucleotides to ribonuclease A. *Biochemistry* **22**, 3884-96.
71. Feher, G. & Kam, Z. (1985). Nucleation and growth of protein crystals: general principles and assays. *Methods Enzymol* **114**, 77-112.
72. Feng, G. H., Lih, C. J. & Cohen, S. N. (2000). TSG101 protein steady-state level is regulated posttranslationally by an evolutionarily conserved COOH-terminal sequence. *Cancer Res* **60**, 1736-41.
73. Feng, S., Chen, J. K., Yu, H., Simon, J. A. & Schreiber, S. L. (1994). Two binding orientations for peptides to the Src SH3 domain: development of a general model for SH3-ligand interactions. *Science* **266**, 1241-7.
74. Feng, S., Kapoor, T. M., Shirai, F., Combs, A. P. & Schreiber, S. L. (1996). Molecular basis for the binding of SH3 ligands with non-peptide elements identified by combinatorial synthesis. *Chem Biol* **3**, 661-70.
75. Feng, S., Kasahara, C., Rickles, R. J. & Schreiber, S. L. (1995). Specific interactions outside the proline-rich core of two classes of Src homology 3 ligands. *Proc Natl Acad Sci U S A* **92**, 12408-15.
76. Ferguson, N., Johnson, C. M., Macias, M., Oschkinat, H. & Fersht, A. (2001). Ultrafast folding of WW domains without structured aromatic clusters in the denatured state. *Proc Natl Acad Sci U S A* **98**, 13002-7.
77. Fernandez-Ballester, G., Blanes-Mira, C. & Serrano, L. (2004). The tryptophan switch: changing ligand-binding specificity from type I to type II in SH3 domains. *J Mol Biol* **335**, 619-29.
78. Ferreon, J. C. & Hilser, V. J. (2003a). The effect of the polyproline II (PPII) conformation on the denatured state entropy. *Protein Sci* **12**, 447-57.
79. Ferreon, J. C. & Hilser, V. J. (2003b). Ligand-induced changes in dynamics in the RT loop of the C-terminal SH3 domain of Sem-5 indicate cooperative conformational coupling. *Protein Sci* **12**, 982-96.
80. Ferreon, J. C. & Hilser, V. J. (2004). Thermodynamics of binding to SH3 domains: the energetic impact of polyproline II (PII) helix formation. *Biochemistry* **43**, 7787-97.
81. Feynman, R. (1963). "Lectures on Physics".
82. Filimonov, V. V., Azuaga, A. I., Viguera, A. R., Serrano, L. & Mateo, P. L. (1999). A thermodynamic analysis of a family of small globular proteins: SH3 domains. *Biophys Chem* **77**, 195-208.
83. Filimonov, V. V., Potekhin, S. A., Matveev, S. V. & Privalov, P. L. (1982). [Thermodynamic analysis of scanning microcalorimetry data. 1. Algorithms for deconvolution of heat absorption curves]. *Mol Biol (Mosk)* **16**, 551-62.
84. Freed, E. O. (2002). Viral late domains. *J Virol* **76**, 4679-87.
85. Freed, E. O. (2003). The HIV-TSG101 interface: recent advances in a budding field. *Trends in Microbiology* **11**, 56-59.
86. Freire, E. (1999). The propagation of binding interactions to remote sites in proteins: analysis of the binding of the monoclonal antibody D1.3 to lysozyme. *Proc Natl Acad Sci U S A* **96**, 10118-22.
87. Freire, E. (2002). Designing drugs against heterogeneous targets. *Nat Biotechnol* **20**, 15-6.
88. Freund, C., Dotsch, V., Nishizawa, K., Reinherz, E. L. & Wagner, G. (1999). The GYF domain is a novel structural fold that is involved in lymphoid signaling through proline-rich sequences. *Nat Struct Biol* **6**, 656-60.

## Bibliografía

89. Fukada, H. & Takahashi, K. (1998). Enthalpy and heat capacity changes for the proton dissociation of various buffer components in 0.1 M potassium chloride. *Proteins* **33**, 159-66.
90. Garbay, C., Liu, W. Q., Vidal, M. & Roques, B. P. (2000). Inhibitors of Ras signal transduction as antitumor agents. *Biochem Pharmacol* **60**, 1165-9.
91. Garcia, A. E. & Hummer, G. (2000). Water penetration and escape in proteins. *Proteins* **38**, 261-72.
92. Garcia-Ruiz, J. M. (1991). *Key Eng. Mater.* **88**, 87-106.
93. Garcia-Ruiz, J. M. (2003). Counterdiffusion methods for macromolecular crystallization. *Methods Enzymol* **368**, 130-54.
94. Garcia-Sosa, A. T., Firth-Clark, S. & Mancera, R. L. (2005). Including tightly-bound water molecules in de novo drug design. Exemplification through the in silico generation of poly(ADP-ribose)polymerase ligands. *J Chem Inf Model* **45**, 624-33.
95. Garcia-Sosa, A. T. & Mancera, R. L. (2006). The effect of a tightly bound water molecule on scaffold diversity in the computer-aided de novo ligand design of CDK2 inhibitors. *J Mol Model* **12**, 422-31.
96. Garcia-Sosa, A. T., Mancera, R. L. & Dean, P. M. (2003). WaterScore: a novel method for distinguishing between bound and displaceable water molecules in the crystal structure of the binding site of protein-ligand complexes. *J Mol Model* **9**, 172-82.
97. Garrus, J. E., von Schwedler, U. K., Pornillos, O. W., Morham, S. G., Zavitz, K. H., Wang, H. E., Wettstein, D. A., Stray, K. M., Cote, M. & Rich, R. L. (2001). Tsg101 and the Vacuolar Protein Sorting Pathway Are Essential for HIV-1 Budding. *Cell* **107**, 55-65.
98. Ghose, R., Shekhtman, A., Goger, M. J., Ji, H. & Cowburn, D. (2001). A novel, specific interaction involving the Csk SH3 domain and its natural ligand. *Nat Struct Biol* **8**, 998-1004.
99. Ghosh, A., Rapp, C. S. & Friesner, R. A. (1998). Generalized Born Model Based on a Surface Integral Formulation. *J. Phys. Chem. B* **102**, 10983-10990.
100. Gill, S. C. & von Hippel, P. H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem* **182**, 319-26.
101. Goila-Gaur, R., Demirov, D. G., Orenstein, J. M., Ono, A. & Freed, E. O. (2003). Defects in human immunodeficiency virus budding and endosomal sorting induced by TSG101 overexpression. *J Virol* **77**, 6507-19.
102. Gomez, J. & Freire, E. (1995). Thermodynamic mapping of the inhibitor site of the aspartic protease endothiapepsin. *J Mol Biol* **252**, 337-50.
103. Gorina, S. & Pavletich, N. P. (1996). Structure of the p53 tumor suppressor bound to the ankyrin and SH3 domains of 53BP2. *Science* **274**, 1001-5.
104. Gottlinger, H., Dorfman, T., Sodroski, J. & Haseltine, W. (1991). Effect of Mutations Affecting the p6 gag Protein on Human Immunodeficiency Virus Particle Release. *PNAS* **88**, 3195-3199.
105. Gottwein, E., Bodem, J., Muller, B., Schmechel, A., Zentgraf, H. & Krausslich, H. G. (2003). The Mason-Pfizer monkey virus PPPY and PSAP motifs both contribute to virus release. *J Virol* **77**, 9474-85.
106. Grest, K., S. & Kremer, K. (1986). Molecular-dynamics simulation for polymers in the presence of a heat bath. *Phys. Rev. A* **33(5)**, 3628-3631.
107. Guo, X., Kameoka, M., Wei, X., Roques, B., Gotte, M., Liang, C. & Wainberg, M. A. (2003). Suppression of an intrinsic strand transfer activity of HIV-1 Tat protein by its second-exon sequences. *Virology* **307**, 154-63.

108. Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp Ser* **266**, 418-427.
109. Halle, B. (2004). Biomolecular cryocrystallography: structural changes during flash-cooling. *Proc Natl Acad Sci U S A* **101**, 4793-8.
110. Harty, R. N., Brown, M. E., Wang, G., Huibregtse, J. & Hayes, F. P. (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**, 13871-6.
111. Harvey, K. F., Dinudom, A., Komwatana, P., Jolliffe, C. N., Day, M. L., Parasivam, G., Cook, D. I. & Kumar, S. (1999). All three WW domains of murine Nedd4 are involved in the regulation of epithelial sodium channels by intracellular Na<sup>+</sup>. *J Biol Chem* **274**, 12525-30.
112. Harvey, K. F. & Kumar, S. (1999). Nedd4-like proteins: an emerging family of ubiquitin-protein ligases implicated in diverse cellular functions. *Trends Cell Biol* **9**, 166-9.
113. Helms, V. (2007). Protein dynamics tightly connected to the dynamics of surrounding and internal water molecules. *Chemphyschem* **8**, 23-33.
114. Henriques, D. A. & Ladbury, J. E. (2001). Inhibitors to the Src SH2 domain: a lesson in structure--thermodynamic correlation in drug design. *Arch Biochem Biophys* **390**, 158-68.
115. Hershko, A. & Ciechanover, A. (1998). The ubiquitin system. *Annu Rev Biochem* **67**, 425-79.
116. Hilser, V. J., Dowdy, D., Oas, T. G. & Freire, E. (1998). The structural distribution of cooperative interactions in proteins: analysis of the native state ensemble. *Proc Natl Acad Sci U S A* **95**, 9903-8.
117. Hilser, V. J., Townsend, B. D. & Freire, E. (1997). Structure-based statistical thermodynamic analysis of T4 lysozyme mutants: structural mapping of cooperative interactions. *Biophys Chem* **64**, 69-79.
118. Hofmann, R. M. & Pickart, C. M. (1999). Noncanonical MMS2-encoded ubiquitin-conjugating enzyme functions in assembly of novel polyubiquitin chains for DNA repair. *Cell* **96**, 645-53.
119. Holt, M. R. & Koffer, A. (2001). Cell motility: proline-rich proteins promote protrusions. *Trends Cell Biol* **11**, 38-46.
120. Horn, J. R., Russell, D., Lewis, E. A. & Murphy, K. P. (2001). Van't Hoff and calorimetric enthalpies from isothermal titration calorimetry: are there significant discrepancies? *Biochemistry* **40**, 1774-8.
121. 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, J. J. (2004). A map of WW domain family interactions. *Proteomics* **4**, 643-55.
122. Huang, M., Orenstein, J. M., Martin, M. A. & Freed, E. O. (1995). p6Gag is required for particle production from full-length human immunodeficiency virus type 1 molecular clones expressing protease. *J Virol* **69**, 6810-8.
123. Huang, X., Poy, F., Zhang, R., Joachimiak, A., Sudol, M. & Eck, M. J. (2000). Structure of a WW domain containing fragment of dystrophin in complex with beta-dystroglycan. *Nat Struct Biol* **7**, 634-8.
124. Humphrey, W., Dalke, A. & Schulten, K. (1996). VMD: visual molecular dynamics. *J Mol Graph* **14**, 33-8, 27-8.

## Bibliografia

125. Irie, T., Licata, J. M., McGettigan, J. P., Schnell, M. J. & Harty, R. N. (2004). Budding of PPxY-containing rhabdoviruses is not dependent on host proteins TGS101 and VPS4A. *J Virol* **78**, 2657-65.
126. Jager, M., Nguyen, H., Crane, J. C., Kelly, J. W. & Gruebele, M. (2001). The folding mechanism of a beta-sheet: the WW domain. *J Mol Biol* **311**, 373-93.
127. Jager, M., Nguyen, H., Dendle, M., Gruebele, M. & Kelly, J. W. (2007). Influence of hPin1 WW N-terminal domain boundaries on function, protein stability, and folding. *Protein Sci* **16**, 1495-501.
128. Janin, J. (1999). Wet and dry interfaces: the role of solvent in protein-protein and protein-DNA recognition. *Structure* **7**, R277-9.
129. Jelen, F., Oleksy, A., Smietana, K. & Otlewski, J. (2003). PDZ domains - common players in the cell signaling. *Acta Biochim Pol* **50**, 985-1017.
130. Jelesarov, I. & Bosshard, H. R. (1999). Isothermal titration calorimetry and differential scanning calorimetry as complementary tools to investigate the energetics of biomolecular recognition. *J Mol Recognit* **12**, 3-18.
131. Jiang, L., Kuhlman, B., Kortemme, T. & Baker, D. (2005). A "solvated rotamer" approach to modeling water-mediated hydrogen bonds at protein-protein interfaces. *Proteins* **58**, 893-904.
132. Jorgensen, W. L., Chandrasekar, J., Madura, J. D., Impey, R. W. & M.L., K. (1983). *J. Chem. Phys.* **79**, 9266.
133. Jorgensen, W. L. & Gao, J. (1988). Cis-trans energy difference for the peptide bond in the gas phase and in aqueous solution. *J. Am. Chem. Soc.* **110**, 4212-4216.
134. Kam, Z., Shore, H. B. & Feher, G. (1978). On the crystallization of proteins. *J Mol Biol* **123**, 539-55.
135. Kanelis, V., Bruce, M. C., Skrynnikov, N. R., Rotin, D. & Forman-Kay, J. D. (2006). Structural determinants for high-affinity binding in a Nedd4 WW3\* domain-Comm PY motif complex. *Structure* **14**, 543-53.
136. Kang, H., Freund, C., Duke-Cohan, J. S., Musacchio, A., Wagner, G. & Rudd, C. E. (2000). SH3 domain recognition of a proline-independent tyrosine-based RKxxYxxY motif in immune cell adaptor SKAP55. *Embo J* **19**, 2889-99.
137. Karanicolas, J. & Brooks, C. L., 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**, 3432-7.
138. Kardinal, C., Konkol, B., Schulz, A., Posern, G., Lin, H., Adermann, K., Eulitz, M., Estrov, Z., Talpaz, M., Arlinghaus, R. B. & Feller, S. M. (2000). Cell-penetrating SH3 domain blocker peptides inhibit proliferation of primary blast cells from CML patients. *Faseb J* **14**, 1529-38.
139. Karplus, M. & McCammon, J. A. (2002). Molecular dynamics simulations of biomolecules. *Nat Struct Biol* **9**, 646-52.
140. Karplus, M. & Sali, A. (1995). Theoretical studies of protein folding and unfolding. *Curr Opin Struct Biol* **5**, 58-73.
141. Kato, Y., Miyakawa, T., Kurita, J. & Tanokura, M. (2006). Structure of FBP11 WW1-PL ligand complex reveals the mechanism of proline-rich ligand recognition by group II/III WW domains. *J Biol Chem* **281**, 40321-9.
142. Katzmann, D. J., Babst, M. & Emr, S. D. (2001). Ubiquitin-dependent sorting into the multivesicular body pathway requires the function of a conserved endosomal protein sorting complex, ESCRT-I. *Cell* **106**, 145-55.
143. Kay, B. K., Willianson, M. P. & 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.

144. Kellogg, G. E., Semus, S. F. & Abraham, D. J. (1991). HINT: a new method of empirical hydrophobic field calculation for CoMFA. *J Comput Aided Mol Des* **5**, 545-52.
145. Kelly, M. A., Chellgren, B. W., Rucker, A. L., Troutman, J. M., Fried, M. G., Miller, A. F. & Creamer, T. P. (2001). Host-guest study of left-handed polyproline II helix formation. *Biochemistry* **40**, 14376-83.
146. Kishan, K. V., Scita, G., Wong, W. T., Di Fiore, P. P. & Newcomer, M. E. (1997). The SH3 domain of Eps8 exists as a novel intertwined dimer. *Nat Struct Biol* **4**, 739-43.
147. Kliger, D. S., Lewis, J. W. & Randal, C. E. (1990). New York, Academy Press.
148. Kline, M. (1990). Mathematical Thought from Ancient to Modern Times. *Oxford University Press.*, 35-37.
149. Koepf, E. K., Petrassi, H. M., Ratnaswamy, G., Huff, M. E., Sudol, M. & Kelly, J. W. (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**, 14338-51.
150. Koepf, E. K., Petrassi, H. M., Sudol, M. & Kelly, J. W. (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**, 841-53.
151. Kollman, P. (1993). Free energy calculations: Applications to chemical and biochemical phenomena. *Chem. Rev.* **93**, 2395-2417.
152. Komada, M. & Kitamura, N. (2001). Hrs and hbp: possible regulators of endocytosis and exocytosis. *Biochem Biophys Res Commun* **281**, 1065-9.
153. Komuro, A., Saeki, M. & Kato, S. (1999). Association of two nuclear proteins, Npw38 and NpwBP, via the interaction between the WW domain and a novel proline-rich motif containing glycine and arginine. *J Biol Chem* **274**, 36513-9.
154. Konev, S. V. (1967). *Fluorescence and Phosphorescence of Proteins and Nucleic Acids*, Plenum Press, New York (USA).
155. Koonin, E. V. & Abagyan, R. A. (1997). TSG101 may be the prototype of a class of dominant negative ubiquitin regulators. *Nat Genet* **16**, 330-1.
156. Kortt, A. A., Gruen, L. C. & Oddie, G. W. (1997a). Influence of mass transfer and surface ligand heterogeneity on quantitative BIAcore binding data. Analysis of the interaction of NC10 Fab with an anti-idiotypic Fab'. *J Mol Recognit* **10**, 148-58.
157. Kortt, A. A., Oddie, G. W., Iliades, P., Gruen, L. C. & Hudson, P. J. (1997b). Nonspecific amine immobilization of ligand can be a potential source of error in BIAcore binding experiments and may reduce binding affinities. *Anal Biochem* **253**, 103-11.
158. Kostelansky, M. S., Sun, J., Lee, S., Kim, J., Ghirlando, R., Hierro, A., Emr, S. D. & Hurley, J. H. (2006). Structural and Functional Organization of the ESCRT-I Trafficking Complex. *Cell* **125**, 113-126.
159. Kuriyan, J. & Cowburn, D. (1997). Modular peptide recognition domains in eukaryotic signaling. *Annu Rev Biophys Biomol Struct* **26**, 259-88.
160. Lachenmann, M. J., Ladbury, J. E., Dong, J., Huang, K., Carey, P. & Weiss, M. A. (2004). Why zinc fingers prefer zinc: ligand-field symmetry and the hidden thermodynamics of metal ion selectivity. *Biochemistry* **43**, 13910-25.
161. Ladbury, J. E. & Arold, S. (2000). Searching for specificity in SH domains. *Chem Biol* **7**, R3-8.
162. Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-5.

## Bibliografía

163. Lafont, V., Armstrong, A. A., Ohtaka, H., Kiso, Y., Mario Amzel, L. & Freire, E. (2007). Compensating enthalpic and entropic changes hinder binding affinity optimization. *Chem Biol Drug Des* **69**, 413-22.
164. Lam, P. Y., Jadhav, P. K., Eyermann, C. J., Hodge, C. N., Ru, Y., Bacheler, L. T., Meek, J. L., Otto, M. J., Rayner, M. M., Wong, Y. N. & et al. (1994). Rational design of potent, bioavailable, nonpeptide cyclic ureas as HIV protease inhibitors. *Science* **263**, 380-4.
165. Landgraf, C., Panni, S., Montecchi-Palazzi, L., Castagnoli, L., Schneider-Mergener, J., Volkmer-Engert, R. & Cesareni, G. (2004). Protein interaction networks by proteome peptide scanning. *PLoS Biol* **2**, E14.
166. Laskowski, R. A., Moss, D. S. & Thornton, J. M. (1993). Main-chain bond lengths and bond angles in protein structures. *J Mol Biol* **231**, 1049-67.
167. Lawrence, D. S. (2005). Signaling protein inhibitors via the combinatorial modification of peptide scaffolds. *Biochim Biophys Acta* **1754**, 50-7.
168. Lazaridis, T. (1998a). Inhomogeneous fluid approach to solvation thermodynamics. 1. Theory. *Journal of Physical Chemistry B* **102**, 3531-3541.
169. Lazaridis, T. (1998b). Inhomogeneous fluid approach to solvation thermodynamics. 2. Applications to simple fluids. *Journal of Physical Chemistry B* **102**, 3542-3550.
170. Lee, K. Y., Hyeok Yoon, J. H., Kim, M., Roh, S., Lee, Y. S., Seong, B. L. & 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**, 434-42.
171. Lemmon, S. K. & Traub, L. M. (2000). Sorting in the endosomal system in yeast and animal cells. *Curr Opin Cell Biol* **12**, 457-66.
172. Levitt, M. & Park, B. H. (1993). Water: now you see it, now you don't. *Structure* **1**, 223-6.
173. Levy, Y. & Onuchic, J. N. (2006). Water mediation in protein folding and molecular recognition. *Annu Rev Biophys Biomol Struct* **35**, 389-415.
174. Li, S. S. (2005). Specificity and versatility of SH3 and other proline-recognition domains: structural basis and implications for cellular signal transduction. *Biochem J* **390**, 641-53.
175. Li, Z. & Lazaridis, T. (2003). Thermodynamic contributions of the ordered water molecule in HIV-1 protease. *J Am Chem Soc* **125**, 6636-7.
176. Li, Z. & Lazaridis, T. (2005). The effect of water displacement on binding thermodynamics: concanavalin A. *J Phys Chem B* **109**, 662-70.
177. Li, Z. & Lazaridis, T. (2006). Thermodynamics of Buried Water Clusters at a Protein-Ligand Binding Interface. *J. Phys. Chem. B* **110**, 1464-1475.
178. Li, Z. & Lazaridis, T. (2007). Water at biomolecular binding interfaces. *Phys Chem Chem Phys* **9**, 573-81.
179. Licata, J. M., Simpson-Holley, M., Wright, N. T., Han, Z., Paragas, J. & Harty, R. N. (2003). Overlapping Motifs (PTAP and PPEY) within the Ebola Virus VP40 Protein Function Independently as Late Budding Domains: Involvement of Host Proteins TSG101 and VPS-4. *J. Virol.* **77**, 1812-1819.
180. Lim, W. A. & Richards, F. M. (1994). Critical residues in an SH3 domain from Sem-5 suggest a mechanism for proline-rich peptide recognition. *Nat Struct Biol* **1**, 221-5.
181. Lim, W. A., Richards, F. M. & Fox, R. O. (1994). Structural determinants of peptide-binding orientation and of sequence specificity in SH3 domains. *Nature* **372**, 375-9.

182. Liu, Q., Berry, D., Nash, P., Pawson, T., McGlade, C. J. & Li, S. S. (2003). Structural basis for specific binding of the Gads SH3 domain to an RxxK motif-containing SLP-76 peptide: a novel mode of peptide recognition. *Mol Cell* **11**, 471-81.
183. Liu, Y. & Sturtevant, J. M. (1995). Significant discrepancies between van't Hoff and calorimetric enthalpies. II. *Protein Sci* **4**, 2559-61.
184. Lloyd, T. E., Atkinson, R., Wu, M. N., Zhou, Y., Pennetta, G. & Bellen, H. J. (2002). Hrs regulates endosome membrane invagination and tyrosine kinase receptor signaling in Drosophila. *Cell* **108**, 261-9.
185. Lo Conte, L., Chothia, C. & Janin, J. (1999). The atomic structure of protein-protein recognition sites. *J Mol Biol* **285**, 2177-98.
186. Loris, R., Langhorst, U., De Vos, S., Decanniere, K., Bouckaert, J., Maes, D., Transue, T. R. & Steyaert, J. (1999). Conserved water molecules in a large family of microbial ribonucleases. *Proteins* **36**, 117-34.
187. Lounnas, V. & Pettitt, B. M. (1994a). A connected-cluster of hydration around myoglobin: correlation between molecular dynamics simulations and experiment. *Proteins* **18**, 133-47.
188. Lounnas, V. & Pettitt, B. M. (1994b). Distribution function implied dynamics versus residence times and correlations: solvation shells of myoglobin. *Proteins* **18**, 148-60.
189. Lowry, T. M. (1935). *London, Green*.
190. Lu, Y., Wang, R., Yang, C. Y. & Wang, S. (2007). Analysis of ligand-bound water molecules in high-resolution crystal structures of protein-ligand complexes. *J Chem Inf Model* **47**, 668-75.
191. Luque, I. & Freire, E. (1998). Structure-based prediction of binding affinities and molecular design of peptide ligands. *Methods Enzymol* **295**, 100-27.
192. Luque, I. & Freire, E. (2002). Structural parameterization of the binding enthalpy of small ligands. *Proteins* **49**, 181-90.
193. Luque, I., Leavitt, S. A. & Freire, E. (2002). The linkage between protein folding and functional cooperativity: two sides of the same coin? *Annu Rev Biophys Biomol Struct* **31**, 235-56.
194. Luque, I., Todd, M. J., Gomez, J., Semo, N. & Freire, E. (1998). Molecular basis of resistance to HIV-1 protease inhibition: a plausible hypothesis. *Biochemistry* **37**, 5791-7.
195. Lybrand, T. P., Ghosh, I. & McCammon, J. A. (1985). Hydration of chloride and bromide anions: determination of relative free energy by computer simulation. *J. Am. Chem. Soc.* **107**, 7793-7794.
196. Macias, M. J., 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**, 646-9.
197. Macias, M. J., Wiesner, S. & Sudol, M. (2002). WW and SH3 domains, two different scaffolds to recognize proline-rich ligands. *FEBS Lett* **513**, 30-7.
198. Makhatadze, G. I., Gill, S. J. & Privalov, P. L. (1990). Partial Molar Heat-Capacities of the Side-Chains of Some Amino-Acid-Residues in Aqueous-Solution - the Influence of the Neighboring Charges. *Biophysical Chemistry* **38**, 33-37.
199. Makhatadze, G. I. & Privalov, P. L. (1995). Energetics of protein structure. *Adv Protein Chem* **47**, 307-425.
200. Manavalan, P. & Johnson, W. C. (1983). Sensitivity of circular dichroism to protein tertiary structure class. **305**, 831-832.

## Bibliografía

201. Mancera, R. L. (2002). De novo ligand design with explicit water molecules: an application to bacterial neuraminidase. *J Comput Aided Mol Des* **16**, 479-99.
202. Martinez, J. C. (1995). *Tesis Doctoral, Universidad de Granada*.
203. Martinez, J. C., Pisabarro, M. T. & Serrano, L. (1998). Obligatory steps in protein folding and the conformational diversity of the transition state. *Nat Struct Biol* **5**, 721-9.
204. Martinez, J. C., Viguera, A. R., Berisio, R., Wilmanns, M., Mateo, P. L., Filimonov, V. V. & 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**, 549-59.
205. Martin-Garcia, J. M., Luque, I., Mateo, P. L., Ruiz-Sanz, J. & Camara-Artigas, A. (2007). Crystallographic structure of the SH3 domain of the human c-Yes tyrosine kinase: Loop flexibility and amyloid aggregation. *FEBS Letters In Press, Corrected Proof*.
206. Martin-Serrano, J., Zang, T. & Bieniasz, P. D. (2001). HIV-1 and Ebola virus encode small peptide motifs that recruit Tsg101 to sites of particle assembly to facilitate egress. *Nat Med* **7**, 1313-9.
207. Mason, A. C. & Jensen, J. H. (2007). Protein-protein binding is often associated with changes in protonation state. *Proteins: Structure, Function, and Bioinformatics* **9999**, NA.
208. Mayer, B. J. (2001). SH3 domains: complexity in moderation. *J Cell Sci* **114**, 1253-63.
209. McCammon, J. A., Gelin, B. R. & Karplus, M. (1977). Dynamics of folded proteins. *Nature* **267**, 585-90.
210. McCammon, J. A. & Harvey, S. C. (1997). Dynamics of Proteins and Nucleic acids. *Cambridge University Press. ISBN 0521307503*.
211. McKinnon, I. R., Fall, L., Parody-Morreale, A. & Gill, S. J. (1984). A twin titration microcalorimeter for the study of biochemical reactions. *Anal Biochem* **139**, 134-9.
212. McPherson, A. (1982). *New York*.
213. McPherson, A. (1999). *Cold Spring Harbor, New York, Cold Spring Harbor Laboratory Press*.
214. Meiyappan, M., Birrane, G. & Ladas, J. A. (2007). Structural Basis for Polyproline Recognition by the FE65 WW Domain. *J Mol Biol* **372**, 970-80.
215. Meyer, E. (1992). Internal water molecules and H-bonding in biological macromolecules: a review of structural features with functional implications. *Protein Sci* **1**, 1543-62.
216. Michel, J. & Thulstrup, E. W. (1986). *New York*.
217. Minke, W. E., Diller, D. J., Hol, W. G. & Verlinde, C. L. (1999). The role of waters in docking strategies with incremental flexibility for carbohydrate derivatives: heat-labile enterotoxin, a multivalent test case. *J Med Chem* **42**, 1778-88.
218. Mongioli, A. M., Romano, P. R., Panni, S., Mendoza, M., Wong, W. T., Musacchio, A., Cesareni, G. & Di Fiore, P. P. (1999). A novel peptide-SH3 interaction. *Embo J* **18**, 5300-9.
219. Moraes, T. F., Edwards, R. A., McKenna, S., Pastushok, L., Xiao, W., Glover, J. N. & Ellison, M. J. (2001). Crystal structure of the human ubiquitin conjugating enzyme complex, hMms2-hUbc13. *Nat Struct Biol* **8**, 669-73.
220. Morales, B., Ramirez-Espain, X., Shaw, A. Z., Martin-Malpartida, P., Yraola, F., Sanchez-Tillo, E., Farrera, C., Celada, A., Royo, M. & Macias, M. J. (2007). NMR structural studies of the ItchWW3 domain reveal that phosphorylation at

- T30 inhibits the interaction with PPxY-containing ligands. *Structure* **15**, 473-83.
221. Morton, C. J., Pugh, D. J., Brown, E. L., Kahmann, J. D., Renzoni, D. A. & Campbell, I. D. (1996). Solution structure and peptide binding of the SH3 domain from human Fyn. *Structure* **4**, 705-14.
222. Mu, Y., Nordenskiold, L. & Tam, J. P. (2006). Folding, misfolding, and amyloid protofibril formation of WW domain FBP28. *Biophys J* **90**, 3983-92.
223. Murphy, K. P. & Freire, E. (1992). Thermodynamics of structural stability and cooperative folding behavior in proteins. *Adv Protein Chem* **43**, 313-61.
224. Murzin, A. G., Brenner, S. E., Hubbard, T. & Chothia, C. (1995). SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J Mol Biol* **247**, 536-40.
225. Musacchio, A., Noble, M., Pauptit, R., Wierenga, R. & Saraste, M. (1992). Crystal structure of a Src-homology 3 (SH3) domain. *Nature* **359**, 851-5.
226. Musacchio, A., Saraste, M. & Wilmanns, M. (1994). High-resolution crystal structures of tyrosine kinase SH3 domains complexed with proline-rich peptides. *Nat Struct Biol* **1**, 546-51.
227. Nagar, B., Hantschel, O., Seeliger, M., Davies, J. M., Weis, W. I., Superti-Furga, G. & Kuriyan, J. (2006). Organization of the SH3-SH2 unit in active and inactive forms of the c-Abl tyrosine kinase. *Mol Cell* **21**, 787-98.
228. Nagar, B., Hantschel, O., Young, M. A., Scheffzek, K., Veach, D., Bornmann, W., Clarkson, B., Superti-Furga, G. & Kuriyan, J. (2003). Structural basis for the autoinhibition of c-Abl tyrosine kinase. *Cell* **112**, 859-71.
229. Naghibi, H., Tamura, A. & Sturtevant, J. M. (1995). Significant discrepancies between van't Hoff and calorimetric enthalpies. *Proc Natl Acad Sci U S A* **92**, 5597-9.
230. Nakasako, M. (1999). Large-scale networks of hydration water molecules around bovine beta-trypsin revealed by cryogenic X-ray crystal structure analysis. *J Mol Biol* **289**, 547-64.
231. Neduva, V., Linding, R., Su-Angrand, I., Stark, A., de Masi, F., Gibson, T. J., Lewis, J., Serrano, L. & Russell, R. B. (2005). Systematic discovery of new recognition peptides mediating protein interaction networks. *PLoS Biol* **3**, e405.
232. Ng, J. D., Gavira, J. A. & Garcia-Ruiz, J. M. (2003). Protein crystallization by capillary counterdiffusion for applied crystallographic structure determination. *J Struct Biol* **142**, 218-31.
233. Nguyen, J. T., Porter, M., Amoui, M., Miller, W. T., Zuckermann, R. N. & Lim, W. A. (2000). Improving SH3 domain ligand selectivity using a non-natural scaffold. *Chem Biol* **7**, 463-73.
234. Nguyen, J. T., Turck, C. W., Cohen, F. E., Zuckermann, R. N. & Lim, W. A. (1998). Exploiting the basis of proline recognition by SH3 and WW domains: design of N-substituted inhibitors. *Science* **282**, 2088-92.
235. Noble, M. E., Musacchio, A., Saraste, M., Courtneidge, S. A. & Wierenga, R. K. (1993). Crystal structure of the SH3 domain in human Fyn; comparison of the three-dimensional structures of SH3 domains in tyrosine kinases and spectrin. *Embo J* **12**, 2617-24.
236. Ohnishi, S., Guntert, P., Koshiha, S., Tomizawa, T., Akasaka, R., Tochio, N., Sato, M., Inoue, M., Harada, T., Watanabe, S., Tanaka, A., Shirouzu, M., Kigawa, T. & Yokoyama, S. (2007). Solution structure of an atypical WW domain in a novel beta-clam-like dimeric form. *FEBS Lett* **581**, 462-8.

## Bibliografía

237. Ohtaka, H., Muzammil, S., Schon, A., Velazquez-Campoy, A., Vega, S. & Freire, E. (2004). Thermodynamic rules for the design of high affinity HIV-1 protease inhibitors with adaptability to mutations and high selectivity towards unwanted targets. *Int J Biochem Cell Biol* **36**, 1787-99.
238. O'Leary, J. M., Bromek, K., Black, G. M., Uhrinova, S., Schmitz, C., Wang, X., Krych, M., Atkinson, J. P., Uhrin, D. & Barlow, P. N. (2004). Backbone dynamics of complement control protein (CCP) modules reveals mobility in binding surfaces. *Protein Sci* **13**, 1238-50.
239. Oneyama, C., Nakano, H. & Sharma, S. V. (2002). UCS15A, a novel small molecule, SH3 domain-mediated protein-protein interaction blocking drug. *Oncogene* **21**, 2037-50.
240. Ortega Roldan, J. L., Romero Romero, M. L., Ora, A., Ab, E., Lopez Mayorga, O., Azuaga, A. I. & van Nuland, N. A. (2007). The high resolution NMR structure of the third SH3 domain of CD2AP. *J Biomol NMR* **39**, 331-6.
241. Otte, L., Wiedemann, U., Schlegel, B., Pires, J. R., 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.
242. Otting, G., Liepinsh, E., Farmer, B. T., 2nd & Wuthrich, K. (1991a). Protein hydration studied with homonuclear 3D <sup>1</sup>H NMR experiments. *J Biomol NMR* **1**, 209-15.
243. Otting, G., Liepinsh, E. & Wuthrich, K. (1991b). Protein hydration in aqueous solution. *Science* **254**, 974-80.
244. Pal, S. K., Peon, J., Bagchi, B. & Zewail, A. H. (2002). Biological Water: Femtosecond Dynamics of Macromolecular Hydration. *J. Phys. Chem. B* **106**, 12376-12395.
245. Pal, S. K., Peon, J. & Zewail, A. H. (2002). Biological water at the protein surface: dynamical solvation probed directly with femtosecond resolution. *Proc Natl Acad Sci U S A* **99**, 1763-8.
246. Palencia, A., Cobos, E. S., Mateo, P. L., Martinez, J. C. & 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**, 527-37.
247. Palencia, A., Martinez, J. C., Mateo, P. L., Luque, I. & Camara-Artigas, A. (2006). Structure of human TSG101 UEV domain. *Acta Crystallogr D Biol Crystallogr* **62**, 458-64.
248. Panni, S., Dente, L. & Cesareni, G. (2002). In vitro evolution of recognition specificity mediated by SH3 domains reveals target recognition rules. *J Biol Chem* **277**, 21666-74.
249. Park, S. & Saven, J. G. (2005). Statistical and molecular dynamics studies of buried waters in globular proteins. *Proteins* **60**, 450-63.
250. Paschek, D., Nymeyer, H. & Garcia, A. E. (2007). Replica exchange simulation of reversible folding/unfolding of the Trp-cage miniprotein in explicit solvent: on the structure and possible role of internal water. *J Struct Biol* **157**, 524-33.
251. Pawson, T. (1995). Protein modules and signalling networks. *Nature* **373**, 573-80.
252. Pawson, T. (2004). Specificity in Signal Transduction: From Phosphotyrosine-SH2 Domain Interactions to Complex Cellular Systems. *Cell* **116**, 191-203.

253. Pawson, T., Raina, M. & Nash, P. (2002). Interaction domains: from simple binding events to complex cellular behavior. *FEBS Lett* **513**, 2-10.
254. Pawson, T. & Scott, J. D. (1997). Signaling through scaffold, anchoring, and adaptor proteins. *Science* **278**, 2075-80.
255. Peranen, J., Rikkinen, M., Hyvonen, M. & Kaariainen, L. (1996). T7 vectors with modified T7lac promoter for expression of proteins in Escherichia coli. *Anal Biochem* **236**, 371-3.
256. Perl, D., Welker, C., Schindler, T., Schroder, K., Marahiel, M. A., Jaenicke, R. & Schmid, F. X. (1998). Conservation of rapid two-state folding in mesophilic, thermophilic and hyperthermophilic cold shock proteins. *Nat Struct Biol* **5**, 229-35.
257. Permyakov, E. A. (1993). *Luminiscent Spectroscopy of Proteins*, CRC Press, Boca Raton (Florida, USA).
258. Petrone, P. M. & Garcia, A. E. (2004). MHC-Peptide Binding is Assisted by Bound Water Molecules. *Journal of Molecular Biology* **338**, 419-435.
259. Petrovich, M., Jonsson, A. L., Ferguson, N., Daggett, V. & Fersht, A. R. (2006). Phi-analysis at the experimental limits: mechanism of beta-hairpin formation. *J Mol Biol* **360**, 865-81.
260. Pickart, C. M. (2004). Back to the Future with Ubiquitin. *Cell* **116**, 181-190.
261. Pires, J. R., 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.
262. Pires, J. R., Taha-Nejad, F., Toepert, F., Ast, T., Hoffmuller, U., Schneider-Mergener, J., Kuhne, R., Macias, M. J. & Oschkinat, H. (2001). Solution structures of the YAP65 WW domain and the variant L30 K in complex with the peptides GTPPPYTVG, N-(n-octyl)-GPPPY and PLPPY and the application of peptide libraries reveal a minimal binding epitope. *J Mol Biol* **314**, 1147-56.
263. Pisabarro, M. T., Ortiz, A. R., Viguera, A. R., Gago, F. & Serrano, L. (1994). Molecular modeling of the interaction of polyproline-based peptides with the Abl-SH3 domain: rational modification of the interaction. *Protein Eng* **7**, 1455-62.
264. Pisabarro, M. T. & Serrano, L. (1996). Rational design of specific high-affinity peptide ligands for the Abl-SH3 domain. *Biochemistry* **35**, 10634-40.
265. Pisabarro, M. T., Serrano, L. & Wilmanns, M. (1998). Crystal structure of the abl-SH3 domain complexed with a designed high-affinity peptide ligand: implications for SH3-ligand interactions. *J Mol Biol* **281**, 513-21.
266. Poornima, C. S. & Dean, P. M. (1995a). Hydration in drug design. 2. Influence of local site surface shape on water binding. *J Comput Aided Mol Des* **9**, 513-20.
267. Poornima, C. S. & Dean, P. M. (1995b). Hydration in drug design. 3. Conserved water molecules at the ligand-binding sites of homologous proteins. *J Comput Aided Mol Des* **9**, 521-31.
268. Pornillos, O., Alam, S. L., Davis, D. R. & Sundquist, W. I. (2002). Structure of the Tsg101 UEV domain in complex with the PTAP motif of the HIV-1 p6 protein. *Nat Struct Biol* **9**, 812-7.
269. Pornillos, O., Alam, S. L., Rich, R. L., Myszka, D. G., Davis, D. R. & Sundquist, W. I. (2002a). Structure and functional interactions of the Tsg101 UEV domain. *EMBO J.* **21**, 2397-2406.
270. Pornillos, O., Garrus, J. E. & Sundquist, W. I. (2002b). Mechanisms of enveloped RNA virus budding. *Trends in Cell Biology* **12**, 569-579.

## Bibliografía

271. Pornillos, O., Higginson, D. S., Stray, K. M., Fisher, R. D., Garrus, J. E., Payne, M., He, G. P., Wang, H. E., Morham, S. G. & Sundquist, W. I. (2003). HIV Gag mimics the Tsg101-recruiting activity of the human Hrs protein. *J Cell Biol* **162**, 425-34.
272. Privalov, P. L. (1979). Stability of proteins: small globular proteins. *Adv Protein Chem* **33**, 167-241.
273. Privalov, P. L. & Potekhin, S. A. (1986). Scanning microcalorimetry in studying temperature-induced changes in proteins. *Methods Enzymol* **131**, 4-51.
274. Rarey, M., Kramer, B. & Lengauer, T. (1999). The particle concept: placing discrete water molecules during protein-ligand docking predictions. *Proteins* **34**, 17-28.
275. Raschke, T. M. (2006). Water structure and interactions with protein surfaces. *Curr Opin Struct Biol* **16**, 152-9.
276. Raymer, M. L., Sanschagrin, P. C., Punch, W. F., Venkataraman, S., Goodman, E. D. & Kuhn, L. A. (1997). Predicting conserved water-mediated and polar ligand interactions in proteins using a K-nearest-neighbors genetic algorithm. *J Mol Biol* **265**, 445-64.
277. Reddy, C. K., Das, A. & Jayaram, B. (2001). Do water molecules mediate protein-DNA recognition? *J Mol Biol* **314**, 619-32.
278. Reiss, D. J. & Schwikowski, B. (2004). Predicting protein-peptide interactions via a network-based motif sampler. *Bioinformatics* **20 Suppl 1**, I274-I282.
279. Ren, R., Mayer, B. J., Cicchetti, P. & Baltimore, D. (1993). Identification of a ten-amino acid proline-rich SH3 binding site. *Science* **259**, 1157-61.
280. Renzoni, D. A., Pugh, D. J., Siligardi, G., Das, P., Morton, C. J., Rossi, C., Waterfield, M. D., Campbell, I. D. & Ladbury, J. E. (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**, 15646-53.
281. Richardson, J. S., Bryan, W. A., 3rd & Richardson, D. C. (2003). New tools and data for improving structures, using all-atom contacts. *Methods Enzymol* **374**, 385-412.
282. Rodier, F., Bahadur, R. P., Chakrabarti, P. & Janin, J. (2005). Hydration of protein-protein interfaces. *Proteins* **60**, 36-45.
283. Rubin, G. M., Yandell, M. D., Wortman, J. R., Gabor Miklos, G. L., Nelson, C. R., Hariharan, I. K., Fortini, M. E., Li, P. W., Apweiler, R., Fleischmann, W., Cherry, J. M., Henikoff, S., Skupski, M. P., Misra, S., Ashburner, M., Birney, E., Boguski, M. S., Brody, T., Brokstein, P., Celniker, S. E., Chervitz, S. A., Coates, D., Cravchik, A., Gabrielian, A., Galle, R. F., Gelbart, W. M., George, R. A., Goldstein, L. S., Gong, F., Guan, P., Harris, N. L., Hay, B. A., Hoskins, R. A., Li, J., Li, Z., Hynes, R. O., Jones, S. J., Kuehl, P. M., Lemaitre, B., Littleton, J. T., Morrison, D. K., Mungall, C., O'Farrell, P. H., Pickeral, O. K., Shue, C., Vossell, L. B., Zhang, J., Zhao, Q., Zheng, X. H. & Lewis, S. (2000). Comparative genomics of the eukaryotes. *Science* **287**, 2204-15.
284. Ruiz-Sanz, J., Simoncsits, A., Toro, I., Pongor, S., Mateo, P. L. & Filimonov, V. V. (1999). A thermodynamic study of the 434-repressor N-terminal domain and of its covalently linked dimers. *Eur J Biochem* **263**, 246-53.
285. Ryckaert, J. P., Ciccotti, G. & Berendsen, H. J. C. (1977). *J. Comput. Phys.*, **23**, 327.
286. Sadasivan, C., Nagendra, H. G. & Vijayan, M. (1998). Plasticity, hydration and accessibility in ribonuclease A. The structure of a new crystal form and its low-humidity variant. *Acta Crystallogr D Biol Crystallogr* **54**, 1343-52.

287. Sadqi, M., Casares, S., Abril, M. A., Lopez-Mayorga, O., Conejero-Lara, F. & Freire, E. (1999). The native state conformational ensemble of the SH3 domain from alpha-spectrin. *Biochemistry* **38**, 8899-906.
288. Sadqi, M., Casares, S., Lopez-Mayorga, O., Martinez, J. C. & Conejero-Lara, F. (2002). pH dependence of the hydrogen exchange in the SH3 domain of alpha-spectrin. *FEBS Lett* **514**, 295-9.
289. 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 and Infection* **6**, 150-156.
290. Sanchez-Ruiz, J. M. (1995). Differential scanning calorimetry of proteins. *Subcell Biochem* **24**, 133-76.
291. Schleinkofer, K., Wiedemann, U., Otte, L., Wang, T., Krause, G., Oschkinat, H. & Wade, R. C. (2004). Comparative structural and energetic analysis of WW domain-peptide interactions. *J Mol Biol* **344**, 865-81.
292. Schoepfer, R. (1993). The pRSET family of T7 promoter expression vectors for Escherichia Coli. *Gene* **124**, 83-85.
293. Schumacher, T. N., Mayr, L. M., Minor, D. L., Jr., Milhollen, M. A., Burgess, M. W. & Kim, P. S. (1996). Identification of D-peptide ligands through mirror-image phage display. *Science* **271**, 1854-7.
294. Schweimer, K., Hoffmann, S., Bauer, F., Friedrich, U., Kardinal, C., Feller, S. M., Biesinger, B. & Sticht, H. (2002). Structural investigation of the binding of a herpesviral protein to the SH3 domain of tyrosine kinase Lck. *Biochemistry* **41**, 5120-30.
295. Sharrow, S. D., Edmonds, K. A., Goodman, M. A., Novotny, M. V. & Stone, M. J. (2005). Thermodynamic consequences of disrupting a water-mediated hydrogen bond network in a protein:pheromone complex. *Protein Sci* **14**, 249-56.
296. Sicheri, F., Moarefi, I. & Kuriyan, J. (1997). Crystal structure of the Src family tyrosine kinase Hck. *Nature* **385**, 602-9.
297. Sigurskjold, B. W. (2000). Exact analysis of competition ligand binding by displacement isothermal titration calorimetry. *Anal Biochem* **277**, 260-6.
298. Sigurskjold, B. W., Berland, C. R. & Svensson, B. (1994). Thermodynamics of inhibitor binding to the catalytic site of glucoamylase from *Aspergillus niger* determined by displacement titration calorimetry. *Biochemistry* **33**, 10191-9.
299. 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**, 406-18.
300. Sleigh, S. H., Seavers, P. R., Wilkinson, A. J., Ladbury, J. E. & Tame, J. R. H. (1999). Crystallographic and Calorimetric Analysis of Peptide Binding to OppA Protein. *Journal of Molecular Biology* **291**, 393-415.
301. Smith, J. D., Cappa, C. D., Wilson, K. R., Messer, B. M., Cohen, R. C. & Saykally, R. J. (2004). Energetics of Hydrogen Bond Network Rearrangements in Liquid Water. *Science* **306**, 851-853.
302. Sparks, A. B., Rider, J. E., Hoffman, N. G., Fowlkes, D. M., Quillam, L. A. & Kay, B. K. (1996). Distinct ligand preferences of Src homology 3 domains from Src, Yes, Abl, Cortactin, p53bp2, PLCgamma, Crk, and Grb2. *Proc Natl Acad Sci U S A* **93**, 1540-4.
303. Spolar, R. S., Livingstone, J. R. & Record, M. T., Jr. (1992). Use of liquid hydrocarbon and amide transfer data to estimate contributions to

## Bibliografía

- thermodynamic functions of protein folding from the removal of nonpolar and polar surface from water. *Biochemistry* **31**, 3947-55.
304. Stapley, B. J. & Creamer, T. P. (1999). A survey of left-handed polyproline II helices. *Protein Sci* **8**, 587-95.
305. Staub, O., Dho, S., Henry, P., Correa, J., Ishikawa, T., McGlade, J. & Rotin, D. (1996). WW domains of Nedd4 bind to the proline-rich PY motifs in the epithelial Na<sup>+</sup> channel deleted in Liddle's syndrome. *EMBO J.* **15**, 2371-2380.
306. Stauffer, T. P., Martenson, C. H., Rider, J. E., Kay, B. K. & Meyer, T. (1997). Inhibition of Lyn function in mast cell activation by SH3 domain binding peptides. *Biochemistry* **36**, 9388-94.
307. Still, W. C., Tempczyk, A., Hawley, R. C. & Hendrickson, T. (1990). Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* **112**, 6127-6129.
308. Strack, B., Calistri, A., Accola, M. A., Palu, G. & Gottlinger, H. G. (2000). A role for ubiquitin ligase recruitment in retrovirus release. *PNAS* **97**, 13063-13068.
309. Sturtevant, J. M. (1977). Heat capacity and entropy changes in processes involving proteins. *Proc Natl Acad Sci U S A* **74**, 2236-40.
310. 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**, 2145-52.
311. Sudol, M. (1998). From Src Homology domains to other signaling modules: proposal of the 'protein recognition code'. *Oncogene* **17**, 1469-74.
312. Sudol, M. & Hunter, T. (2000). NeW wrinkles for an old domain. *Cell* **103**, 1001-4.
313. Sudol, M., Sliwa, K. & Russo, T. (2001). Functions of WW domains in the nucleus. *FEBS Lett* **490**, 190-5.
314. Sundquist, W. I., Schubert, H. L., Kelly, B. N., Hill, G. C., Holton, J. M. & Hill, C. P. (2004). Ubiquitin Recognition by the Human TSG101 Protein. *Molecular Cell* **13**, 783-789.
315. Tame, J. R., Murshudov, G. N., Dodson, E. J., Neil, T. K., Dodson, G. G., Higgins, C. F. & Wilkinson, A. J. (1994). The structural basis of sequence-independent peptide binding by OppA protein. *Science* **264**, 1578-81.
316. Tame, J. R., Sleigh, S. H., Wilkinson, A. J. & Ladbury, J. E. (1996). The role of water in sequence-independent ligand binding by an oligopeptide transporter protein. *Nat Struct Biol* **3**, 998-1001.
317. Tamm, I. E. (1979). Fundamentals of the Theory of electricity; Mir Publishers: Moscow.
318. Tarek, M. & Tobias, D. J. (2002). Role of protein-water hydrogen bond dynamics in the protein dynamical transition. *Phys Rev Lett* **88**, 138101.
319. Teo, H., Veprintsev, D. B. & Williams, R. L. (2004). Structural insights into endosomal sorting complex required for transport (ESCRT-I) recognition of ubiquitinated proteins. *J Biol Chem* **279**, 28689-96.
320. Teyra, J., Doms, A., Schroeder, M. & Pisabarro, M. T. (2006). SCOWLP: a web-based database for detailed characterization and visualization of protein interfaces. *BMC Bioinformatics* **7**, 104.
321. Teyra, J. & Teresa Pisabarro, M. (2007). Characterization of interfacial solvent in protein complexes and contribution of wet spots to the interface description. *Proteins*.
322. Timmins, J., Schoehn, G., Ricard-Blum, S., Scianimanico, S., Vernet, T., Ruigrok, R. W. & Weissenhorn, W. (2003). Ebola virus matrix protein VP40

- interaction with human cellular factors Tsg101 and Nedd4. *J Mol Biol* **326**, 493-502.
323. Todd, M. J., Luque, I., Velazquez-Campoy, A. & Freire, E. (2000). Thermodynamic basis of resistance to HIV-1 protease inhibition: calorimetric analysis of the V82F/I84V active site resistant mutant. *Biochemistry* **39**, 11876-83.
324. Tomasi, J. & Persico, M. (1994). *M. Chem. Rev.*, **94**, 2007.
325. Tong, A. H., Drees, B., Nardelli, G., Bader, G. D., Brannetti, B., Castagnoli, L., Evangelista, M., Ferracuti, S., Nelson, B., Paoluzi, S., Quondam, M., Zucconi, A., Hogue, C. W., Fields, S., Boone, C. & Cesareni, G. (2002). A combined experimental and computational strategy to define protein interaction networks for peptide recognition modules. *Science* **295**, 321-4.
326. VanDemark, A. P., Hofmann, R. M., Tsui, C., Pickart, C. M. & Wolberger, C. (2001). Molecular insights into polyubiquitin chain assembly: crystal structure of the Mms2/Ubc13 heterodimer. *Cell* **105**, 711-20.
327. Vega, S., Kang, L. W., Velazquez-Campoy, A., Kiso, Y., Amzel, L. M. & Freire, E. (2004). A structural and thermodynamic escape mechanism from a drug resistant mutation of the HIV-1 protease. *Proteins* **55**, 594-602.
328. Velazquez Campoy, A. & Freire, E. (2005). ITC in the post-genomic era...? Priceless. *Biophys Chem* **115**, 115-24.
329. Velazquez-Campoy, A. & Freire, E. (2006). Isothermal titration calorimetry to determine association constants for high-affinity ligands. *Nat Protoc* **1**, 186-91.
330. Velazquez-Campoy, A., Freire, E. & Kiso, Y. (2001). Incorporating target heterogeneity in drug design
- The binding energetics of first- and second-generation HIV-1 protease inhibitors: implications for drug design. *J Cell Biochem Suppl Suppl* **37**, 82-8.
331. Velazquez-Campoy, A., Leavitt, S. A. & Freire, E. (2004). Characterization of protein-protein interactions by isothermal titration calorimetry. *Methods Mol Biol* **261**, 35-54.
332. Velazquez-Campoy, A., Luque, I., Todd, M. J., Milutinovich, M., Kiso, Y. & Freire, E. (2000a). Thermodynamic dissection of the binding energetics of KNI-272, a potent HIV-1 protease inhibitor. *Protein Sci* **9**, 1801-9.
333. Velazquez-Campoy, A., Todd, M. J. & Freire, E. (2000b). HIV-1 protease inhibitors: enthalpic versus entropic optimization of the binding affinity. *Biochemistry* **39**, 2201-7.
334. Verdecia, M. A., Bowman, M. E., Lu, K. P., Hunter, T. & Noel, J. P. (2000). Structural basis for phosphoserine-proline recognition by group IV WW domains. *Nat Struct Biol* **7**, 639-43.
335. Verdonk, M. L., Cole, J. C. & Taylor, R. (1999). SuperStar: a knowledge-based approach for identifying interaction sites in proteins. *J Mol Biol* **289**, 1093-108.
336. Verlet, L. (1967). *Phys. Rev.* **159**, 98 (1967); *Phys. Rev.* **165**, 201 (1967).
337. Vidal, M., Gigoux, V. & Garbay, C. (2001). SH2 and SH3 domains as targets for anti-proliferative agents. *Crit Rev Oncol Hematol* **40**, 175-86.
338. Viguera, A. R., Arrondo, J. L., Musacchio, A., Saraste, M. & Serrano, L. (1994). Characterization of the interaction of natural proline-rich peptides with five different SH3 domains. *Biochemistry* **33**, 10925-33.
339. Viguera, A. R., Blanco, F. J. & Serrano, L. (1995). The order of secondary structure elements does not determine the structure of a protein but does affect its folding kinetics. *J Mol Biol* **247**, 670-81.
340. Viguera, A. R., Serrano, L. & Wilmanns, M. (1996). Different folding transition states may result in the same native structure. *Nat Struct Biol* **3**, 874-80.

## Bibliografia

341. Vijayakumar, M., Qian, H. & Zhou, H. X. (1999). Hydrogen bonds between short polar side chains and peptide backbone: prevalence in proteins and effects on helix-forming propensities. *Proteins* **34**, 497-507.
342. Wang, C., Pawley, N. H. & Nicholson, L. K. (2001). The role of backbone motions in ligand binding to the c-Src SH3 domain. *J Mol Biol* **313**, 873-87.
343. Wang, Y. X., Freedberg, D. I., Grzesiek, S., Torchia, D. A., Wingfield, P. T., Kaufman, J. D., Stahl, S. J., Chang, C. H. & Hodge, C. N. (1996a). Mapping hydration water molecules in the HIV-1 protease/DMP323 complex in solution by NMR spectroscopy. *Biochemistry* **35**, 12694-704.
344. Wang, Y. X., Freedberg, D. I., Yamazaki, T., Wingfield, P. T., Stahl, S. J., Kaufman, J. D., Kiso, Y. & Torchia, D. A. (1996b). Solution NMR evidence that the HIV-1 protease catalytic aspartyl groups have different ionization states in the complex formed with the asymmetric drug KNI-272. *Biochemistry* **35**, 9945-50.
345. Watson, K. A., Mitchell, E. P., Johnson, L. N., Son, J. C., Bichard, C. J., Orchard, M. G., Fleet, G. W., Oikonomakos, N. G., Leonidas, D. D., Kontou, M. & et al. (1994). Design of inhibitors of glycogen phosphorylase: a study of alpha- and beta-C-glucosides and 1-thio-beta-D-glucose compounds. *Biochemistry* **33**, 5745-58.
346. Weber, P. C. (1991). *Adv. Prot. Chem.*
347. Weber, P. C., Pantoliano, M. W., Simons, D. M. & Salemme, F. R. (1994). Structure-Based Design of Synthetic Azobenzene Ligands for Streptavidin. *J. Am. Chem. Soc.* **116**, 2717-2724.
348. Weinryb, I. S., R.F. (1971). The Luminescence of aromatic amino acids. In *Excited States of Proteins and Nucleic Acids* (Weinryb, R. F. S. a. I., ed.), pp. 277-318. Plenum Press, New York (USA).
349. Wernet, P., Nordlund, D., Bergmann, U., Cavalleri, M., Odelius, M., Ogasawara, H., Naslund, L. A., Hirsch, T. K., Ojamae, L., Glatzel, P., Pettersson, L. G. & Nilsson, A. (2004). The structure of the first coordination shell in liquid water. *Science* **304**, 995-9.
350. Williams, J. C., Weijland, A., Gonfloni, S., Thompson, A., Courtneidge, S. A., Superti-Furga, G. & Wierenga, R. K. (1997). The 2.35 Å crystal structure of the inactivated form of chicken Src: a dynamic molecule with multiple regulatory interactions. *J Mol Biol* **274**, 757-75.
351. Williams, M. A., Goodfellow, J. M. & Thornton, J. M. (1994). Buried waters and internal cavities in monomeric proteins. *Protein Sci* **3**, 1224-35.
352. Winn, M. D., Isupov, M. N. & Murshudov, G. N. (2001). Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr D Biol Crystallogr* **57**, 122-33.
353. Wiseman, T., Williston, S., Brandts, J. F. & Lin, L. N. (1989). Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal Biochem* **179**, 131-7.
354. Wittekind, M., Mapelli, C., Farmer, B. T., 2nd, Suen, K. L., Goldfarb, V., Tsao, J., Lavoie, T., Barbacid, M., Meyers, C. A. & Mueller, L. (1994). Orientation of peptide fragments from Sos proteins bound to the N-terminal SH3 domain of Grb2 determined by NMR spectroscopy. *Biochemistry* **33**, 13531-9.
355. Woody, R. W., Sugeta, H. & Kodama, T. S. (1996). [Circular dichroism of proteins: recent developments in analysis and prediction]. *Tanpakushitsu Kakusan Koso* **41**, 56-69.

356. Wu, J., Yang, J. T. & Wu, C. S. (1992). Beta-II conformation of all-beta proteins can be distinguished from unordered form by circular dichroism. *Anal Biochem* **200**, 359-64.
357. Wu, X., Knudsen, B., Feller, S. M., Zheng, J., Sali, A., Cowburn, D., Hanafusa, H. & Kuriyan, J. (1995). Structural basis for the specific interaction of lysine-containing proline-rich peptides with the N-terminal SH3 domain of c-Crk. *Structure* **3**, 215-26.
358. Xu, W., Doshi, A., Lei, M., Eck, M. J. & Harrison, S. C. (1999). Crystal structures of c-Src reveal features of its autoinhibitory mechanism. *Mol Cell* **3**, 629-38.
359. Xu, W., Harrison, S. C. & Eck, M. J. (1997). Three-dimensional structure of the tyrosine kinase c-Src. *Nature* **385**, 595-602.
360. 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.
361. Yasuda, J., Nakao, M., Kawaoka, Y. & Shida, H. (2003). Nedd4 regulates egress of Ebola virus-like particles from host cells. *J Virol* **77**, 9987-92.
362. Yu, H., Chen, J. K., Feng, S., Dalgarno, D. C., Brauer, A. W. & Schreiber, S. L. (1994). Structural basis for the binding of proline-rich peptides to SH3 domains. *Cell* **76**, 933-45.
363. Yu, H., Rosen, M. K., Shin, T. B., Seidel-Dugan, C., Brugge, J. S. & Schreiber, S. L. (1992). Solution structure of the SH3 domain of Src and identification of its ligand-binding site. *Science* **258**, 1665-8.
364. Zarrinpar, A., Bhattacharyya, R. P. & Lim, W. A. (2003a). The structure and function of proline recognition domains. *Sci STKE* **2003**, RE8.
365. Zarrinpar, A. & Lim, W. A. (2000). Converging on proline: the mechanism of WW domain peptide recognition. *Nat Struct Biol* **7**, 611-3.
366. Zarrinpar, A., Park, S. H. & Lim, W. A. (2003b). Optimization of specificity in a cellular protein interaction network by negative selection. *Nature* **426**, 676-80.
367. Zauhar, R. J. & Morgan, R. S. (1985). A new method for computing the macromolecular electric potential. *Journal of Molecular Biology* **186**, 815-820.
368. Zhang, L. Y., Gallicchio, E., Friesner, R. A. & Levy, R. M. (2001). Solvent models for protein-ligand binding: Comparison of implicit solvent poisson and surface generalized born models with explicit solvent simulations. *Journal of Computational Chemistry* **22**, 591-607.
369. Zhang, Y. L. & Zhang, Z. Y. (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**, 139-48.

