

UNIVERSIDAD DE GRANADA

FACULTAD DE CIENCIAS



DEPARTAMENTO DE QUÍMICA FÍSICA

**DISEÑO Y CARACTERIZACIÓN DE PROTEÍNAS QUIMÉRICAS
PARA EL ESTUDIO DEL RECONOCIMIENTO MOLECULAR
ENTRE LIGANDOS PEPTÍDICOS Y DOMINIOS SH3**

ADELA M^a CANDEL RAMÓN

**Tesis Doctoral
GRANADA
2008**

Editor: Editorial de la Universidad de Granada
Autor: Adela María Candell Ramón
D.L.: Gr. 133-2008
ISBN: 978-84-338-4752-2

UNIVERSIDAD DE GRANADA
FACULTAD DE CIENCIAS
DEPARTAMENTO DE QUÍMICA FÍSICA

Memoria presentada para aspirar al grado de Doctor en Química

Fdo: **Adela M^a Candel Ramón**

Licenciada en Química por la Universidad de Granada

Granada, a 21 de Enero de 2008

V^oB^o
DIRECTORES DE LA TESIS

Fdo: **Francisco Conejero Lara**
Profesor Titular
Departamento de Química Física
Universidad de Granada

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

“Una vez alguien me dijo que los peces no tienen memoria, que en apenas unos segundos olvidan lo que han vivido momentos antes. La memoria de hombres y mujeres me recuerda a menudo a la de los peces, hombres y mujeres que olvidan su historia, lo que han sentido, hombres y mujeres con amnesia abocados a repetir los mismos errores...”

(Ismael Serrano)

A mis padres
A mis hermanos

AGRADECIMIENTOS

Como es sabido por todos aquellos que han realizado y escrito una Tesis, no es tarea fácil, por lo que en este momento en el que parece que está llegando el final, no puedo olvidar dedicar unas palabras a todos y cada uno de los que en cierta medida, profesional o personal, han contribuido a ello.

En primer lugar, quiero agradecer a Pedro el haberme permitido realizar esta Tesis en su grupo de investigación.

A mis dos directores de Tesis, Quico y Jose, por su dedicación, por transmitirme su pasión por la ciencia, y porque ellos son los responsables de que yo esté aquí ahora, ya que sin su apoyo y su confianza, nada de esto hubiera sido posible. Gracias a ambos por darme la oportunidad de trabajar con vosotros y por dejarme descubrir cómo me gusta este trabajo.

A Nico por su apoyo incondicional, por transmitirme su pasión por el RMN, por haberme abierto las puertas de su casa y haberme brindado su amistad desde el primer momento.

A Anabel, porque me ha demostrado su cariño y su amistad día a día, y porque siempre ha sabido escucharme y aconsejarme cuando lo he necesitado, dejando de un lado lo que estaba haciendo para dedicarme su tiempo.

A Salva ("*mi currito*"), me faltan las palabras para agradecerle todo lo que ha hecho por mí. Gracias por todo lo que me has enseñado, por tu paciencia y sobre todo por tu amistad.

A Eva, mi compañera de despacho durante esta última etapa. Gracias por nuestras largas conversaciones y consejos, por tu ayuda y por tu tiempo.

A M^a del Mar, por ser tu, porque eres única, me alegro de haberte conocido.

A Irene, por sus sabios consejos; a Javi, porque es el primero que está cuando surge cualquier problema; a Abdulio, por compartir conmigo sus conocimientos; a Enrique, por ser mi tutor de doctorado y por solucionar siempre todos los problemas burocráticos; a Beatriz, por su apoyo y por sus ánimos... y al resto de profesores del departamento, por toda la ayuda prestada...

A Ana Cámara y a M^a Paz, que aunque no pertenecen a este Departamento, siempre han mostrado mucho interés, preocupándose en todo momento por mí.

A mis compañeros de laboratorio, Andrés, Jose Manuel, Bertrand, Manu, Lorena, Jose Luis, Sisi, Javi Murciano y Carles, porque sin ellos no habría llegado hasta aquí. Gracias por todos los momentos que hemos pasado, porque siempre habéis sabido sacarme una sonrisa cuando lo necesitaba, escucharme, ayudarme y sobre todo en muchos casos demostrarme que no sois solo compañeros de trabajo, sino amigos dentro y fuera del laboratorio. A Andrés, porque desde que nos conocimos cuando empezamos a estudiar Química, has estado ahí, creo que sobran las palabras entre nosotros; a Jose Manuel, por todas las risas y salidas nocturnas que hemos compartido, ¡que hubiera hecho sin ti!; a Bertrand, por todo lo que me ha ayudado, siempre has estado cuando te he necesitado; a Manu, por su interés en todo lo que he

hecho; a Lorena ("*mi niña*"), por nuestras conversaciones y confidencias y sobre todo por tu amistad, porque desde que nos conocimos conectamos de una forma especial; a Jose Luis ("*mi compi* *erremenero*"), por su ayuda y apoyo en todo momento, y por ser mi amigo; a Sisi, por hacerme reír; a Javi Murciano y a Carles, porque aunque no he tenido mucho tiempo de conocerlos todavía, se que estáis ahí para lo que necesite.

A todos mis compañeros del grupo vecino: David, Héctor, Asun, Rocío e Inma, por vuestra ayuda y por ser compañeros de juergas; y a los que ya no están, como Raúl, por todos los ratos y conversaciones que hemos compartido juntos, y porque aún en la distancia me ha demostrado su amistad, que estoy segura de que durará mucho tiempo.

A Jose Miguel, por toda la ayuda recibida en lo que ha papeleo se refiere y por las largas conversaciones mantenidas en la secretaria. A Araceli, por haber hecho el trabajo diario más fácil gracias a su apoyo.

A toda la gente del grupo de Espectroscopia y Estructura Molecular del Instituto de Química Física "Rocasolano" (CSIC) de Madrid, donde he tenido la oportunidad de realizar dos estancias, en especial a Marta, por todo lo que me ha enseñado a nivel profesional y personal, por sus consejos y por su apoyo. A Manolo, por haberme dado la oportunidad de pertenecer a su "*gran familia erremenera*" durante un tiempo. A Jorge, M^a Ángeles, Carlos, Douglas, Miguel Treviño, Jose Manuel, Clara y Cristina, por toda la ayuda prestada. A Anjos, parte del "*gineceo*", junto con Inés y Yolanda, por los buenos momentos pasados en el "*chalecito*". A mis compañeros becarios, Inés, Yolanda, Jorge Pedro, Miguel Ángel, Jaime, Esther y Yasmina, por haberme hecho sentir uno más y haberme demostrado su amistad. Gracias a todos por los ratos que hemos pasado juntos, por las juergas y por los cursos y congresos, porque sin vuestra presencia no hubieran sido lo mismo, y porque mi paso por el Rocasolano ha merecido la pena, más que nada por haberos conocido a todos, y como una vez me escribisteis: "*Ohú, estamos temblando porque esto sin ti va a ser una tortura. Nos dejas, con la camisa negra y te vas, de perreo caminando por la vida, pero... nada de esto fue un error!*". Inés, ya sabes lo importante que has sido y que eres para mí.

A mis amigos de Granada, Selu, Inma, Kukas, Miguel, David, Yoli y Blanca, Salva, Nico, Ana y Víctor, porque más que amigos, habéis sido mi "*familia*" en Granada. Gracias a todos por escucharme y ayudarme cuando lo he necesitado y por todos los buenos ratos pasados y los que nos quedan por pasar.

A mis amigos Químicos, Vane, Olga, Rafa, Emilio, Mariano, Nico, Jose Manuel, Alberto, Blas y Dori, por seguir estando ahí por mucho tiempo que pase, y porque aunque en algunos casos no nos podamos ver tan a menudo como quisiéramos, siempre encontramos el momento para juntarnos y seguir recordando viejos tiempos.

A mis amigas del Colegio Mayor, Blanca, Lola, Ayo, Estela, María Muñoz, María Moreno, Yolanda y María de la Cruz, por haberme hecho pasar los mejores años de estudiante en Granada, y porque por mucho que pase el tiempo, no perdemos el contacto y seguimos siendo las de siempre.

A mis amigos de Caravaca, Raquel, Mari, Mariola, Susana, Azucena, Jose y Paco, que aunque creo que nunca han entendido muy bien que es lo que hago, siempre han sabido preguntarme e interesarse, apoyándome en todo. Raquel, sabes que eres mi amiga del alma, con la que no me hace falta nada más que una mirada

para saber lo que pensamos. Para mí has sido uno de los apoyos más importantes a lo largo de esta Tesis, y sobre todo a lo largo de mi vida, no cambies nunca, eres la mejor amiga que se puede tener.

A mis amigos del Departamento de Química Analítica, Alegría, Ángel, Rocío y María, por el cariño que me habéis demostrado, y porque desde que os conozco siempre me habéis ayudado en todo lo que he necesitado.

A mis padres, Antonio y Milagros, porque sé lo orgullosos que estáis de mí, aunque lo que soy os lo debo a vosotros. Gracias por ser los mejores padres del mundo, por vuestros consejos, por apoyarme, por haber creído en mí siempre y por permitirme llegar hasta aquí. A mis hermanos, Antonio y May, por haber sido un ejemplo a seguir, por vuestro cariño, por los ánimos recibidos, por desearme siempre todo lo mejor y porque creo que todavía no se han inventado las palabras que puedan expresar lo que significáis para mí. Al resto de mi familia, a mis cuñados, Aurelio y Miriam, por su apoyo incondicional, a mis sobrinas, María, Marta, Ana, Blanca y Miriam, que son la alegría de la casa, y a mis abuelos, que aunque ya no están, se que siempre están a mi lado ayudándome.

Por último, y no por eso menos importante, ya que sin su ayuda nada de esto hubiera sido posible, a Alfonso, porque durante los últimos seis años ha estado día a día a mi lado, apoyándome, animándome, dándome la fuerza necesaria para superar los malos momentos y alegrándose de los buenos. Y a su familia, que siempre ha sabido interesarse y apoyarme en todo.

Y a todos aquellos de los que me haya podido olvidar, les pido perdón y les doy las gracias también.

Gracias a todos

ÍNDICE

CAPÍTULO 1: INTRODUCCIÓN	1
1.1. Introducción	3
1.1.1. Diseño molecular	3
1.1.2. Dominios modulares de reconocimiento molecular	4
1.2. Antecedentes bibliográficos	5
1.2.1. Dominios SH3 y procesos de reconocimiento molecular	5
1.2.2. El dominio SH3: estructura, estabilidad y plegamiento	6
1.2.3. Unión de ligandos al dominio SH3	8
1.2.4. El problema del diseño de ligandos para dominios SH3	11
1.2.5. Proteínas quiméricas como herramientas para el diseño de ligandos	12
1.3. Objetivos de la Tesis	14
CAPÍTULO 2: DISEÑO Y CARACTERIZACIÓN PRELIMINAR DE PROTEÍNAS QUIMÉRICAS	17
2.1. Diseño de quimeras basadas en la topología del permutante circular S19P20s	19
2.2. Caracterización preliminar de la estabilidad de las quimeras basadas en S19P20s	21
2.3. Caracterización preliminar de la conformación de la quimera SPCp41 mediante técnicas espectroscópicas	22
2.3.1. Espectroscopia de fluorescencia	22
2.3.2. Dicroísmo circular	24
2.4. Diseño de proteínas quiméricas con otras topologías	26
CAPÍTULO 3: DETERMINACIÓN DE LA ESTRUCTURA TRIDIMENSIONAL DE LA QUIMERA SPCp41	31
3.1. Asignación de los espectros de RMN de SPCp41	33
3.2. Estructura secundaria de la quimera SPCp41	35
3.3. Cálculo de la estructura tridimensional de SPCp41	38
3.4. Calidad de las estructuras tridimensionales de SPCp41	39
3.5. Dinámica del esqueleto de la quimera SPCp41	44
3.6. Comparación de la estructura de SPCp41 con otras estructuras	44

CAPÍTULO 4: ANÁLISIS DE LA TERMODINÁMICA DE UNIÓN DE DOMINIOS SH3 A PÉPTIDOS RICOS EN PROLINA USANDO LA QUIMERA SPCp41 **49**

4.1. Descripción de los mutantes Pro-Ala de la quimera SPCp41	52
4.2. Caracterización estructural de la quimera SPCp41 y los mutantes Pro-Ala por espectroscopia de fluorescencia y dicroísmo circular	52
4.3. Estudio del desplegamiento térmico de la quimera SPCp41 y de sus mutantes Pro-Ala por calorimetría diferencial de barrido (CDB)	57
4.4. Análisis por RMN de la quimera SPCp41 y el mutante P74A	63
4.5. Análisis de la dinámica de la cadena principal de las proteínas quiméricas mediante experimentos de relajación de RMN ¹⁵ N	65
4.6. Análisis termodinámico de las curvas de capacidad calorífica molar parcial de desplegamiento utilizando un modelo de tres estados	67
4.7. Estudio termodinámico de la unión del permutante circular S19P20s al péptido p41 por calorimetría isotérmica de titulación	73
4.8. Proteínas quiméricas de fusión como herramientas útiles para el análisis termodinámico de las interacciones débiles proteína-ligando	76
4.9. Importancia de los restos Pro en la unión SH3-p41	77
4.10. El impacto de las mutaciones Pro-Ala sobre la termodinámica de la interacción SH3-ligando	77
4.11. La interacción intramolecular SH3-ligando muestra una fuerte compensación entalpía-entropía	79
4.12. Importancia biológica de la unión intramolecular frente a la unión intermolecular de secuencias de baja afinidad	80

CAPÍTULO 5: ESTUDIO CINÉTICO DEL PLEGAMIENTO DE LA QUIMERA SPCp41 **83**

5.1. Cinéticas de plegamiento de la quimera SPCp41 mediante el método de flujo detenido (<i>stopped-flow</i>)	85
5.2. Análisis cinético de la fase lenta del plegamiento de SPCp41	88
5.2.1. Catálisis por prolil isomerasas	90
5.2.2. Análisis mutacional de las prolinas del fragmento p41	91
5.3. La isomerización de prolinas como posible origen de la fase lenta	92
5.4. Análisis cinético de la fase rápida de plegamiento de SPCp41	94
5.4.1. Análisis de los gráficos de Chevron	94

5.4.2. Análisis de las amplitudes de las cinéticas de plegamiento de la quimera SPCp41	98
5.5. Fase ultra-rápida en el plegamiento de SPCp41	101
5.6. Análisis del estado de transición de la quimera SPCp41	106
5.7. Análisis de los errores en los parámetros cinéticos y en los valores de ϕ	113
5.7.1. Desarrollo de una base de datos experimentales	113
5.7.2. Errores en las constantes cinéticas	114
5.7.3. Errores en los parámetros cinéticos obtenidos del análisis de los gráficos de Chevron	116
5.7.4. Análisis de los errores en las energías de Gibbs y en los valores de ϕ	119
5.7.5. Estrategia para el trabajo experimental y el análisis de errores	120
CAPÍTULO 6: MATERIALES Y MÉTODOS	121
6.1. Clonación de las quimeras basadas en la topología del permutante circular S19P20s por mutagénesis dirigida	123
6.1.1. Diseño e introducción de alteraciones en fragmentos de ADN	125
6.1.2. Digestión de los productos de la RCP	128
6.1.3. Ligación del ADN digerido al vector de expresión	128
6.1.4. Introducción del plásmido en el organismo de expresión	129
6.1.5. Ensayos de expresión	129
6.1.6. Purificación del plásmido	129
6.1.7. Secuenciación	130
6.2. Clonación de mutantes de la quimera SPCp41	130
6.3. Clonación de las quimeras basadas en la topología de Itk-SH3	133
6.4. Expresión y purificación de proteínas	133
6.4.1. Cromatografía de exclusión molecular	134
6.4.2. Cromatografía de afinidad	135
6.4.3. Purificación de proteínas marcadas con ^{15}N y/o ^{13}C	135
6.4.4. Expresión y purificación de la ciclofilina humana <i>hCyp18</i>	136
6.5. Preparación de muestras	136
6.5.1. Preparación de tampones y de disoluciones de proteína	136
6.5.2. Preparación de disoluciones concentradas de agente desnaturizante	137
6.6. Calorimetría diferencial de barrido (CDB)	137
6.6.1. El calorímetro diferencial de barrido	138
6.6.2. Realización del experimento calorimétrico	138
6.6.2.1. Cálculo de la capacidad calorífica molar parcial de la proteína	138

6.6.2.2. Análisis de las trazas calorimétricas según la termodinámica de equilibrio	139
6.6.2.2.1. Modelo de equilibrio de dos estados	139
6.6.2.2.2. Modelo de equilibrio de tres estados	142
6.7. Dicroísmo circular	145
6.7.1. Dicroísmo circular en la zona del UV-lejano (170-250 nm)	146
6.7.2. Dicroísmo circular en la zona del UV-cercano (250-350 nm)	147
6.7.3. Obtención de espectros de dicroísmo circular	148
6.8. Espectroscopia de fluorescencia	148
6.8.1. Análisis de la unión proteína-ligando mediante espectroscopia de fluorescencia	149
6.9. Estructura tridimensional de proteínas mediante Resonancia Magnética Nuclear (RMN)	151
6.9.1. Registro de los espectros de RMN	153
6.9.2. Procesado y análisis de los espectros de RMN	155
6.9.3. Asignación de los espectros de RMN	155
6.9.3.1. Asignación de proteínas mediante métodos homonucleares	156
6.9.3.2. Asignación de proteínas mediante métodos heteronucleares	156
6.9.4. Cálculo de la estructura	159
6.9.5. Análisis y calidad de las estructuras	160
6.10. Dinámica de proteínas por RMN	161
6.10.1. Medidas de relajación de ^{15}N en NH amídicos	163
6.10.2. Registro de los espectros	165
6.10.3. Procesado y análisis de los datos	167
6.11. Calorimetría isotérmica de titulación (CIT)	167
6.12. Estudio de la cinética de plegamiento-desplegamiento según la teoría del estado de transición	168
6.12.1. Preparación y realización de los experimentos	169
6.12.2. Determinación de las constantes de velocidad	171
6.12.3. Análisis de las cinéticas de plegamiento-desplegamiento según la teoría del estado de transición	172
6.12.4. Gráficos de Chevron	174
6.12.5. Estados de transición y cálculo de los valores de ϕ mediante el método de ingeniería de proteínas	176
CAPÍTULO 7: RESUMEN Y CONCLUSIONES	183

BIBLIOGRAFÍA	189
APÉNDICES	209
APÉNDICE I: TABLA DE DESPLAZAMIENTOS QUÍMICOS	211
APÉNDICE II: ARTÍCULOS	217

ABREVIATURAS

ABREVIATURAS

[ET][‡]	Estado de transición
[urea]_{1/2}	Concentración de urea en el punto medio de la transición
[θ]	Elipticidad molar
2D	Bidimensional
3D	Tridimensional
A	Amplitud total
A₁	Amplitud de la exponencial asociada al cambio de fluorescencia que ocurre durante la primera fase del proceso cinético
A₂	Amplitud de la exponencial asociada al cambio de fluorescencia que ocurre durante la segunda fase del proceso cinético
Abs	Absorbancia
A_D	Amplitud del estado desnaturalizado
A_{d1}	Amplitud de la exponencial asociada al cambio de fluorescencia que ocurre durante la primera fase del proceso cinético de desplegamiento
A_{d2}	Amplitud de la exponencial asociada al cambio de fluorescencia que ocurre durante la segunda fase del proceso cinético de desplegamiento
ADN	Ácido desoxirribonucleico
A_N	Amplitud del estado nativo
A_{norm}	Amplitud total normalizada
A_{p1}	Amplitud de la exponencial asociada al cambio de fluorescencia que ocurre durante la primera fase del proceso cinético de plegamiento
A_{p2}	Amplitud de la exponencial asociada al cambio de fluorescencia que ocurre durante la segunda fase del proceso cinético de plegamiento
ASA	Área superficial accesible al disolvente
B₀	Intensidad de campo magnético
B_{efectivo}	Campo magnético real que experimenta el núcleo
C	Concentración de proteína
CDB	Calorimetría diferencial de barrido
CIT	Calorimetría isotérmica de titulación
C_{p,D}	Capacidad calorífica del estado D
C_{p,I}	Capacidad calorífica del estado I
C_{p,N}	Capacidad calorífica del estado N

$C_{p,s}$	Capacidad calorífica parcial del disolvente
CSA	Anisotropía de desplazamiento químico
CSI	Índice de desplazamiento químico
D	Estado desnaturalizado de la proteína
d	Proceso de despleamiento
Da	Dalton
DC	Dicroísmo circular
DD	Interacción dipolar
dNTP's	Desoxinucleótidos trifosfato
DSS	2,2-dimetil-2-silapentano-5-sulfonato sódico
EDTA	Ácido etilendiamino tetracético
f_D	Población del estado desnaturalizado
f_I	Población del estado intermedio
f_N	Población del estado nativo
h	Constante de Planck
I	Estado intermedio de la proteína
I_∞	Intensidad de fluorescencia una vez finalizado el proceso
IPTG	Isopropil β -D-tiogalactopiranosido
I_t	Intensidad de fluorescencia
I_t	Intensidad de fluorescencia a tiempo t
J(ω)	Función de densidad espectral
K^\ddagger	Constante de equilibrio químico que tiene lugar entre el estado inicial (D ó N) y el estado de transición ($[ET]^\ddagger$)
$k_{\ddagger-D}$	Constante de velocidad del proceso de plegamiento
$k_{\ddagger-D,urea}$	Constante de velocidad del proceso de plegamiento a una concentración dada de urea
$k_{\ddagger-N}$	Constante de velocidad del proceso de despleamiento
$k_{\ddagger-N,urea}$	Constante de velocidad del proceso de despleamiento a una concentración dada de urea
K_1	Constante de equilibrio entre los estados N e I
k_1	Constante de velocidad de la primera fase del proceso cinético
K_2	Constante de equilibrio entre los estados I y D
k_2	Constante de velocidad de la primera fase del proceso cinético
K_a	Constante de equilibrio de asociación
k_B	Constante de Boltzmann

K_D	Constante de equilibrio de desplegamiento de la proteína
K_d	Constante de equilibrio de disociación del complejo proteína-ligando
k_{d1}	Constante de velocidad de la primera fase del proceso cinético de desplegamiento
k_{d2}	Constante de velocidad de la segunda fase del proceso cinético de desplegamiento
k_{p1}	Constante de velocidad de la primera fase del proceso cinético de plegamiento
k_{p2}	Constante de velocidad de la segunda fase del proceso cinético de plegamiento
k_{urea} = k_{‡-N,urea} + k_{‡-D,urea}	
I	Paso de luz de la célula espectrofotométrica
LB	Medio de cultivo bacteriano Laura Bertani
M	Proteína (macromolécula) libre
m_{‡-D}	Pendiente de la rama de una cinética de plegamiento de un gráfico de Chevron
m_{‡-N}	Pendiente de la rama de una cinética de desplegamiento de un gráfico de Chevron
m_{d1}	Pendiente de la rama de la primera fase de una cinética de desplegamiento en un gráfico de Chevron
m_{D1}	Pendiente de equilibrio de la primera fase de una cinética
m_{d2}	Pendiente de la rama de la segunda fase de una cinética de plegamiento en un gráfico de Chevron
ML	Proteína (macromolécula) complejada
M_p	Masa molecular de la proteína
m_{p1}	Pendiente de la rama de la primera fase de una cinética de plegamiento en un gráfico de Chevron
m_{p2}	Pendiente de la rama de la segunda fase de una cinética de plegamiento en un gráfico de Chevron
N	Estado nativo de la proteína
n	Estequiometría de la interacción
NOE	Efecto Overhauser nuclear
NOESY	Espectroscopia de efecto Overhauser nuclear
OC	Orden de contacto
p	Proceso de plegamiento

PDB	Banco de datos de proteínas
q	Función de partición
R₁	Velocidad de relajación longitudinal
R₂	Velocidad de relajación transversal
RMN	Resonancia magnética nuclear
RMSD	Desviación cuadrática media
S²	Parámetro de orden
SDS	Lauril sulfato sódico
SDS-PAGE	Electroforesis en gel de poliacrilamida en presencia de lauril sulfato sódico
T	Temperatura
t	Tiempo
T_f	Temperatura de fusión de un oligonucleótido
T_m	Temperatura de la transición entre el estado N y D
T_{m,1}	Temperatura de la transición entre el estado N e I
T_{m,2}	Temperatura de la transición entre el estado I y D
TOCSY	Espectroscopia de correlación total
TSP	3-trimetilsilipropionato de sodio
UV	Ultravioleta
V	Función blanco
V_p(T)	Volumen específico parcial de la proteína
V_s(T)	Volumen específico parcial del disolvente
ΔAbs(λ)	Diferencia de absorbancia medida en DC
ΔC_p	Cambio de capacidad calorífica entre el estado N y D
ΔC_{p,1}	Cambio de capacidad calorífica entre el estado N e I
ΔC_{p,2}	Cambio de capacidad calorífica entre el estado I y D
ΔC_{p,ap}	Diferencia de capacidad calorífica aparente entre la disolución de macromolécula y la de disolvente
ΔC_p^{exc}	Exceso de capacidad calorífica
ΔC_p^{int}	Cambio de capacidad calorífica interna de la proteína
ΔG[‡]	Cambio de energía de Gibbs asociado a la constante de equilibrio K [‡]
ΔG_{‡-D}	Cambio de energía de Gibbs asociado a la constante de equilibrio K [‡] entre el estado inicial (D) y el estado de transición ([ET] [‡]) en el proceso de plegamiento

$\Delta G_{\ddagger-N}$	Cambio de energía de Gibbs asociado a la constante de equilibrio K^{\ddagger} entre el estado inicial (D) y el estado de transición ($[ET]^{\ddagger}$) en el proceso de desplegamiento
ΔG_1	Cambio de energía de Gibbs entre el estado N e I
ΔG_2	Cambio de energía de Gibbs entre el estado I y D
ΔG_a	Cambio de energía de Gibbs de la formación del complejo
ΔG_D	Cambio de energía de Gibbs entre el estado N y D
$\Delta G_{D,1}$	Cambio de la energía de Gibbs que tiene lugar durante el desplegamiento para la primera fase de una cinética
ΔH^{\ddagger}	Cambio de entalpía asociado a la constante de equilibrio K^{\ddagger}
ΔH_1	Cambio de entalpía entre el estado N e I
ΔH_1	Cambio de entalpía entre el estado N e I
ΔH_2	Cambio de entalpía entre el estado I y D
ΔH_2	Cambio de entalpía entre el estado I y D
ΔH_a	Cambio de entalpía de unión
ΔH_D	Cambio de entalpía entre el estado N y D
ΔH_m	Cambio de entalpía entre el estado N y D a la T_m
$\Delta H_{m,1}$	Cambio de entalpía entre el estado N e I a la T_m
$\Delta H_{m,2}$	Cambio de entalpía entre el estado I y D a la T_m
ΔS^{\ddagger}	Cambio de entropía asociado a la constante de equilibrio K^{\ddagger}
ΔS_1	Cambio de entropía entre el estado N e I
ΔS_2	Cambio de entropía entre el estado I y D
ΔS_a	Cambio de entropía de la formación del complejo
ΔS_D	Cambio de entropía entre el estado N y D
ΔS_m	Cambio de entropía entre el estado N y D a la T_m
$\Delta S_{m,1}$	Cambio de entropía entre el estado N e I a la T_m
$\Delta S_{m,2}$	Cambio de entropía entre el estado I y D a la T_m
$\Delta \Delta G_{\ddagger-D}$	Cambio en la energía del estado de transición entre la especie de referencia y el mutante en el proceso de plegamiento
$\Delta \Delta G_{\ddagger-N}$	Cambio en la energía del estado de transición entre la especie de referencia y el mutante en el proceso de desplegamiento
$\Delta \Delta G_{N-D}$	Energía de desestabilización inducida por una mutación
$\Delta \varepsilon = \varepsilon_I(\lambda) - \varepsilon_D(\lambda)$	

Abreviaturas

δ	Desplazamiento químico
ϵ_D	Coefficiente de extinción molar para la luz polarizada a la derecha
ϵ_I	Coefficiente de extinción molar para la luz polarizada a la izquierda
$\phi_{\pm D}$	Grado de formación de cada interacción en el estado de transición
γ	Constante giromagnética del núcleo
λ	Longitud de onda
ν	Frecuencia de resonancia de un núcleo
$\nu_{\text{referencia}}$	Frecuencia de resonancia de un núcleo elegido como referencia
θ	Elipticidad
σ	Constante de apantallamiento
σ_x	Error de x obtenido del ajuste de los datos experimentales
τ_m	Tiempo de correlación global

CAPÍTULO 1: INTRODUCCIÓN

CAPÍTULO 1: INTRODUCCIÓN

1.1. Introducción

1.1.1. Diseño molecular

El establecimiento de interacciones transitorias entre moléculas biológicas es crucial para el desarrollo correcto del ciclo de vida celular. Como consecuencia, la elucidación de los factores que determinan el reconocimiento molecular posee un indudable interés académico, además de enormes repercusiones biotecnológicas y terapéuticas. Ejemplos de estas interacciones son la unión de sustratos e inhibidores a enzimas, la identificación de antígenos por sus correspondientes anticuerpos, la interacción entre hormonas y sus receptores celulares o el establecimiento de redes de transducción de señales. La comprensión profunda y detallada de estos procesos nos permitiría predecir y modular la respuesta fisiológica o incluso patológica de la célula a un estímulo específico y facilitaría el desarrollo de estrategias racionales de diseño molecular que conduzcan a la obtención de nuevas proteínas con las funciones y estabildades deseadas, o al desarrollo de nuevos fármacos optimizados en cuanto a su efectividad y sus efectos secundarios.

En la última década hemos sido testigos de una gran revolución en el campo de la genómica y proteómica que ha hecho aún más patente la necesidad de profundizar en el estudio del reconocimiento molecular. La resolución del *Genoma Humano*, junto con los de otros organismos modelo y de varios agentes patógenos de gran importancia, ha generado un nuevo panorama científico que requiere la revisión y actualización de las metodologías tradicionales de diseño molecular. Los miles de genes identificados en cada genoma codifican al menos otras tantas 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 constituyen potenciales dianas para el desarrollo de nuevos agentes terapéuticos contra una gran variedad de enfermedades. Para cada nueva diana será necesario identificar o diseñar, en su caso, nuevos compuestos que regulen su actividad y que, tras ser optimizados, puedan convertirse en nuevos agentes terapéuticos con la afinidad de unión, especificidad, selectividad, bio-disponibilidad y propiedades toxicológicas requeridas para un tratamiento efectivo de la enfermedad.

Se hace necesario, por tanto, incrementar la eficacia de los algoritmos de diseño molecular, lo que requiere una mayor racionalización del proceso. Para ello es importante profundizar en la relación existente entre la estructura y la estabilidad de una proteína y sus características funcionales, e incorporar al proceso de diseño no sólo aspectos estructurales, sino consideraciones termodinámicas y dinámicas (estabilidad y cooperatividad), genéticas y funcionales, lo que requiere un profundo análisis multidisciplinar de los fenómenos de reconocimiento molecular.

1.1.2. Dominios modulares de reconocimiento molecular

En estos procesos de reconocimiento molecular, a través de los que se organiza y coordina todo el ciclo de vida celular, frecuentemente intervienen como intermediarios una serie de pequeños dominios modulares, altamente conservados, que forman parte de diversas proteínas, en número variable y en distintas combinaciones (Pawson y Scott, 1997; Pawson *et al.*, 2002). Se trata de pequeños dominios (entre 30 y 150 aminoácidos) que se encuentran plegados en estructuras compactas y estables, caracterizadas por uno o más sitios de unión. Cada familia de dominios reconoce en sus proteínas diana ciertas secuencias específicas (generalmente de entre 3 y 6 aminoácidos) y muy conservadas. Así, por ejemplo, los dominios SH3 (*Src-Homology region-3*) reconocen secuencias ricas en restos de prolina, los dominios SH2 (*Src-Homology region-2*) secuencias peptídicas con restos de tirosina fosforilada y los dominios PDZ reconocen secuencias correspondientes a extremos carboxilo terminal (Kuriyan y Cowburn, 1997; Kay, B. K. *et al.*, 2000; Yan *et al.*, 2002; Jelen *et al.*, 2003; Ball *et al.*, 2005). Estas secuencias características se encuentran flanqueadas por aminoácidos que interactúan con elementos variables del sitio de unión del dominio y que parecen ser responsables de la especificidad de unión dentro de cada familia de dominios, es decir, de la selección de las proteínas funcionalmente complementarias “*in vivo*”. Las interacciones mediadas por dominios modulares no requieren de grandes superficies de contacto sino que se trata de superficies relativamente pequeñas, bastante similares a las que se encuentran en los sitios de unión de pequeños ligandos a enzimas, lo que facilita el diseño de pequeños compuestos que bloqueen con alta especificidad estas interacciones y hacen de estos módulos de reconocimiento interesantes dianas para el desarrollo de nuevos agentes terapéuticos (Kardinal *et al.*, 1999; Kardinal *et al.*, 2000; Lee *et al.*, 2002; Lawrence, 2005).

Sin embargo, aún no están suficientemente claros los factores que determinan la afinidad y la especificidad de la interacción entre estos dominios de reconocimiento y sus dianas, de tal forma que resulta imprescindible profundizar en el estudio de los diferentes elementos que dirigen las interacciones entre estos dominios y sus ligandos. Así, en los últimos años se ha podido establecer de forma mucho más concluyente la enorme interdependencia que existe entre la estructura y la estabilidad de una proteína y su función biológica. La importancia de incorporar consideraciones termodinámicas (contribuciones entálpicas y entrópicas a la afinidad de unión, estabilidad y cooperatividad, etc.) al proceso de diseño ha quedado demostrada recientemente (Luque y Freire, 2000; Velazquez-Campoy *et al.*, 2001; Luque *et al.*, 2002; Velazquez Campoy y Freire, 2005).

Por otro lado, en la era de la proteómica, se hace necesario el desarrollo de métodos fiables de predicción de posibles dianas celulares dentro de un genoma concreto, así como de identificación de qué secuencias serán reconocidas por un dominio en particular. En este sentido la integración de toda la información disponible en un “*código de reconocimiento molecular*” es clave para la identificación y predicción de ligandos potenciales de un dominio determinado en el contexto de los distintos genomas (Sudol, 1998; Brannetti *et al.*, 2000; Tong *et al.*, 2002).

1.2. Antecedentes bibliográficos

1.2.1. Dominios SH3 y procesos de reconocimiento molecular

Como se ha indicado anteriormente, el ciclo de vida de la célula y su respuesta ante los estímulos que recibe tanto del interior como del exterior se encuentra organizado según una compleja red de procesos de reconocimiento molecular en los que intervienen los dominios SH3, que probablemente sean los módulos de reconocimiento más extendidos del proteoma (Mayer, 2001). Mediante algoritmos de cribado de bases de datos se han logrado identificar hasta la fecha más de 1500 dominios SH3 diferentes, que forman parte de proteínas tales como quinasas, lipasas, GTPasas, proteínas adaptadoras y estructurales (o proteínas víricas) que participan en las redes de comunicación intracelular, en la organización del esqueleto citoplasmático y en el tráfico de membranas (Skorski *et al.*, 1998; Stein, 1998; McPherson, 1999). Los dominios SH3 están por tanto presentes en numerosas proteínas modulares, muchas de las cuales ni siquiera están funcionalmente relacionadas, tanto en organismos unicelulares como pluricelulares, donde no poseen una posición topológica definida (Koch *et al.*, 1991; Musacchio *et al.*, 1992a; Musacchio *et al.*, 1992b; Pawson y Schlessinger, 1993),

Se sabe que los dominios SH3 actúan frecuentemente como sitios de anclaje para la captura de sustratos o para la formación de complejos supramoleculares que a menudo conducen a la modificación enzimática de algunos de sus componentes, dando lugar a la aparición de nuevos sitios de interacción así como a la propagación y amplificación de señales químicas intracelulares (Kiyokawa *et al.*, 1997; Klejman *et al.*, 2002). En ocasiones, estos dominios 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 (Barila y Superti-Furga, 1998; Arold, S. T. *et al.*, 2001; Brasher *et al.*, 2001; Brabek *et al.*, 2002). Este es el caso de las tirosinas kinasas de la familia Src, algunos de cuyos miembros, especialmente c-Src y c-Yes, están claramente implicados en el desarrollo de procesos cancerígenos (Miyazaki *et al.*, 1999). Además del cáncer, los dominios SH3 están asociados con otras patologías como el SIDA (Lyn, Hck, Lck), la leucemia (Abl), la osteoporosis (Src) o ciertos procesos inflamatorios, alérgicos y asmáticos (Dalgarno *et al.*, 1997; Skorski *et al.*, 1998). De ahí que el diseño de ligandos que interfieran en el establecimiento de las interacciones entre estos dominios y sus dianas se ha consolidado en los últimos años como una estrategia viable para el desarrollo de nuevos fármacos para el tratamiento de este tipo de patologías (Stauffer *et al.*, 1997; Kardinal *et al.*, 2000; Vidal *et al.*, 2001; Lee *et al.*, 2002; Oneyama *et al.*, 2002).

Así, además de formar parte de proteínas implicadas en rutas específicas de transducción de señales, en interacciones proteína-proteína, funciones de reconocimiento a nivel molecular, etc. (Viguera *et al.*, 1994a; Pawson, 1995; Xu *et al.*, 1997; Skorski *et al.*, 1998; Stein, 1998; McPherson, 1999), los dominios SH3 también están presentes en numerosas proteínas del citoesqueleto (α -Spc) proporcionando de esta manera un posible nexo entre las distintas rutas de transducción de señales y los cambios morfológicos que sufre la célula (Lehto *et al.*, 1988; Musacchio *et al.*, 1992b).

1.2.2. El dominio SH3: estructura, estabilidad y plegamiento

Además de por su importancia biológica, los estudios de los dominios SH3 presentan un enorme interés para el estudio estructural y energético de su plegamiento, estabilidad e interacciones, ya que se trata de modelos muy adecuados dado su pequeño tamaño, con un peso molecular inferior a los 10 kDa. Además se pueden sobre-expresar fácilmente mediante protocolos convencionales de ingeniería de proteínas y pueden purificarse y cristalizarse de forma totalmente independiente del resto de la proteína de la que forman parte, lo que constituye una gran ventaja para su estudio sistemático.

Se han resuelto las estructuras tridimensionales de un gran número de dominios SH3 tanto en sus formas naturales, como en una gran variedad de mutantes y quimeras. Concretamente, la estructura del de α -espectrina (Spc-SH3) se ha conseguido resolver haciendo uso tanto de la cristalografía de rayos X (Musacchio *et al.*, 1992b) como de la espectroscopia de RMN en disolución (Blanco *et al.*, 1997). Todos los dominios SH3 contienen entre 55 y 75 aminoácidos y comparten una estructura tridimensional común que consiste en dos láminas β antiparalelas dispuestas ortogonalmente entre sí. El sitio de unión de estos dominios contiene una región hidrofóbica muy conservada, flanqueada por dos lazos (conocidos como *RT* y *n-Src*) de composición bastante más variable y de naturaleza más polar.

Más concretamente, en la Figura 1.2.1 se representa la estructura tridimensional de rayos X del dominio Spc-SH3, prácticamente idéntica a la obtenida por RMN. Este dominio está constituido por 62 aminoácidos distribuidos en cinco hebras β denominadas β_a (del residuo 8 a 11), β_b (del 30 a 35), β_c (del 41 a 46), β_d (del 49 a 54) y β_e (del 58 a 61). Estas hebras β forman dos láminas β antiparalelas denominadas β_I y β_{II} . La lámina más pequeña, la β_I , está constituida por las hebras β_b - β_a - β_e , y la mayor de ellas, la β_{II} , la forman las hebras β_b - β_c - β_d . Otro elemento estructural de este dominio es un largo lazo de 19 residuos denominado *lazo RT* que conecta las dos primeras hebras β_a y β_b formando una estructura a modo de horquilla. La conexión entre las hebras β_d y β_e se realiza mediante una vuelta de hélice 3_{10} (residuos 55 a 57) que se inicia tras la Pro54, muy conservada entre las distintas secuencias de la familia SH3. Entre las hebras β_c y β_d existe un giro β de tipo II' llamado *giro distal*; en este giro los residuos Val46 y Arg49 forman un puente de hidrógeno mediante átomos de la cadena principal. Este giro β podría inducir una cierta tensión conformacional en su estructura, ya que el residuo Asp47 ocupa una posición no permitida en el mapa de Ramachandran (Vega *et al.*, 2000). Finalmente, el lazo denominado *n-Src* conecta las hebras β_b y β_c .

Los restos más conservados de la secuencia están próximos entre sí desde un punto de vista estructural. De hecho, se localizan entre los lazos *RT* y *n-Src* formando parte del sitio de unión en las interacciones proteína-proteína (Argos, 1988). No obstante, algunos de los aminoácidos que forman el sitio de unión no son exactamente los mismos en todos los dominios SH3, sino que presentan ligeras variaciones, lo cual nos da idea de que éstos reconocen secuencias similares pero no idénticas, de manera que esta diversidad les otorga una cierta especificidad estructural (Ren *et al.*, 1993). De hecho, los dominios SH3 presentan una mayor homología estructural que de secuencia.

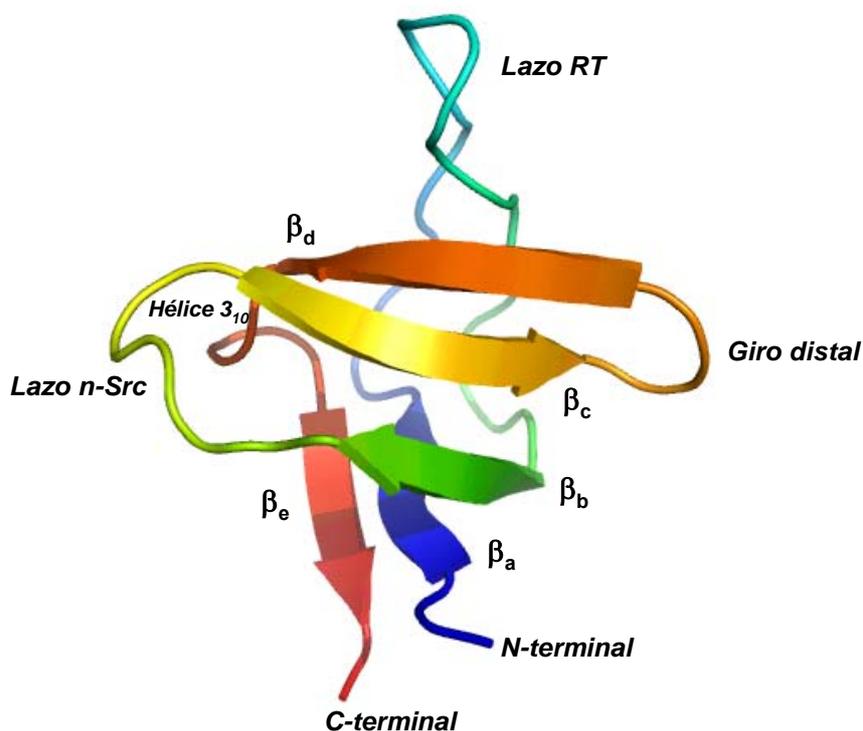


Figura 1.2.1. Representación en forma de cinta de la estructura tridimensional de la forma natural del dominio Spc-SH3 obtenida por cristalografía de rayos X (Musacchio *et al.*, 1992b) y depositada en el Protein Data Bank (PDB) con el código 1SHG. Se indican algunos elementos de estructura sobre la misma

En cuanto a la estabilidad estructural y el plegamiento de este tipo de dominios, cabe mencionar que son varios los grupos de investigación que han llevado a cabo estudios de este tipo, empleando para ello diversas técnicas como las cinéticas de plegamiento mediante flujo detenido (*stopped-flow*) utilizando observables espectroscópicos como el dicroísmo circular o la fluorescencia, la calorimetría diferencial de barrido (CDB), o el análisis de la dependencia del plegamiento y la estabilidad con el pH o con la concentración de agentes desnaturantes (Viguera *et al.*, 1994b; Guijarro *et al.*, 1998; Plaxco *et al.*, 1998a; Filimonov *et al.*, 1999; Sadqi *et al.*, 1999). Los estudios cinéticos y de equilibrio del plegamiento-desplegamiento de estos dominios indican que, en la mayoría de las condiciones experimentales, es un proceso aparentemente sencillo, que puede describirse adecuadamente según un *modelo de dos estados*, en el que no aparecen estados intermedios significativamente poblados.

Los estudios cinéticos llevados a cabo sobre el plegamiento este dominio, aplicando la técnica de flujo detenido, han puesto de manifiesto la ausencia de intermedios durante los procesos de plegamiento y desplegamiento. El núcleo de plegamiento se encuentra localizado en la horquilla β que forman las hebras β_c y β_d en torno al *giro distal* y en la hélice 3_{10} (Viguera *et al.*, 1996). La velocidad de plegamiento de este dominio resulta ser menor que la observada para otras proteínas de estructura β del mismo tamaño (Perl *et al.*, 1998). Según estos estudios, uno de los factores que limita la velocidad de plegamiento de este dominio es la formación de dicho giro tipo β . Así, los estudios con permutantes circulares del dominio Spc-SH3 (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 β presentes en la estructura de dicho dominio. En dichos estudios se analizó la estabilidad y el plegamiento de un conjunto de permutantes circulares, realizados sobre el dominio Spc-SH3, en los que se trató de interrumpir la unión covalente entre las dos hebras β que se encuentran formando cada una de las tres horquillas en la estructura del dominio natural, a la vez que se formaba una nueva horquilla uniendo las hebras antiparalelas de los extremos N y C-terminales, próximas entre sí en la estructura tridimensional (Viguera *et al.*, 1996; Martínez *et al.*, 1999). La disposición estructural de estos permutantes resultó ser similar a la del dominio Spc-SH3 natural y, aunque todos ellos presentaron una estabilidad parecida, se comprobó que sus constantes de plegamiento y desplegado diferían notablemente. Estos datos, junto con la aplicación del método de ingeniería de proteínas, demostró que la ruta de plegamiento cambia en los mutantes permutados circularmente, lo que indica que ni el orden de la estructura secundaria en la cadena (alterado a través de las permutaciones) ni la conservación de cualquiera de las horquillas β presentes en el dominio natural juegan un papel crucial para que se produzca el plegamiento, aunque ambos factores podrían influir en las cinéticas de plegamiento y de desplegado alterando su ruta. A partir de aquí, los estudios cinéticos de estos permutantes variando la longitud de la nueva horquilla (formada por unión de los extremos N y C-terminales de la proteína natural) mediante la inserción de restos de glicina, demostraron que la difusión intramolecular también desempeña un papel importante en el proceso de plegamiento (Viguera y Serrano, 1997), sugiriendo que la dificultad para alcanzar el estado de transición aumenta al aumentar la longitud de la nueva horquilla. Posteriormente, se demostró mediante estudios cinéticos y termodinámicos (Martínez y Serrano, 1999; Martínez *et al.*, 1999) que la longitud de la horquilla estudiada produce un efecto de carácter entrópico en el proceso de plegamiento, y que el estado de transición es bastante compacto y ordenado. De estos estudios se dedujo que puede lograrse un aumento en la velocidad de plegamiento de las proteínas disminuyendo la barrera entrópica del proceso.

A pesar de la aparente simplicidad macroscópica de los procesos de plegamiento-desplegado de estos pequeños dominios, los estudios recientes de intercambio hidrógeno-deuterio (IHD) seguido por RMN con el dominio Spc-SH3, llevados a cabo por nuestro grupo de investigación (Sadqi *et al.*, 1999; Sadqi *et al.*, 2002a; Sadqi *et al.*, 2002b; Casares *et al.*, 2007) indican que en condiciones nativas el dominio se encuentra en un equilibrio conformacional de naturaleza estadística que implica conformaciones muy variadas, que comprenden desde el estado nativo, totalmente 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 transmita de forma eficiente al resto de la estructura (Casares *et al.*, 2003; Casares *et al.*, 2007).

1.2.3. Unión de ligandos al dominio SH3

Los dominios SH3 reconocen secuencias polipeptídicas ricas en prolina, a las que se unen por lo general con una afinidad moderada o baja (K_d entre 10^{-6} y 10^{-4} M $^{-1}$). Los ligandos de estos dominios se unen a los mismos adoptando una conformación en hélice de tipo poliprolina II (PPII) con dos posibles orientaciones sobre el sitio de unión.

Estos ligandos, con motivos característicos del tipo PxxP, pueden interactuar con secuencias consenso del tipo RxLPP#P y Px#PxR para los ligandos de *clase I* y de *clase II* respectivamente, donde # es frecuentemente un residuo hidrofóbico y x representa a cualquier otro residuo (Alexandropoulos *et al.*, 1995). Los ligandos de *clase I* y *clase II* se unen a los dominios SH3 con orientaciones opuestas, de modo que los de *clase I* (+) lo hacen con su extremo N-terminal en el lazo RT y los de *clase II* (-) lo hacen con el extremo C-terminal en esa posición, como podemos observar en la **Figura 1.2.2**.

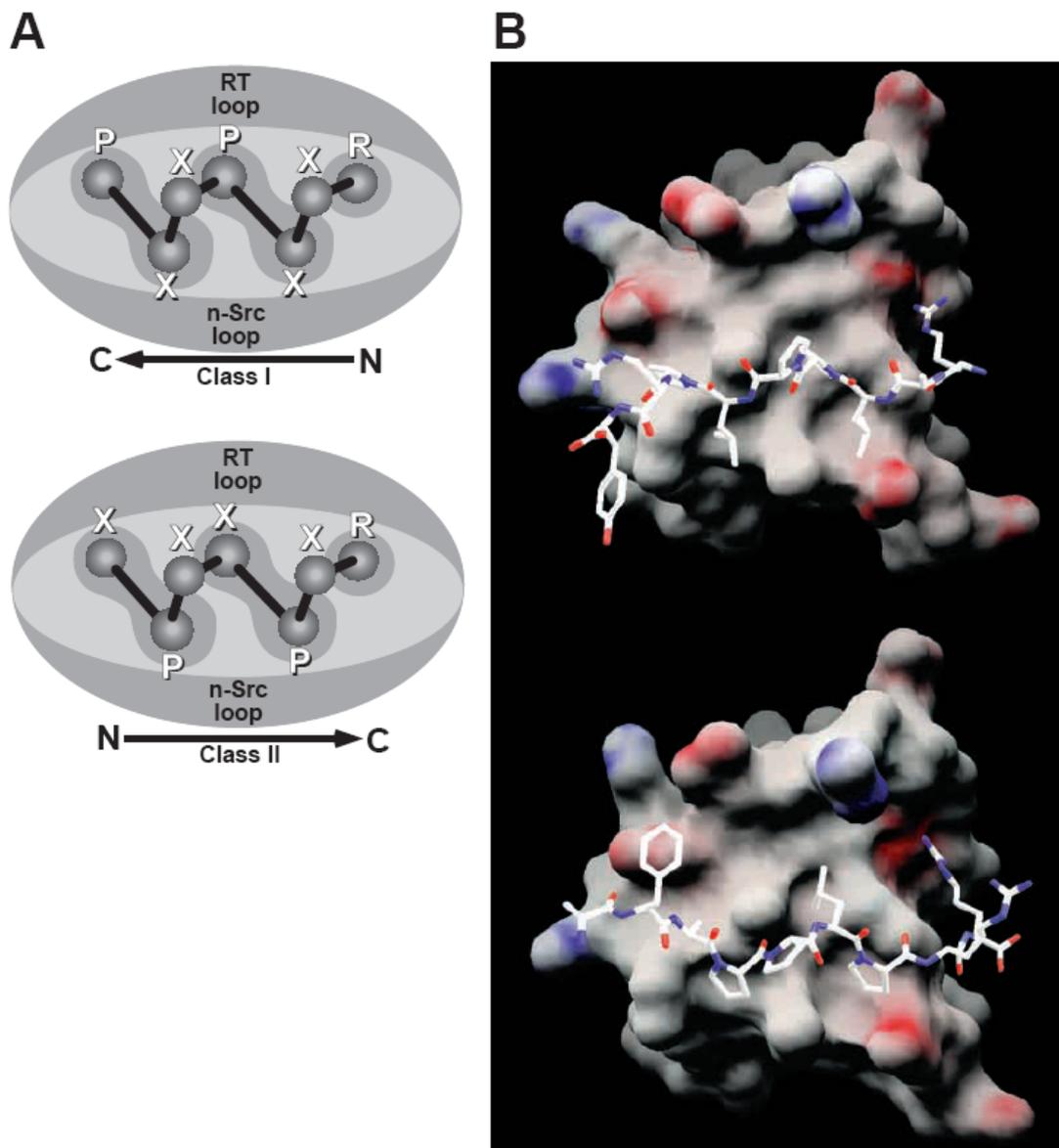


Figura 1.2.2. Interacción de los dominios SH3 con sus ligandos, tomada de la referencia (Mayer, 2001). **A.** Representación esquemática de la unión de ligandos de *clase I* (arriba) y *clase II* (abajo) al dominio SH3. **B.** Ejemplo de la unión del dominio Src-SH3 a un ligando de *clase I* (arriba) y a un ligando de *clase II* (abajo).

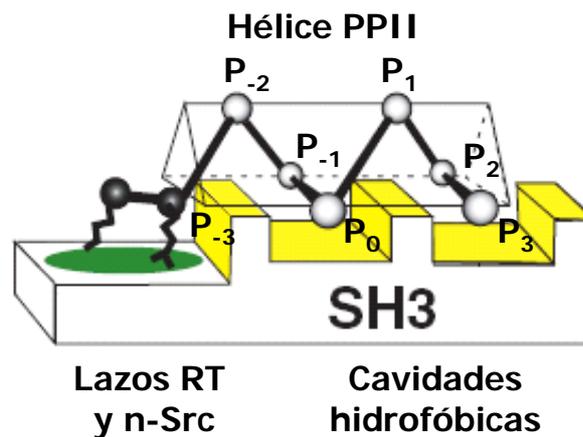


Figura 1.2.3. Representación esquemática del sitio de unión del dominio SH3, en la que se observan las dos cavidades hidrofóbicas y el bolsillo de especificidad donde se alojan los ligandos ricos en prolina. Las esferas y varillas corresponden a la estructura del ligando.

La región del dominio SH3 implicada en la interacción con secuencias ricas en prolina consiste en una superficie hidrofóbica formada por las cadenas laterales de residuos que se encuentran muy conservados en estos dominios (Musacchio *et al.*, 1992b; Yu *et al.*, 1992; Musacchio *et al.*, 1994; Yu *et al.*, 1994). Los residuos que forman parte del sitio de unión son: los dos residuos aromáticos que se encuentran en el motivo ALYDY, que está muy conservado en todos los dominios SH3, el primer W del motivo WW, también muy conservado, y residuos de la hélice 3_{10} , formada por el motivo PXXY. El sitio de unión presenta dos cavidades hidrofóbicas poco profundas revestidas por las cadenas laterales de los residuos citados. En estas cavidades se insertan las cadenas laterales de las prolinas del ligando. Esta interacción parece conferir un marcado carácter hidrofóbico a la unión de este tipo de ligandos ricos en prolina. Para establecer esta interacción, el péptido debe adoptar una conformación en hélice de poliprolina II. Este tipo de interacción está altamente conservada en las interacciones entre las diferentes variantes de dominios SH3 y sus ligandos peptídicos. Existe una tercera cavidad en el sitio de unión, llamada bolsillo de especificidad (bolsillo “compás”), que se encuentra flanqueada por los lazos RT y n-Src (Musacchio, 2002; Zarrinpar *et al.*, 2003)}. En la Figura 1.2.3 hemos representado un esquema del sitio de unión del dominio SH3, en el que se observan las dos cavidades hidrofóbicas y el bolsillo de especificidad donde se alojan los ligandos ricos en prolina. En esta Figura podemos observar que en la primera cavidad se alojan las posiciones P₃ y P₂, y en la segunda P₀ y P₋₁. La parte restante del sitio de unión es el bolsillo de especificidad, la posición P₋₃. Las posiciones P₁ y P₂ se orientan hacia fuera de la proteína.

El cambio en la posición de unión del ligando de (+) a (-) provoca que los residuos de Pro ocupen sitios diferentes en la superficie de SH3. Aún cuando la hélice de PPII es pseudo-simétrica, y presenta grupos aceptores de enlaces de H (C=O) dirigidos hacia los mismos grupos donores en cualquiera de las dos orientaciones, hay una preferencia en el tipo de unión de los residuos de Pro en sitios concretos de unión. Así, para los ligandos de *clase I* (+), la Pro se prefiere en las posiciones P₀ y P₃, mientras que en los ligandos de *clase II* (-) se prefiere en las posiciones P₋₁ y P₂. Esta especificidad de los residuos Pro por sitios de unión concretos es consecuencia de las diferencias en el empaquetamiento de los ligandos de *clase I* y *II* con la superficie de la proteína.

Según el modelo de Lim (Lim *et al.*, 1994) los residuos de las hélices de PPII se pueden empaquetar en el bolsillo de unión de dos formas diferentes: con el C_α apuntando hacia la superficie de la proteína (un sitio con “*empaquetamiento interno*”) o con el mismo apuntando hacia fuera de la proteína (un sitio con “*empaquetamiento externo*”). Los sitios con “*empaquetamiento externo*” favorecidos para los residuos Pro en la unión *clase I (+)* son P₀ y P₃, mientras que en la unión *clase II (-)* estos sitios son P₋₁ y P₂. Por el contrario, los sitios con “*empaquetamiento interno*” favorecidos para los residuos Pro en la unión *clase I (+)* son los sitios P₋₁ y P₂, mientras que en la unión *clase II (-)* estos sitios son P₀ y P₃.

En los últimos años, la identificación de ligandos con motivos diferentes a los descritos anteriormente parece cuestionar la generalidad global del modelo y sugiere un potencial de interacción para la familia de dominios SH3 más amplio de lo inicialmente establecido (Lewitzky *et al.*, 2001; Agrawal y Kishan, 2002; Kami *et al.*, 2002; Lewitzky *et al.*, 2004). Así, la identificación de algunos ligandos naturales, ya sean intra o intermoleculares, que carecen de prolina en su secuencia pero que, sin embargo, interaccionan de forma favorable con estos dominios, parece señalar que la condición realmente crítica para la unión de estos ligandos es su capacidad para adoptar una conformación en hélice de poliprolina en el complejo, siempre que el coste energético del proceso no sea demasiado elevado. Son varios los grupos que han investigado los determinantes de la afinidad y la especificidad en la unión de ligandos de poliprolina a este tipo de dominios. Por ejemplo, el grupo del Dr. Cesareni ha desarrollado una clasificación de los diferentes dominios SH3 de acuerdo con sus preferencias por distintas secuencias en los ligandos y además han desarrollado también un primer algoritmo computacional para la identificación de posibles ligandos de los dominios SH3 en las bases de datos de secuencias (Brannetti *et al.*, 2000; Cesareni *et al.*, 2002).

1.2.4. El problema del diseño de ligandos para dominios SH3

A pesar del interés que la inhibición de las interacciones entre estos módulos de reconocimiento molecular de secuencias ricas en prolina 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. Muchos han sido los esfuerzos dirigidos a la identificación y diseño de ligandos de alta afinidad para dominios SH3 (Dalgarno *et al.*, 1997; Garbay *et al.*, 2000) que incluyen estrategias diversas como el uso de bibliotecas de compuesto sintéticos, la sustitución sistemática de las distintas posiciones del ligando por alanina (Chen, J. K. *et al.*, 1993; Ren *et al.*, 1993; Feng *et al.*, 1994), técnicas de “*phage display*” (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; Witter *et al.*, 1998) 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 combinatoriales con bibliotecas de peptoides que incluyen sustitutos sintéticos de las prolina clave en el ligando ha proporcionado algunos ligandos de muy alta afinidad (Nguyen, Jack T. *et al.*, 1998; Aghazadeh y Rosen, 1999; Nguyen, J. T. *et al.*, 2000).

La aplicación de técnicas de diseño racional ha proporcionado algunos resultados satisfactorios, como los obtenidos por el grupo del Dr. Serrano (EMBL, Heidelberg) que ha conseguido secuencias peptídicas ricas en prolina altamente afines y específicas para el dominio SH3 de Abl (Pisabarro *et al.*, 1994; Pisabarro y Serrano, 1996). Sin embargo, el éxito de estas metodologías de diseño racional

seguirá estando muy limitado mientras no se posea una comprensión profunda y detallada de los determinantes de la afinidad y especificidad de unión para el reconocimiento de ligandos ricos en prolina por los dominios SH3. 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. En este sentido, durante los últimos años nuestro grupo de investigación, en colaboración con el grupo del Dr. Freire (Johns Hopkins University, Baltimore, USA) han trabajado en la incorporación de estas consideraciones termodinámicas al proceso de diseño, lo que ha permitido abordar el análisis molecular no sólo de los orígenes de la afinidad y especificidad de los ligandos, sino del desarrollo de resistencia a fármacos y de las redes de transmisión de información y cooperatividad (Luque y Freire, 1998; Luque *et al.*, 1998; Velazquez-Campoy y Freire, 2001; Velazquez-Campoy *et al.*, 2001; Luque y Freire, 2002; Luque *et al.*, 2002; Velazquez Campoy y Freire, 2005).

1.2.5. Proteínas quiméricas como herramientas para el diseño de ligandos

Posiblemente la técnica que se considera más idónea para investigar la unión de ligandos a las proteínas es la calorimetría isotérmica de titulación (CIT), ya que permite la caracterización termodinámica completa de la interacción, proporcionando no sólo información sobre la afinidad de unión, sino también sobre las contribuciones entálpicas y entrópicas al proceso. Una adecuada interpretación de estos parámetros termodinámicos en combinación con datos de tipo estructural y dinámico sería la clave para entender en profundidad los determinantes moleculares responsables del proceso de unión.

Así, por ejemplo, la interacción del dominio SH3 con secuencias ricas en prolina ha sido considerada hasta la fecha como de naturaleza principalmente hidrofóbica debida al empaquetamiento de restos hidrofóbicos del dominio con los del péptido, fundamentalmente Pro. De hecho, en los bolsillos conservados del sitio de unión para la interacción con secuencias consenso sólo nos encontramos una interacción de tipo polar, que consiste en un puente de hidrógeno formado por una tirosina muy conservada, localizada en la hélice 3_{10} del dominio SH3 y el oxígeno del carbonilo en los ligandos (Zarrinpar *et al.*, 2003; Ball *et al.*, 2005). No obstante, se han observado otras interacciones polares o electrostáticas que generalmente se establecen con el bolsillo de especificidad o con restos de los lazos *RT* o *n-Src*.

El patrón termodinámico esperado para una interacción mayoritariamente hidrofóbica debería estar gobernado por una contribución entrópica favorable (positiva), debida principalmente a la liberación de las moléculas de agua que solvatan las superficies hidrofóbicas del ligando y proteína que pasan a interactuar entre sí en el complejo (Velazquez-Campoy *et al.*, 2000). La contribución entálpica debería ser desfavorable (también positiva) o, en el mejor de los casos, ligeramente favorable debido a que los contactos de van der Waals formados en una interacción entre superficies de tipo hidrofóbico en el complejo y las interacciones de solvatación de estas mismas superficies en las especies libres se compensan prácticamente entre sí.

Los análisis calorimétricos llevados a cabo sobre los complejos SH3-péptido revelan un patrón termodinámico característico gobernado por una entalpía de unión muy favorable (fuertemente exotérmica), a la que se opone la contribución entrópica (Feng *et al.*, 1996; Palencia *et al.*, 2004; Casares *et al.*, 2007). Esto es justo lo contrario a lo que cabría esperarse para una interacción de tipo hidrofóbico, tal y como hemos descrito anteriormente. Éste hecho es difícil de explicar teniendo en cuenta la naturaleza hidrofóbica de las superficies interaccionantes, y pone de manifiesto la complejidad asociada a las interacciones establecidas por dominios SH3. El análisis termodinámico llevado a cabo en nuestro laboratorio para la interacción del péptido p41 con el dominio Abl-SH3 ha proporcionado también resultados sorprendentes, tales como el origen entálpico del incremento de afinidad asociado a las sustituciones de Leu a Pro en ligandos peptídicos o los efectos cooperativos de carácter entálpico originados por sustituciones en posiciones expuestas al disolvente del ligando (Palencia *et al.*, 2004).

Por tanto, además de las interacciones de tipo hidrofóbico establecidas en el sitio de unión, deberíamos considerar la posibilidad de factores adicionales que contribuyen en la interacción, tales como la presencia de moléculas de agua ocluidas en la interfase de unión (Palencia *et al.*, 2004) y la modulación de las propiedades dinámicas y conformacionales de la proteína como consecuencia de la formación del complejo (Ferreon y Hilser, 2004; Casares *et al.*, 2007). Éstos factores podrían ser importantes para la interacción, tal y como reflejan los parámetros termodinámicos obtenidos en diferentes trabajos, los cuales no somos capaces de explicar para una interacción de tipo hidrofóbico.

Uno de los grandes problemas en el análisis termodinámico de las interacciones SH3-ligando es que sus afinidades normalmente están por debajo del intervalo accesible para su estudio por CIT (Wiseman *et al.*, 1989). Para realizar estos experimentos de CIT se necesitan decenas de miligramos de ligando peptídico y de la proteína, lo que hace inviable el uso de la CIT para un estudio en profundidad de la energética de unión, debido al elevado coste de los péptidos de síntesis y a problemas colaterales como la limitada solubilidad y/o agregación de las muestras. Por otra parte, el análisis de la termodinámica de la interacción proteína-ligando mediante CDB se ha utilizado con anterioridad, tanto en nuestro grupo de investigación como en otros, presentando ciertas ventajas con respecto a la CIT en muchos casos (Robert *et al.*, 1988; Brandts y Lin, 1990; Shrake y Ross, 1990; Straume y Freire, 1992; Martinez *et al.*, 1994; Martinez *et al.*, 1995; Casares *et al.*, 2007).

En la bibliografía existen algunos ejemplos de proteínas quiméricas artificiales que han sido utilizadas para analizar estructuralmente las interacciones entre dominios SH3 y péptidos ricos en prolina (Gmeiner *et al.*, 2001; Freund *et al.*, 2003). Gmeiner *et al.* han fusionado la región de la hélice de poliprolina II (PPII) de la proteína GAP humana al extremo C-terminal de dominio SH3 de Hck. Esta proteína de fusión, SH3_{Hck}:PPII_{GAP}, se organiza espontáneamente adoptando una estructura en la que la hélice de poliprolina II está unida intramolecularmente a la cavidad hidrofóbica del dominio SH3 (Gmeiner *et al.*, 2001). Por otro lado, Freund *et al.* han unido covalentemente la secuencia del péptido SHRPPPPGHRV de la proteína CD2 al extremo N-terminal del dominio GYF de CD2BP2 a través de una secuencia de conexión, y han demostrado por RMN que ésta permite la unión intramolecular del péptido al dominio (Freund *et al.*, 2003). La estructura es similar a la obtenida para el complejo formado por el dominio GYF y el péptido SHRPPPPGHRV (Freund *et al.*, 2002). Una de las ventajas de la unión covalente entre péptidos ricos en prolina y estos dominios es que no es necesario utilizar altas concentraciones de ligandos para

estudiar la interacción y permite obtener NOEs para el cálculo estructural entre proteínas y péptidos que presentan baja afinidad y/o solubilidad limitada.

Por otro lado, la unión intramolecular entre dominios SH3 y secuencias ricas en prolina posee gran importancia biológica, ya que juega un papel esencial en la regulación de muchos procesos de proteínas como, por ejemplo, en varias familias de tirosina quinasas. Estas interacciones intramoleculares son relativamente débiles y se sabe que los ligandos intramoleculares de los dominios SH3 de estas proteínas son habitualmente distintos de las secuencias de elevada afinidad para los dominios SH3 (Andreotti *et al.*, 1997; Nguyen, J. T. y Lim, 1997; Boggon y Eck, 2004). Además, se ha demostrado que la longitud de las secuencias de conexión entre los dominios SH3 y los péptidos ricos en prolina es esencial para determinar la competición entre la unión intra e intermolecular, como ocurre por ejemplo en las quinasas Rlk e Itk de familia de las Tec-quininas (Laederach *et al.*, 2003). Así, la comprensión a nivel cuantitativo de los efectos termodinámicos de la ligadura covalente entre dominios SH3 y sus ligandos peptídicos posee una importancia fundamental para entender en profundidad los factores que intervienen en la regulación de estos procesos biológicos.

1.3. Objetivos de la Tesis

En esta Tesis Doctoral se propone un abordaje original para un estudio de las interacciones SH3-péptido que consiste en la obtención, mediante ingeniería genética, de un conjunto de proteínas quiméricas de diseño que integrarán el dominio SH3 y la secuencia del ligando peptídico en una única cadena polipeptídica. El objetivo general de este trabajo es disponer de una herramienta para optimizar las interacciones en el sitio de unión de forma simple y eficiente mediante el estudio de la termodinámica de su desplegamiento combinado con técnicas de mutagénesis dirigida. La ventaja principal que presentan estas quimeras consiste en la posibilidad de evaluar los cambios en la termodinámica de las interacciones SH3-ligando mediante el estudio de su desplegamiento térmico por calorimetría diferencial de barrido (CDB). Esta técnica sería más rápida, directa, y a la vez menos costosa, que los experimentos de unión entre dominios SH3 y los ligandos peptídicos aislados mediante calorimetría isotérmica de titulación.

Este trabajo se ha enmarcado principalmente en el seno del proyecto de investigación BI003-04274 del Ministerio de Educación y Ciencia, que ha perseguido profundizar en la comprensión de los determinantes termodinámicos, dinámicos y estructurales de las interacciones proteína-ligando con el fin de incorporar este conocimiento en las metodologías actuales de diseño de ligandos.

Como objetivos específicos perseguidos en este trabajo se enumeran los siguientes:

1. Obtener mediante técnicas de ingeniería genética una serie de proteínas quiméricas recombinantes que incluyan en una única polipeptídica un dominio SH3 y su ligando peptídico rico en prolina.
2. Seleccionar el mejor diseño mediante la caracterización del desplegamiento de las proteínas quiméricas y de su conformación mediante calorimetría diferencial de barrido, espectroscopia de fluorescencia y dicroísmo circular.
3. Determinar la estructura tridimensional a alta resolución de la quimera seleccionada mediante RMN y analizar su dinámica conformacional.

4. Obtener una serie de mutantes Pro-Ala en la secuencia del ligando rico en prolina para analizar la importancia relativa de las diferentes Pro sobre la termodinámica de la unión mediante calorimetría diferencial de barrido, lo que nos permitirá evaluar la validez de nuestra metodología para el análisis termodinámico de las interacciones SH3-ligando.
5. Analizar en profundidad el mecanismo de plegamiento de la quimera mediante el estudio de la cinética del plegamiento utilizando la técnica de flujo detenido.

CAPÍTULO 2: DISEÑO Y CARACTERIZACIÓN PRELIMINAR DE PROTEÍNAS QUIMÉRICAS

CAPÍTULO 2: DISEÑO Y CARACTERIZACIÓN PRELIMINAR DE PROTEÍNAS QUIMÉRICAS

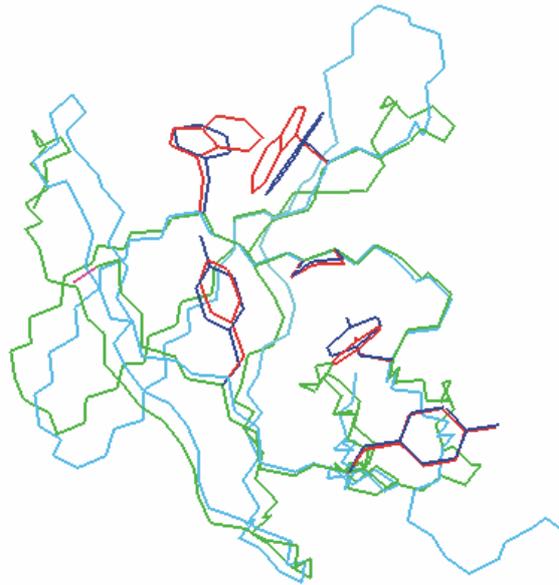
2.1. Diseño de quimeras basadas en la topología del permutante circular S19P20s

El diseño se realizó a partir de la distribución topológica de un permutante circular del dominio SH3 de la proteína α -espectrina (Spc-SH3), que fue diseñado y estudiado previamente en nuestro grupo de investigación (Martinez *et al.*, 1999). En este permutante circular, conocido como S19P20s, la cadena polipeptídica se ha abierto entre los residuos 19 y 20 ubicados en el lazo *RT* y los extremos N- y C-terminales se han conectado con una secuencia peptídica corta. Como se puede observar en la Figura 2.1.1 el dominio permutante S19P20s mantiene prácticamente la misma estructura tridimensional que el dominio natural, aunque presenta algunas diferencias en su estabilidad y en su velocidad de plegamiento (Viguera *et al.*, 1996; Martinez *et al.*, 1999).

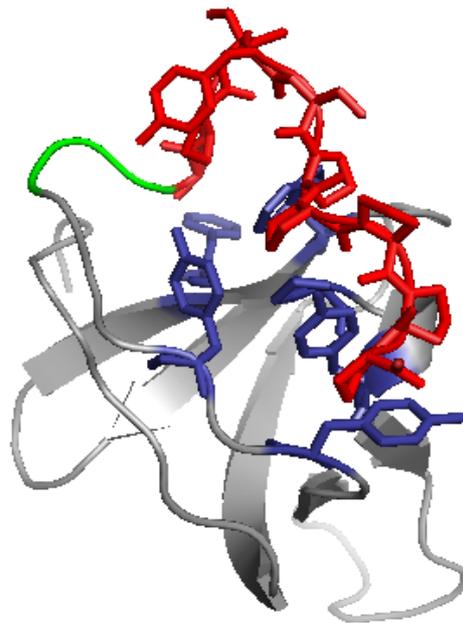
La ventaja de utilizar el permutante circular S19P20s como base del diseño en lugar del dominio SH3 natural se puede deducir fácilmente a partir del análisis de la estructura tridimensional de S19P20s y de los complejos SH3- péptido existentes en el PDB. Así, en la Figura 2.1.1 se observa que la localización espacial del nuevo extremo C-terminal de S19P20s es relativamente cercana al sitio de unión y, así, se podría elongar la cadena con una secuencia corta de conexión de dos o tres aminoácidos, que sería suficiente para enlazar con el ligando peptídico.

Para el diseño de esta construcción quimérica se utilizó como modelo de partida la estructura resuelta por cristalografía de rayos X del complejo formado entre el dominio SH3 de Abl y el decapeptido p41 (PDB:1BBZ) (Pisabarro *et al.*, 1998). Sobre la estructura de éste complejo se hizo una superposición de la estructura del permutante circular S19P20s (PDB:1TUC) (Viguera *et al.*, 1996), utilizando el programa Swiss PDB Viewer v3.7 (en su versión para PC, <http://expasy.org/spdbv/>). En la Figura 2.1.1 observamos que las estructuras son muy similares, siendo la desviación cuadrática media (RMSD) entre las coordenadas de la cadena principal de 1.37 Å. Al hacer la superposición se observa que el sitio de unión de ambos dominios SH3 está bastante bien conservado, concretamente, los restos Tyr7, Phe9, Trp36, Trp47, Pro49 y Tyr52 en Abl-SH3 (representados en rojo en la Figura 2.1.1.A) son equivalentes a Tyr57, Tyr59, Trp24, Phe35, Pro37 y Tyr40 en S19P20s (representados en azul en la Figura 2.1.1.A), siendo también su disposición espacial relativa muy similar. Por tanto, creamos un modelo del complejo entre el péptido p41 y el permutante circular S19P20s reemplazando el dominio de Abl-SH3, y probamos secuencias cortas de conexión entre ambos, de diferentes longitudes (de 2 o 3 restos) y composición (Figura 2.1.1.B). En las secuencias de conexión se emplearon aminoácidos polares o glicina, que se encuentran habitualmente en los lazos de las proteínas. Las distintas variantes se clonaron y expresaron en *E. Coli*. y se purificaron como se indica en las Secciones 6.1. y 6.4.1 de Materiales y Métodos.

A



B



MGPREVTMKK GDILTLLNST NKDWWKVEVN DRQGFVPAAY VKKLDSGTGK
ELVLALYDYQ EKS X APSYSPPPPP
X: DNG, GGG, GDN y DN

Figura 2.1.1. A. Superposición de las estructuras tridimensionales del dominio Abl-SH3 (azul) (PDB:1BBZ) (Pisabarro et al., 1998) y el permutante circular S19P20s (verde) (PDB:1TUC) (Viguera et al., 1996). Las cadenas laterales de los restos correspondientes al sitio de unión en ambas estructuras se han representado en color rojo (Abl-SH3) y en color azul (S19P20s). La flecha indica la localización del extremo C-terminal de S19P20s. **B.** Modelo de la estructura de la quimera SPCp41. La cadena representada en color gris corresponde al permutante circular. El péptido p41 está representado en color rojo y la secuencia de unión entre el péptido y el permutante está en color verde. En azul se muestran las cadenas laterales de los restos del permutante que interactúan con el péptido. Debajo de modelo de la estructura se muestra la secuencia de la quimera que se deseaba obtener, siendo X la secuencia de conexión entre el permutante y el péptido.

2.2. Caracterización preliminar de la estabilidad de las quimeras basadas en S19P20s

Una vez purificadas las cuatro variantes, se analizó su estabilidad térmica mediante calorimetría diferencial de barrido (CDB) en tampón glicina 20 mM a pH 3.5, comparándolas a su vez con la del permutante circular S19P20s bajo las mismas condiciones experimentales. Los termogramas de CDB se han representado en la (Figura 2.2.1) y la Tabla 2.2.1 recoge los parámetros de la transición de desplegamiento obtenidos directamente de las curvas, que son la temperatura de transición, T_m , y la entalpía de desnaturalización, ΔH_m . Todas las variantes son más termoestables que el permutante S19P20s en unos 5-8 °C, lo que indica la presencia de interacciones estabilizantes debidas a la extensión de la cadena.

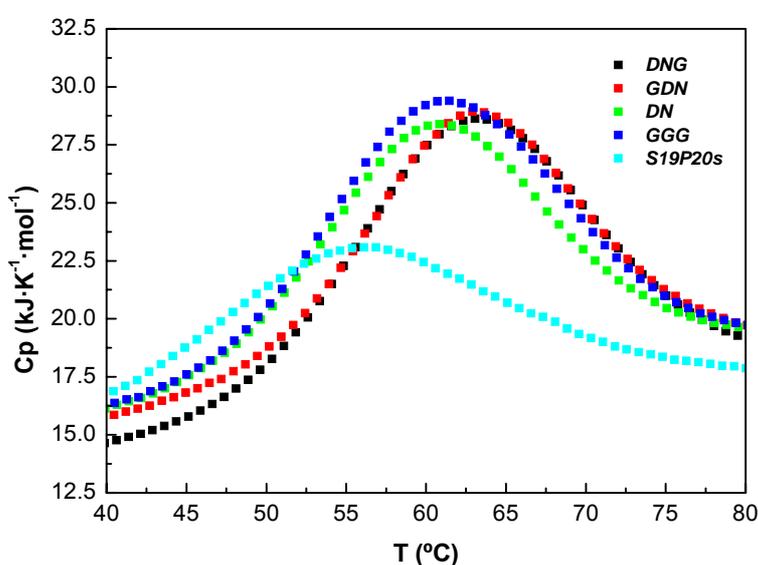


Figura 2.2.1. Capacidad calorífica molar parcial, C_p , en función de la temperatura para las distintas construcciones de las quimeras basadas en la topología del permutante circular S19P20s. En colores se muestran los termogramas obtenidos para las quimeras con distintas secuencias de conexión entre el permutante circular S19P20s y el péptido p41, y en turquesa aparece representado el termograma del permutante S19P20s.

Puesto que una de las principales finalidades de la construcción de estas quimeras es el uso de la CDB para evaluar la termodinámica de la interacción del dominio SH3 con el péptido p41 y, según estudios previos sobre la aplicabilidad de la CDB al análisis de interacciones proteína-ligando (Brandts y Lin, 1990; Martínez *et al.*, 1994; Martínez *et al.*, 1995), el valor de la T_m aumenta conforme la afinidad se hace mayor. De las cuatro variantes, la que resultó ser más estable térmicamente fue aquella cuya secuencia de unión estaba formada por los residuos **GDN**, y que presentaba además una delección de una lisina (K) en la posición 62, cerca del extremo C-terminal, realizada para aliviar posibles tensiones producidas por interacciones carga-carga en esta zona. Así, hemos seleccionado la quimera con la conexión GDN como objeto central de los estudios realizados en esta Memoria y en los siguientes Capítulos nos referiremos a ella, de manera genérica, como SPCp41.

Tabla 2.2.1. *Parámetros termodinámicos obtenidos directamente de los termogramas de las distintas quimeras según la secuencia de conexión y del permutante circular S19P20s.*

Proteína	T_m	ΔH_m
DNG	62.4	207.1
GDN	62.8	198.1
DN	59.6	195.8
GGG	60.5	204.6
S19P20s	54.8	153.8

2.3. Caracterización preliminar de la conformación de la quimera SPCp41 mediante técnicas espectroscópicas

2.3.1. Espectroscopia de fluorescencia

Con el fin de observar los cambios en el espectro de fluorescencia intrínseca de triptófano que produce la interacción entre el péptido p41 y el dominio permutante circular S19P20s, se realizó la titulación de S19P20s con p41 en tampón glicina 20 mM pH 3.0 y a 25°C, mediante espectroscopia de fluorescencia como se explica en la Sección 6.8.1 de Materiales y Métodos. Este experimento nos permitió además estimar la constante de disociación del complejo entre ambas moléculas.

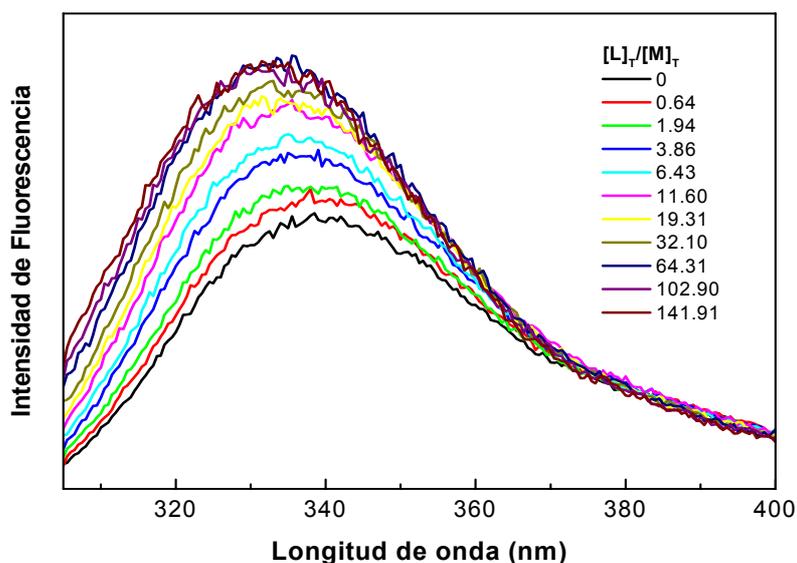


Figura 2.3.1. *Espectros de emisión de fluorescencia del permutante circular S19P20s normalizados por la concentración de proteína, libre (en negro) y en presencia de diferentes concentraciones (en colores) del decapeptido de diseño p41 (APSYSPPPP). El experimento de titulación se realizó por adición de volúmenes crecientes de disolución stock de p41 sobre una disolución madre de proteína, a pH 3.0 y 25°C de temperatura.*

En la Figura 2.3.1 se aprecia cómo, a medida que se añade ligando p41 a la muestra de proteína, la señal de fluorescencia se hace cada vez más intensa y al mismo tiempo se produce un desplazamiento del máximo de los espectros hacia menores longitudes de onda, lo que indica un enterramiento del triptófano desde un medio más polar (disolvente) hasta un medio más apolar (interfase de contacto péptido-proteína), que es justo lo que parece estar pasando con el triptófano en la posición 24 del permutante, el cual está situado en la superficie de unión. Este

fenómeno también se observa al unirse el péptido p41 al dominio Abl-SH3 (Pisabarro *et al.*, 1998) o al dominio Spc-SH3 (Casares *et al.*, 2007). Para obtener la curva de titulación, debido al desplazamiento del máximo de los espectros hacia el azul, en lugar de considerar intensidades de fluorescencia a una determinada longitud de onda, se consideraron las “áreas totales de fluorescencia”, obtenidas mediante integración entre 305 y 400 nm de cada uno de los espectros obtenidos en cada punto de la titulación, obteniéndose la curva de titulación típica que esta representada en la Figura 2.3.2.

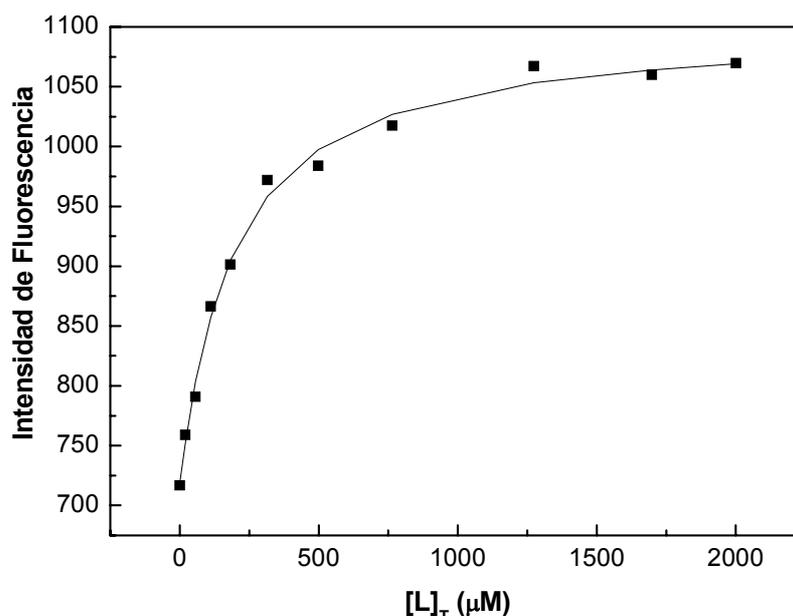


Figura 2.3.2. Curva de titulación del permutante circular S19P20s con p41. Experimento realizado por adición de volúmenes crecientes de disolución stock de p41 sobre una disolución madre de proteína. Los puntos representan datos de “área total de fluorescencia” normalizada por la concentración de proteína en cada punto y la línea continua representa el ajuste por mínimos cuadrados de dichos datos a la ecuación 6.8.4.

La constante de disociación del complejo S19P20s:p41 a pH 3.0, obtenida del experimento de titulación, resultó ser $K_d = 160 \pm 40 \mu\text{M}$, que es bastante mayor que la encontrada para el complejo entre Abl-SH3 y p41 (Palencia *et al.*, 2004), cuyo valor es $K_d = 5 \mu\text{M}$ y que la del complejo entre el mutante R21A de Spc-SH3 y p41 (Casares *et al.*, 2007), cuyo valor es $K_d = 55 \mu\text{M}$. Este resultado es lógico ya que en S19P20s parte del sitio de unión se ha alterado por el corte de la cadena en la zona del lazo RT, que juega un papel importante en la afinidad y especificidad de las interacciones (Roberts, 1999; Mayer, 2001; Wang, C. Y. *et al.*, 2001a; Wang, W. *et al.*, 2001b; Yuzawa *et al.*, 2001).

También se registraron los espectros de fluorescencia de la quimera SPCp41 y el permutante S19P20s aislado (Figura 2.3.3), en las mismas condiciones de los experimentos de titulación. La quimera SPCp41 se tituló también con volúmenes crecientes de péptido p41, observándose que los espectros de fluorescencia no se afectaban, lo que apoya nuestra hipótesis de que el segmento de cadena añadido a la quimera se encuentra ocupando de forma intramolecular el sitio de unión del dominio SH3 e impide la unión intermolecular del péptido p41 añadido. Este hecho también se aprecia si comparamos el espectro de SPCp41 con el de S19P20s saturado con péptido representado en la Figura 2.3.3. Ambos espectros son muy similares, lo que

indica que el entorno del triptófano 24, que se encuentra en la superficie del sitio de unión, es similar en ambos casos.

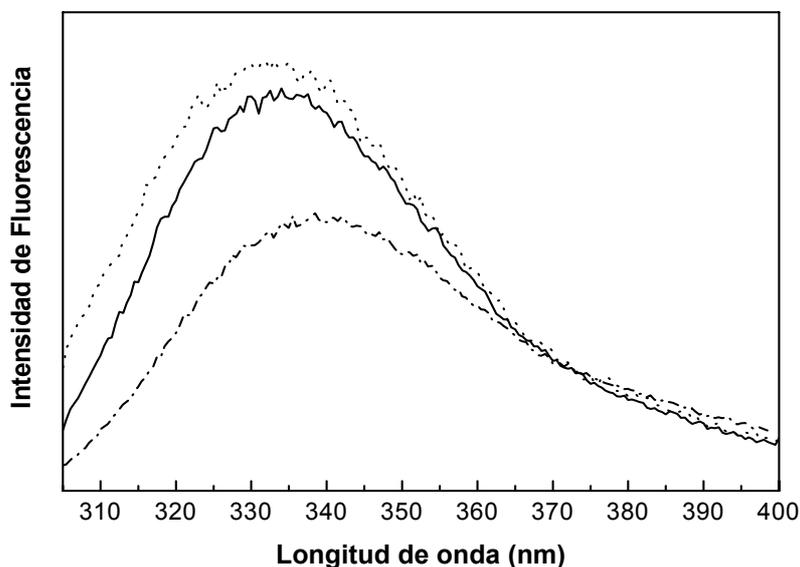


Figura 2.3.3. Espectros de fluorescencia de la quimera SPCp41 (línea continua), del permutante circular S19P20s (línea discontinua) y del permutante en presencia de 1.5 mM de p41 (línea punteada), todos ellos en las mismas condiciones (pH 3.0 y 25 °C).

2.3.2. Dicroísmo circular

Los experimentos de dicroísmo circular (DC) nos permitieron caracterizar la conformación de la quimera SPCp41, comparándola con las conformaciones que presentan S19P20s y el péptido libre p41. Se realizaron para ello los experimentos en condiciones nativas, es decir, en tampón glicina 20 mM pH 3.0, a 25°C, tanto en la región del ultravioleta lejano como en la del cercano, a las concentraciones adecuadas de proteína según la zona del ultravioleta en que se registraron los espectros.

En el ultravioleta lejano, como se aprecia en la Figura 2.3.4, el espectro de DC de S19P20s es muy similar al del dominio natural Spc-SH3, lo cual es lógico atendiendo a la similitud de sus estructuras (Musacchio *et al.*, 1992b; Viguera *et al.*, 1996), mientras que el espectro de SPCp41 presenta una señal de DC negativa mayor, como consecuencia de la presencia del péptido p41 en su secuencia. La diferencia entre los espectros de DC de SPCp41 y S19P20s, una vez normalizados por mol de proteína, es representativa del espectro de DC del segmento de cadena adicional en SPCp41. Este espectro posee las características típicas de los espectros de los péptidos ricos en prolina que adoptan una conformación en hélice de poliprolina II (una banda positiva a 228 nm y una banda negativa a 204 nm) (Manning y Woody, 1991; Rabanal *et al.*, 1993; Viguera *et al.*, 1994a). Esto indica, por tanto, que el segmento p41 en SPCp41 posee un alto grado de hélice de poliprolina II, lo que coincide con la conformación que adopta dicho péptido en el complejo con Abl-SH3 (Pisabarro *et al.*, 1998) y con Spc-SH3 (Casares *et al.*, 2007). Estas características no se aprecian en el espectro de DC del péptido p41 aislado, que muestra un espectro típico de cadena polipeptídica desordenada.

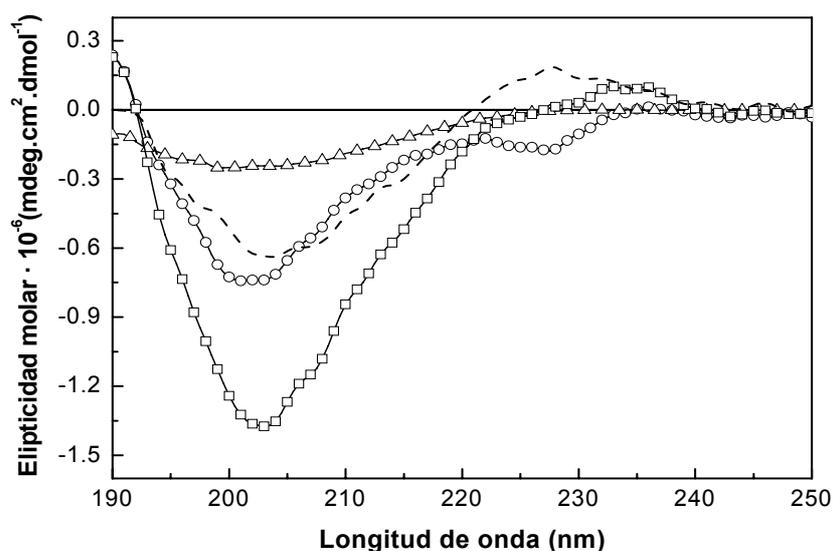


Figura 2.3.4. Espectros de DC de la región del ultravioleta lejano de la quimera SPCp41 (cuadrados), del permutante S19P20s (círculos) y del péptido p41 aislado (triángulos), en glicina 20 mM a pH 3.0 a 25°C. La línea discontinua representa la diferencia entre el espectro de SPCp41 y el de S19P20s.

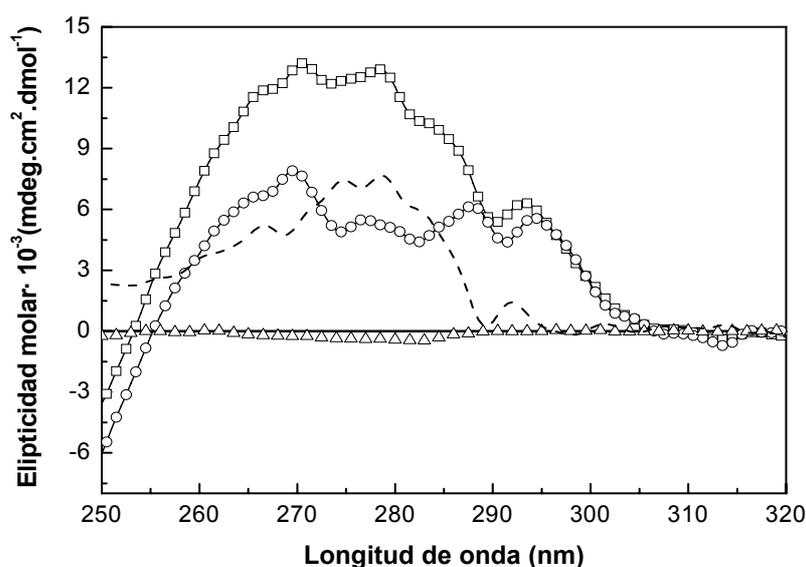


Figura 2.3.5. Espectros de DC en la región de ultravioleta cercano de SPCp41 (cuadrados), de S19P20s (círculos) y del péptido p41 aislado (triángulos), en glicina 20 mM pH 3.0 a 25°C. La línea discontinua representa la diferencia entre los espectros de la SPCp41 y el S19P20s.

Por otro lado, en el espectro de DC en el ultravioleta cercano (Figura 2.3.5), SPCp41 presenta una señal claramente mayor que S19P20s, siendo la diferencia más acusada alrededor de 278 nm, lo cual indica la inmovilización de cadenas laterales de alguna(s) tirosina(s) (ya que aproximadamente esa es la longitud de onda a la que absorben radiación ultravioleta) debida a la interacción con la secuencia de p41. A 294 nm también se observa una diferencia de señal, aunque más pequeña, entre el espectro de SPCp41 y el de S19P20s. Esta diferencia, por la zona del espectro en la que aparece, nos sugiere cambios en el entorno de la cadena lateral de Trp26, que se encuentra en la superficie del sitio de unión. Estas observaciones están de acuerdo

con las estructuras de los complejos entre el mismo péptido y cada uno de los dominios Abl-SH3 (Pisabarro *et al.*, 1998) y Spc-SH3 (Casares *et al.*, 2007).

Por tanto, los resultados obtenidos indican que la proteína quimérica SPCp41 parece imitar las características estructurales de los procesos de unión entre el los dominios SH3 y el ligando p41, con una conformación de hélice de poliprolina II para p41 y un enterramiento e inmovilización de los restos aromáticos en la interfase de unión similar a la que se observa típicamente en la unión de ligandos ricos en prolina con otros dominios SH3.

2.4. Diseño de proteínas quiméricas con otras topologías

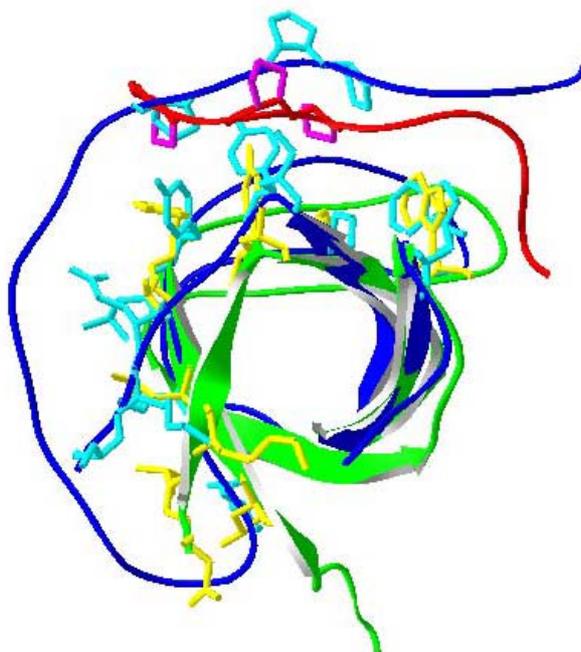
Se diseñaron otras proteínas quiméricas a partir de la topología original del dominio SH3 de la proteína α -espectrina (Spc-SH3), con el propósito de unir este dominio, a través de una secuencia de conexión al péptido p41. Esta nueva topología presenta la ventaja de que mantiene el lazo *RT* intacto, lo que es importante para imitar las interacciones con el ligando de forma más completa, aunque presenta la dificultad adicional de necesitar una mayor longitud en la secuencia de unión, lo que dificulta su diseño.

Para el diseño de esta construcción quimérica, se utilizó como modelo de partida la estructura determinada por RMN del complejo intramolecular que forma el dominio SH3 natural de la proteína Itk con el péptido rico en prolina KPL, de secuencia KKPLPPTP (Andreotti *et al.*, 1997). Inicialmente se llevó a cabo la superposición de las estructuras del complejo Itk-SH3 (PDB:1AWJ), la del dominio Spc-SH3 (PDB:1SHG) y la del péptido p41 en su conformación en el complejo con Abl-SH3 (PDB:1BBZ), utilizando el programa Swiss PDB Viewer v3.7. En la Figura 2.4.1 hemos representado Itk-SH3 en color azul con las cadenas laterales de los residuos que presentan contactos con la secuencia de conexión entre el dominio SH3 y el péptido natural en color turquesa. Spc-SH3 se ha representado en color verde y con las cadenas laterales equivalentes en color amarillo y, por último, el péptido p41 se ha representado en color rojo, con las prolinas que se conservan en el péptido natural de Itk-SH3 en color magenta. Como se observa en dicha figura las estructuras son similares, siendo la desviación cuadrática media (RMSD) entre las coordenadas de las cadenas principales de 1.63 Å. En la superposición se observa que el sitio de unión para la secuencia rica en prolina está bastante conservado. Del análisis de la estructura de Itk-SH3 se deduce que los contactos más claros con la secuencia de conexión aparecen entre las cadenas laterales de Phe17, Ile26 y Leu28, y las cadenas laterales de Glu77, Lys 78 y Ser79 del extremo C-terminal.

Decidimos clonar y expresar dos proteínas quiméricas entre Spc-SH3 y p41 con este tipo de topología; la primera usando la misma secuencia de conexión (EDNRRSFQEPEETLVI) que aparece en la unión intramolecular del péptido KPL al dominio SH3 de Itk. En lugar de la forma natural de Spc-SH3 se utilizó el mutante R21A, puesto que este mutante posee una mayor afinidad por el péptido p41 que el dominio natural (Casares *et al.*, 2007). A partir de ésta primera construcción se realizó un segundo diseño en el que se sustituyó el extremo C-terminal de Spc-SH3 (VKKLD) por el de Itk-SH3 (LVEKS), ya que en el análisis de la estructura se aprecian muchos contactos que se encuentran presentes entre la secuencia de conexión y este extremo

Una vez clonadas, expresadas y purificadas estas dos variantes quiméricas del dominio Spc-SH3, se analizó su desplegamiento térmico por CDB. Ambas quimeras

presentan una transición de desplegamiento casi idéntica a la del mutante R21A de Spc-SH3, tal y como podemos observar en la Figura 2.4.2 para una de ellas, lo que indica que ninguna de las dos proteínas quiméricas muestra una estabilización adicional con respecto al dominio independiente debido a la ausencia de interacción péptido-dominio en dichas quimeras. Estos resultados sugieren que en estas quimeras el segmento p41 y la secuencia de conexión se encuentran desordenados y no influyen en la estabilidad térmica del dominio SH3.



1AWJ	3	KKPL	PPT	PE	NRRSFQ	EE	TL	VIALYDYQ	TNDPQ	ELALR	CDEEYLLDS
1SHG	6							K	EL	VLALYDYQ	EKSPREVTMK
1BBZ	1	APSYSP	PP	PP							
1AWJ	53	SEIHW	WRVQD	KNGHEGYAPS	SYLVEKS						
1SHG	37	TNKDW	WKVEV	NDR-QGFVPA	AYVKKLD						

Figura 2.4.1. Superposición de las estructuras tridimensionales de ItkSH3-KPL (azul) (PDB:1AWJ) con los residuos que presentan contactos con la secuencia de conexión entre el péptido y el dominio SH3 en turquesa, de Spc-SH3 (verde) (PDB:1SHG), con los residuos equivalentes a los representados en color turquesa en color amarillo y del péptido p41 (rojo) con las Pro que se conservan en ambos péptidos en magenta. También se muestra el alineamiento de las secuencias en el mismo código de colores que aparece en la estructura.

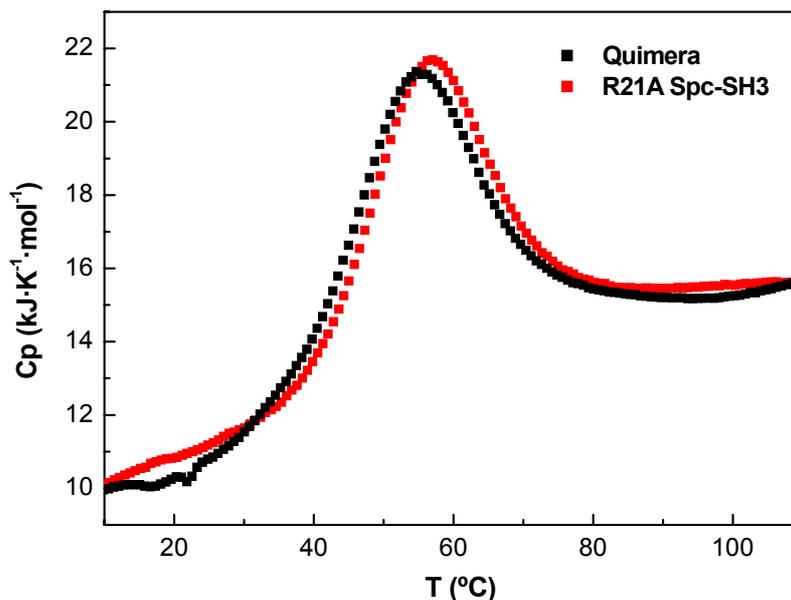
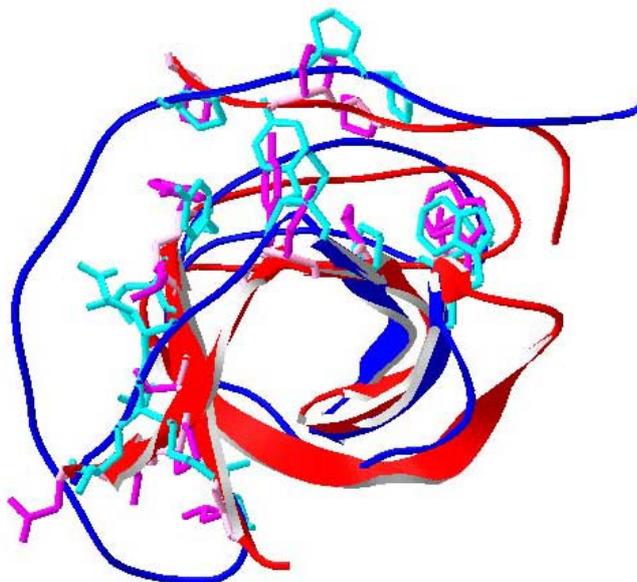


Figura 2.4.2. Capacidad calorífica molar parcial, C_p , en función de la temperatura para la quimera Spc-SH3 basada en Itk-SH3. En color negro se muestra el termograma obtenido para la quimera, y en rojo aparece representado el termograma del mutante R21A Spc-SH3.



1AWJ	3	KKPLPPTPED	NRRSFQEPEE	TLVIALYDYQ	TNDPQELALR	CDEEYLLDS
1BBZ	1			NLFVALYDFV	ASGDNTLSIT	KGEKLRVLGY
1BBZ	1	APSYSPPPPP				
1AWJ	53	SEIHVWRVQ	DKNGHEGYAP	SSYLVEKS		
1BBZ	31	NHNGEWCEAQ	TKNG-QGWVP	SNYITPVNS		

Figura 2.4.3. Superposición de las estructuras tridimensionales de ItkSH3-KPL (azul) (PDB:1AWJ), del complejo formado por Abl-SH3:p41 (rojo) (PDB:1BBZ). En turquesa se muestran los residuos que presentan contactos con la secuencia de conexión entre el péptido KPL y el dominio Itk-SH3, y en magenta se muestran los residuos de Abl-SH3:p41 equivalentes a los representados en turquesa. También se muestra el alineamiento de las secuencias en el mismo código de colores que aparece en la estructura.

A la vista de los resultados obtenidos, se diseñó la misma construcción utilizando el dominio Abl-SH3, ya que este dominio presenta una afinidad por p41 mucho mayor ($K_d=5 \mu\text{M}$) (Palencia *et al.*, 2004) que la observada para el mutante R21A de Spc-SH3 ($K_d=55 \mu\text{M}$) (Casares *et al.*, 2007). Se siguió la misma estrategia que para el diseño de las anteriores quimeras, usando como modelos la estructura del complejo entre Abl-SH3 y p41 (PDB:1BBZ) (Pisabarro *et al.*, 1998) y la del dominio Itk-SH3 (PDB:1AWJ) (Andreotti *et al.*, 1997). En la Figura 2.4.3 se representa la superposición de ambas estructuras donde hemos representado Itk-SH3 en color azul y los residuos que presentan contactos con la secuencia de conexión en turquesa, Abl-SH3:p41 (PDB:1BBZ) en color rojo con los residuos equivalentes a los que presentan contactos con la secuencia de conexión en Itk-SH3 en color magenta. Al igual que en el caso anterior, las estructuras son muy parecidas, siendo el RMSD entre ambas de 1.77 Å. De nuevo se eligió mantener la misma secuencia de conexión que aparece en el complejo intramolecular Itk-SH3:KPL.

Una vez clonadas, expresadas y purificadas estas dos nuevas construcciones, los experimentos de desplegamiento térmico realizados por CDB no muestran transición alguna, lo que sugiere que las proteínas quiméricas se encuentran desplegadas en condiciones nativas. Por tanto, hasta la fecha no hemos conseguido diseñar una proteína quimérica con la topología de Itk-SH3:KPL que presente en su estructura la interacción intramolecular entre el dominio SH3 y el péptido p41. Es posible que la secuencia de conexión utilizada sea demasiado corta o no se adapte correctamente a los dominios Spc-SH3 y Abl-SH3, estableciendo interacciones desfavorables con estos que impidan la interacción intramolecular o, incluso, el correcto plegamiento del dominio SH3. También existe la posibilidad de que el coste entrópico del ordenamiento intramolecular del ligando y la secuencia de conexión sea demasiado alto, dada su gran longitud, no encontrándose compensado por interacciones favorables con los dominios SH3, que sí pudieran encontrarse optimizadas en Itk-SH3. En cualquier caso, responder a estas preguntas requeriría realizar nuevas construcciones con secuencias de conexión de longitud y composición diferentes. En este punto decidimos concentrar nuestros esfuerzos en el estudio de la quimera SPCp41 para profundizar en su estructura, sus propiedades termodinámicas y su utilizad como herramienta para el análisis de las interacciones SH3-ligando, como se describe en adelante.

CAPÍTULO 3:
DETERMINACIÓN DE LA ESTRUCTURA
TRIDIMENSIONAL DE LA QUIMERA SPC_{p41}

CAPÍTULO 3: DETERMINACIÓN DE LA ESTRUCTURA TRIDIMENSIONAL DE LA QUIMERA SPCp41

Para continuar nuestra investigación era esencial resolver la estructura tridimensional a alta resolución de la quimera SPCp41 ya que ésta nos permitiría una interpretación adecuada de las magnitudes termodinámicas observadas y ayudaría en el diseño futuro de mejores interacciones en el sitio de unión. Además, la estructura nos servirá para comprobar en qué grado el plegamiento de la quimera imita las características estructurales de los complejos formados por dominios SH3 y péptidos ricos en prolina. En este Capítulo se describe la determinación de la estructura de SPCp41 mediante RMN heteronuclear multidimensional utilizando muestras de SPCp41 enriquecidas isotópicamente con ^{13}C y ^{15}N . Este trabajo se realizó en gran parte durante dos estancias de tres meses en el Instituto de Química Física “Rocasolano” del Consejo Superior de Investigaciones Científicas (C.S.I.C.) de Madrid, bajo la supervisión de la Dra. Marta Bruix, en cuyo laboratorio se registraron la mayor parte de los espectros de RMN utilizados. La asignación de los espectros y la determinación de la estructura se realizó con la supervisión del Dr. Nico van Nuland.

3.1. Asignación de los espectros de RMN de SPCp41

Las asignaciones de los desplazamientos químicos (o frecuencias de resonancia) de los núcleos ^1H , ^{15}N , y ^{13}C de la proteína SPCp41 se llevaron a cabo utilizando espectros registrados con muestras de $^{15}\text{N},^{13}\text{C}$ -SPCp41, ^{15}N -SPCp41 y SPCp41 en glicina-D5 20 mM, pH 3.5 a 30°C, siguiendo el protocolo que se describe a continuación (véase también el la Sección 6.9.3 de Materiales y Métodos).

En primer lugar, se realizó la asignación secuencial de los desplazamientos químicos de los núcleos de la cadena principal de la proteína (protones ^1HN amida, nitrógenos H^{15}N amida y carbonos ^{13}CO carbonílicos y $^{13}\text{C}_\alpha$), junto con los carbonos $^{13}\text{C}_\beta$ de las cadenas laterales, mediante el análisis conjunto del espectro bidimensional (2D) $^1\text{H},^{15}\text{N}$ -HSQC y de los espectros tridimensionales (3D) de triple resonancia CBCANH, CBCA(CO)NH, HNCA, HN(CO)CA y HNCO. Este análisis permite la asignación secuencial de toda la cadena polipeptídica, salvo para núcleos que presentan solapamiento de señales y en los casos de las conectividades a prolinas, pues éstas carecen de protones amida. La Figura 3.1.1 muestra la base para la estrategia de asignación. Si nos fijamos, por ejemplo, en las dos primeras tiras extraídas de los espectros CBCANH (con fondo azul) y CBCA(CO)NH (con fondo naranja) para el mismo par de frecuencias ^1H - ^{15}N amida de un resto i , observamos 4 picos de cruce en el CBCANH: los dos picos positivos (en negro) corresponden a los núcleos $\text{C}_{\alpha i}$ y $\text{C}_{\alpha i-1}$, y los dos picos negativos (en rojo) corresponden a $\text{C}_{\beta i}$ y $\text{C}_{\beta i-1}$. En cambio, en la tira equivalente del CBCA(CO)NH sólo son visibles los picos de cruce correspondientes al $\text{C}_{\alpha i-1}$ y $\text{C}_{\beta i-1}$, lo que nos permite identificar las frecuencias de C_α y C_β del residuo i . Dos picos de cruce a exactamente esas dos frecuencias deberán buscarse en otra tira del espectro tridimensional correspondiente a otro par de frecuencias ^1H - ^{15}N amida que pertenecerá al residuo $i+1$ (tercera y cuarta tiras en la Figura 3.1.1). De esta forma se va avanzando a través del espectro tridimensional hasta completar la asignación de la cadena principal de la proteína. Las frecuencias de los ^{13}CO de cada residuo de la proteína se extrajeron del espectro HNCO a cada uno de pares de frecuencias ^1H - ^{15}N amida del mismo residuo previamente asignadas.

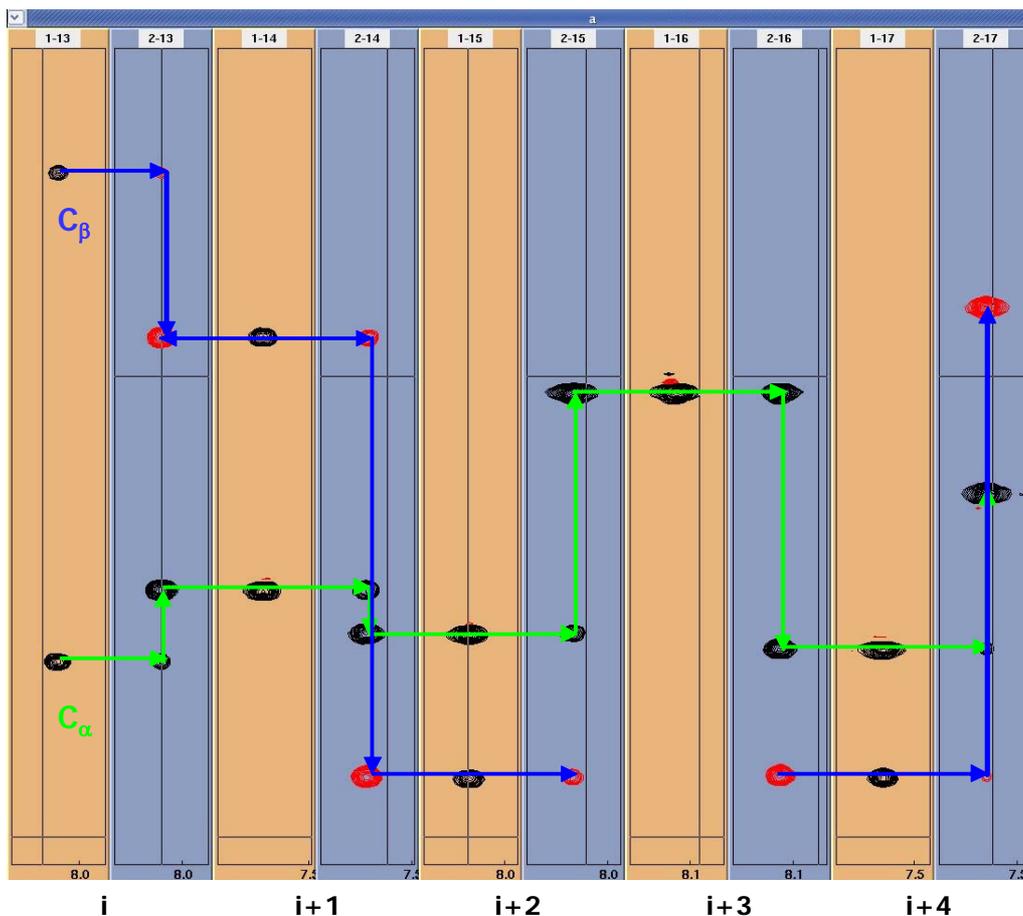


Figura 3.1.1. Tiras de los espectros CBCA(CO)NH (naranja) y CACBNH (azul) centradas en las frecuencias ^1H - ^{15}N de los picos de cruce del espectro ^1H - ^{15}N HSQC. Las flechas indican la asignación secuencial de los residuos a través de los carbonos C_α (flechas en verde) y C_β (flechas en azul).

En la Figura 3.1.2 podemos ver el espectro ^1H , ^{15}N -HSQC de SPCp41 a pH 3.5 en el que cada pico de cruce correspondiente a un par ^1H - ^{15}N amida del esqueleto se ha identificado con el número de residuo al que pertenece. En el espectro se observan todas las señales de correlación esperadas de acuerdo con la secuencia de la proteína. Esta asignación de la cadena principal es la base para la asignación de las frecuencias de ^1H , ^{15}N y ^{13}C de las cadenas laterales de los restos correspondientes usando una combinación de los espectros 2D y 3D descritos en la Sección 6.9.1 de Materiales y Métodos. Para la asignación de las cadenas laterales alifáticas empleamos los espectros 3D ^{15}N -TOCSY-HSQC, ^{15}N -NOESY-HSQC y HCCH-TOCSY. Del 3D ^{15}N -TOCSY-HSQC extraímos tiras centradas en el par de frecuencias ^1H - ^{15}N amida de cada aminoácido. Estas tiras contienen picos de cruce para las frecuencias de todos los protones acoplados escalarmenete de cada cadena lateral. A partir de estas señales pudimos generalmente asignar las frecuencias de los protones H_α , H_β , H_γ , H_δ y/o H_ϵ de cada cadena lateral de cada residuo. Con el HCCH-TOCSY es posible observar picos de cruce para el sistema de espín completo de cada aminoácido en cada par de frecuencias C_i - H_i , de forma que podemos identificar las frecuencias de C_γ , C_δ y/o C_ϵ y de los protones H_γ , H_δ y/o H_ϵ . Con el ^{15}N -NOESY-HSQC comprobamos los NOEs secuenciales que se producen entre núcleos de restos vecinos y asignamos las frecuencias de los restos que presentaban superposición de señales o de aquellos donde la transferencia de magnetización en los experimentos de tipo TOCSY no fue lo

suficientemente efectiva. Las frecuencias cada sistema de espín de protones de los grupos aromáticos se asignaron con los espectros 2D TOCSY y 2D NOESY en D₂O observando las señales de correlación entre dichos protones aromáticos. Posteriormente se asignó su pertenencia a un determinado residuo mediante la identificación de los NOEs entre los protones H_δ del anillo aromático y H_β del mismo residuo en el 2D NOESY.

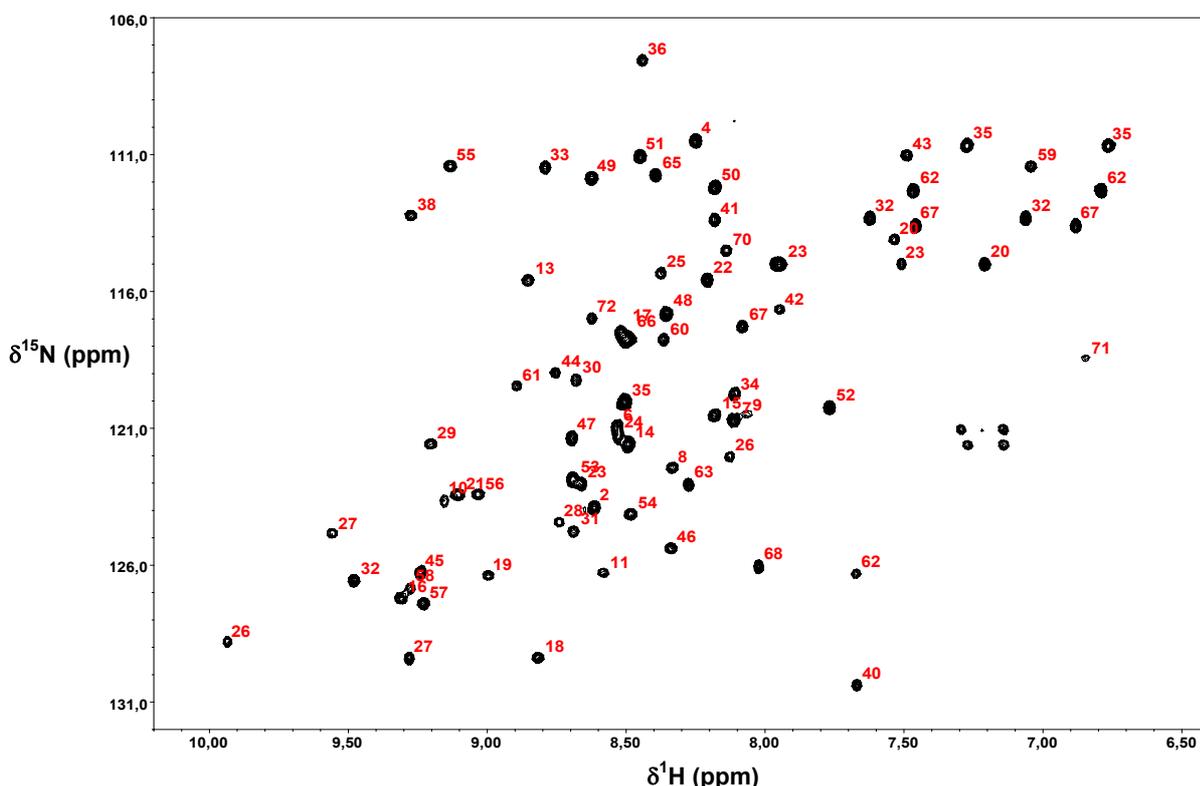


Figura 3.1.2. Espectro ^1H - ^{15}N HSQC de la proteína SPCp41 marcada en ^{15}N (glicina-D5 20 mM pH 3.5, 25°C). Los picos de cruce ^1H - ^{15}N amida están identificados con el correspondiente número de aminoácido en la secuencia.

Se ha conseguido realizar la asignación prácticamente completa de todos los desplazamientos químicos de ^1H , de todos los desplazamientos químicos de ^{15}N y de los desplazamientos químicos de los ^{13}C alifáticos (ver Tabla A.1 en el Apéndice I), mientras que los ^{13}C aromáticos no se asignaron. Estos desplazamientos químicos sirvieron de entrada para el programa CYANA que se utilizó posteriormente para el cálculo de la estructura como se describe más adelante.

3.2. Estructura secundaria de la quimera SPCp41

La estructura secundaria de las proteínas puede identificarse a partir de los desplazamientos conformacionales de los protones $^1\text{H}_\alpha$, de los carbonos $^{13}\text{C}_\alpha$ y $^{13}\text{C}_\beta$, así como de los patrones de NOEs secuenciales (entre núcleos de restos i e $i+1$) y de medio alcance (entre núcleos de restos i e $i+2$, $i+3$ ó $i+4$). Los desplazamientos conformacionales de $^1\text{H}_\alpha$, $^{13}\text{C}_\alpha$, $^{13}\text{C}_\beta$ y ^{13}CO se definen para cada tipo de aminoácido como la diferencia entre el desplazamiento químico observado experimentalmente y su valor en ausencia de estructura ordenada ($\Delta\delta = \delta_{\text{observado}} - \delta_{\text{sin estructura}}$, ppm). Se han

tomado como valores de referencia los desplazamientos químicos medidos para cada aminoácido en hexapéptidos sin estructura (Wishart *et al.*, 1995). Los restos en hélices α se caracterizan por tener desplazamientos conformacionales negativos para los protones $^1\text{H}_\alpha$ y los carbonos $^{13}\text{C}_\beta$, y positivos para los carbonos $^{13}\text{C}_\alpha$ y los carbonos ^{13}CO , mientras que los desplazamientos conformacionales característicos de los restos en hebras β son positivos para los protones $^1\text{H}_\alpha$ y los carbonos $^{13}\text{C}_\beta$, y negativos para los carbonos $^{13}\text{C}_\alpha$ y los carbonos ^{13}CO . En cuanto a los NOEs secuenciales, las hélices α presentan patrones característicos de NOEs de medio alcance entre el protón α de un resto i y los protones amida de los restos $i+3$ e $i+4$ ($d_{\alpha\text{N}(i,i+3)}$, $d_{\alpha\text{N}(i,i+4)}$) y el protón β del resto $i+3$ ($d_{\alpha\beta(i,i+3)}$), junto a NOEs secuenciales entre protones amida ($d_{\text{NN}(i,i+1)}$) intensos y débiles o incluso ausentes entre el protón α del resto i y el amida del resto $i+1$ ($d_{\alpha\text{N}(i,i+1)}$). En las hebras β , los NOEs secuenciales $d_{\text{NN}(i,i+1)}$ no se observan o son muy débiles y los $d_{\alpha\text{N}(i,i+1)}$ son muy intensos (Wüthrich, 1986).

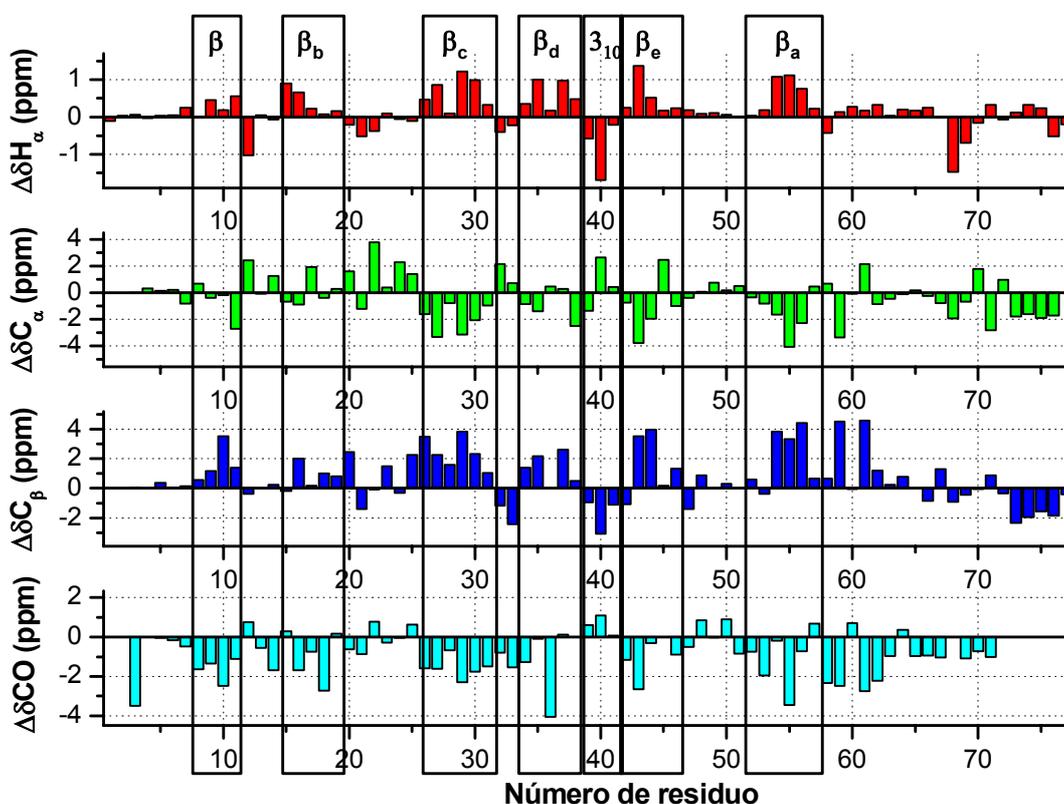


Figura 3.2.1. Desplazamientos conformacionales de los núcleos $^1\text{H}_\alpha$ ($\Delta\delta\text{H}_\alpha = \delta\text{H}_\alpha^{\text{observado}} - \delta\text{H}_\alpha^{\text{desplegado}}$, ppm) en color rojo, $^{13}\text{C}_\alpha$ ($\Delta\delta\text{C}_\alpha = \delta\text{C}_\alpha^{\text{observado}} - \delta\text{C}_\alpha^{\text{desplegado}}$, ppm) en color verde, $^{13}\text{C}_\beta$ ($\Delta\delta\text{C}_\beta = \delta\text{C}_\beta^{\text{observado}} - \delta\text{C}_\beta^{\text{desplegado}}$, ppm) en color azul y ^{13}CO ($\Delta\delta\text{CO} = \delta\text{CO}^{\text{observado}} - \delta\text{CO}^{\text{desplegado}}$, ppm) en color turquesa, en función de la secuencia de la proteína, a pH 3.5 y 25°C. Los valores de $\delta\text{H}_\alpha^{\text{desplegado}}$, $\delta\text{C}_\alpha^{\text{desplegado}}$, $\delta\text{C}_\beta^{\text{desplegado}}$ y $\delta\text{CO}^{\text{desplegado}}$, se tomaron de (Wishart *et al.*, 1995). Los elementos de estructura secundaria aparecen numerados de acuerdo con la numeración que corresponde a la forma natural del dominio SH3 de α -espectrina.

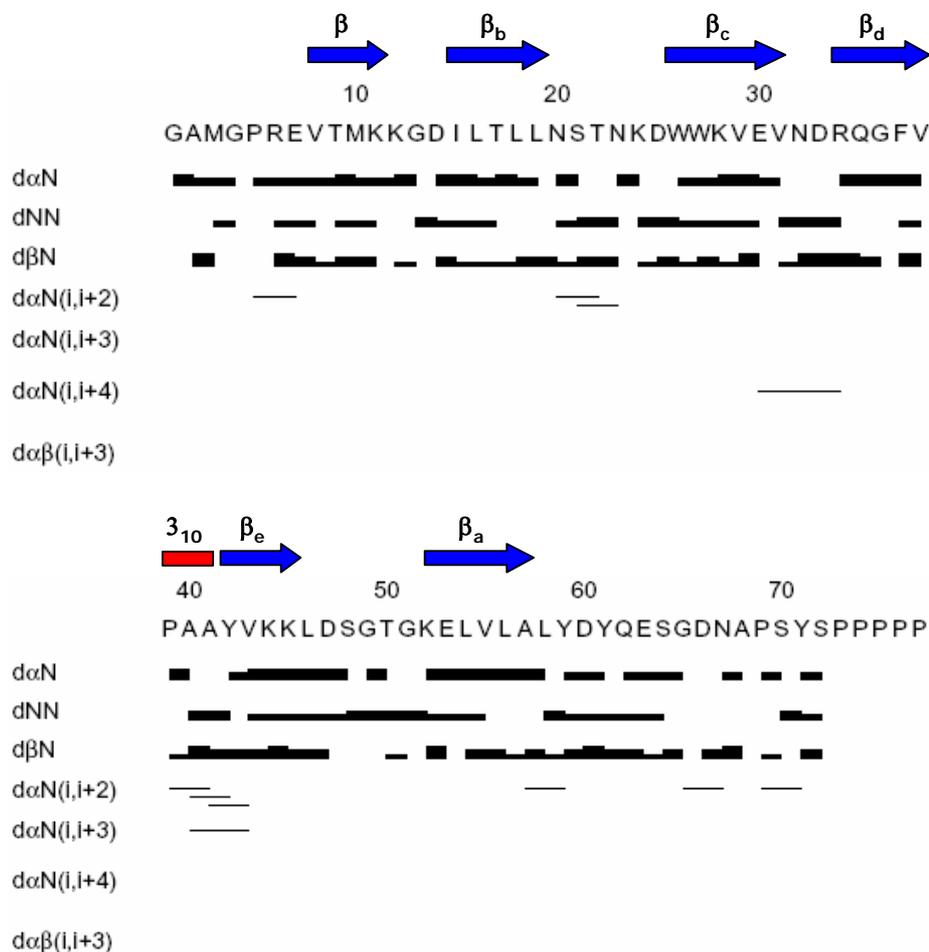


Figura 3.2.2. Resumen de NOEs secuenciales de la quimera SPCp41 en función del número de residuo en la secuencia. También se indican los elementos de estructura secundaria: con flechas azules las hebras β y con una barra roja la región de la hélice 3₁₀.

De acuerdo con los criterios expuestos, en la quimera SPCp41 se identifica la presencia de 6 hebras β que comprenden los aminoácidos 8-11, 15-19, 26-31, 34-38, 44-46 y 52-57 (Figuras 3.2.1 y 3.2.2) y de una pequeña hélice 3₁₀ formada por los aminoácidos 39-41 (Figuras 3.2.1 y 3.2.2). La disposición topológica de las hebras β , que puede deducirse a partir de las correlaciones NOE de largo alcance $d_{\alpha\alpha(i,j)}$, $d_{\alpha N(i,j)}$ y $d_{NN(i,j)}$, corresponde a dos láminas β antiparalelas (Figura 3.2.3). Los elementos de estructura secundaria se establecieron con mayor precisión una vez realizado el cálculo de la estructura tridimensional de SPCp41, como veremos en la siguiente Sección.

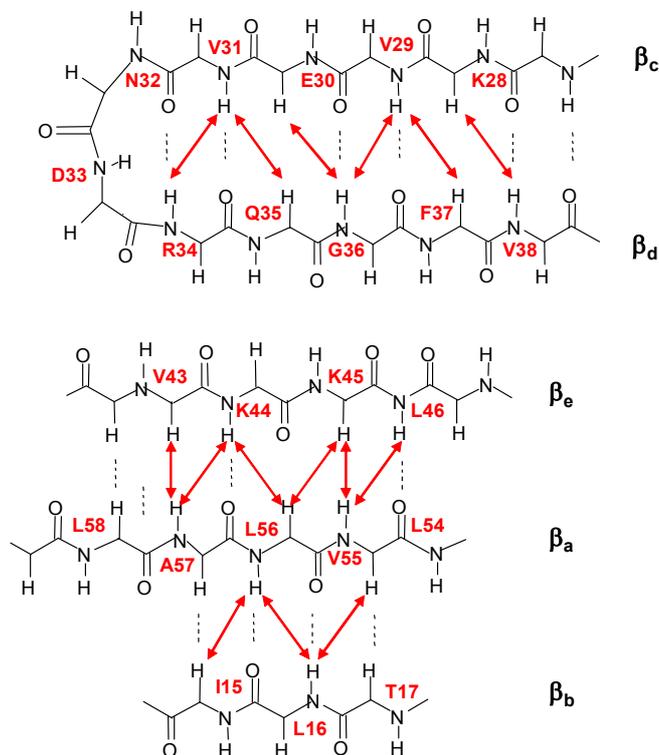


Figura 3.2.3. Diagrama esquemático de la topología de las láminas β antiparalelas de SPCp41. Una flecha entre protones indica una correlación NOE no secuencial observada.

3.3. Cálculo de la estructura tridimensional de SPCp41

La estructura tridimensional en disolución de SPCp41 se calculó tal y como se describe en la Sección 6.9.4 de Materiales y Métodos. Se obtuvieron 1581 restricciones de distancia entre pares de núcleos a partir de los picos de cruce presentes en el espectro 2D NOESY registrado en 10% D₂O:90% H₂O. El volumen de los picos de cruce del espectro 2D NOESY se midió utilizando la subrutina de integración del programa SPARKY (Goddard y Kneller). También se obtuvo un conjunto de restricciones para los ángulos diedros ϕ y ψ de la cadena principal a partir de los desplazamientos químicos de los protones ¹HN, ¹H _{α} , de los nitrógenos H¹⁵N y de los carbonos ¹³C _{α} , ¹³C _{β} y ¹³CO usando el programa PREDITOR (Berjanskii *et al.*, 2006) (http://wishart.biology.ualberta.ca/shifto/cqi-bin/predictor_current.py). El número de ángulos ϕ y ψ restringidos en el cálculo fue de 40 y de 38 respectivamente. Este número incluye a todos los ángulos diedros cuyo intervalo de confianza de predicción fue superior a 0.7, con excepción de las prolinas (que no poseen grupo HN) y las glicinas.

Para el cálculo de la estructura se utilizó el programa CYANA 2.0 (Combined assignment and dYnamics Algorithm for NMR Applications), que usa una combinación de cálculo de estructuras por geometría de distancias, dinámica molecular en ángulos de torsión y asignación automática de los espectros NOESY (<http://quentert.gsc.riken.go.jp/Software/Cyana.html>) (Herrmann *et al.*, 2002). Dentro

del programa CYANA se utilizó un protocolo de asignación semiautomática de las señales NOE entre pares de núcleos basada en la estructura (Herrmann *et al.*, 2002). Este protocolo parte de la asignación de los desplazamientos químicos de la mayoría de los núcleos de la proteína y de las listas de todas las señales NOE presentes en el espectro NOESY y realiza un proceso iterativo de siete ciclos de cálculo de estructuras asignando las señales NOE a determinados pares de núcleos. Después de cada ciclo el programa usa cada estructura generada para recalcular un nuevo conjunto de restricciones de distancia y lo compara con los NOEs experimentales corrigiendo su asignación. El proceso iterativo transcurre hasta que se asignan más del 95% de las señales NOE, de forma que la gran mayoría de las restricciones de distancia derivadas de las estructuras son compatibles con la asignación de los picos de cruce NOE. El primer ciclo de cálculo parte de un conjunto de estructuras generadas al azar y en los siguientes ciclos se utilizan las procedentes del ciclo anterior. Las restricciones angulares se incorporaron en todos los ciclos del cálculo de estructuras. Se calcularon un total de 100 estructuras, considerándose las veinte mejores, es decir, las que tienen un valor menor de la *función blanco* (ver la Sección 6.9.4 de Materiales y Métodos), como las representativas de la estructura de la quimera SPCp41.

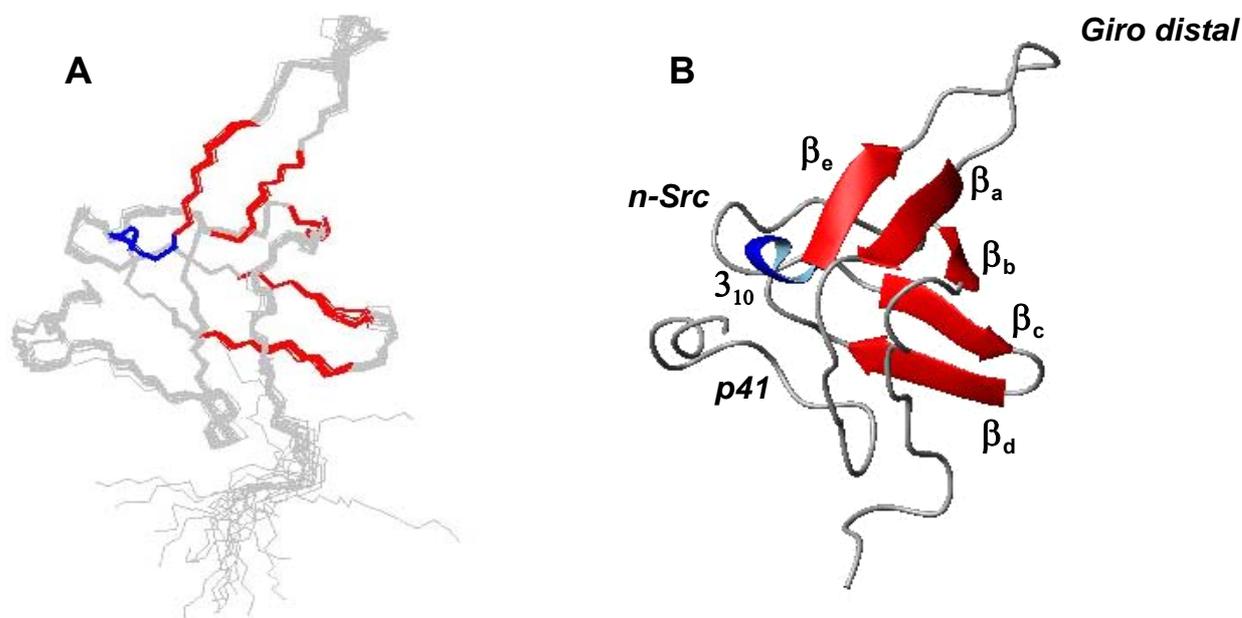


Figura 3.3.1. A. Vista de la superposición de los átomos del esqueleto de las veinte estructuras finales de SPCp41. B. Representación en forma de cinta de una de dichas estructuras, concretamente la de menor energía. Las regiones de hebras β se indican en rojo y la hélice 3_{10} en azul.

El conjunto final de las 20 estructuras calculadas, que satisfacen adecuadamente las restricciones experimentales, está representado en la Figura 3.3.1.A donde además hemos representado un modelo de la estructura tridimensional de SPCp41 en cintas (Figura 3.3.1.B), dibujando los elementos de estructura secundaria con las hebras β representadas en rojo y la hélice 3_{10} en azul.

3.4. Calidad de las estructuras tridimensionales de SPCp41

Para validar las estructuras determinadas se comprobó que las estructuras cumplen los criterios de calidad establecidos para los cálculos realizados con el programa CYANA (Jee y GÅ¼ntert, 2003). La Tabla 3.4.1 recoge la estadística de los

principales parámetros estructurales de las 20 estructuras de menor energía. El análisis indica que no se viola ninguna de las restricciones de distancia o de ángulos diedros. Las desviaciones del conjunto de estructuras se cuantifican mediante la medida del RMSD, es decir, de la raíz de la desviación cuadrática media entre las coordenadas atómicas de todos los pares de estructuras calculadas. En nuestro caso el RMSD por pares de estructuras para los residuos 7 al 77 es de 0.34 ± 0.06 Å para la cadena principal, y de 0.78 ± 0.09 Å teniendo en cuenta todos los átomos pesados. Estos valores se sitúan dentro del intervalo habitual para este tipo de cálculos estructurales, que es $\text{RMSD} \leq \sim 0.8$ Å para la cadena principal y $\text{RMSD} \leq \sim 1.5$ Å para los átomos pesados.

Tabla 3.4.1. Estadística de los principales parámetros estructurales de las 20 estructuras de SPCp41^a.

Restricciones^b	
Distancias	
Corto alcance ($ i-j \leq 1$)	725
Medio alcance ($2 \leq i-j \leq 4$)	144
Largo alcance ($ i-j \geq 5$)	712
Ángulos diedros	
<i>Phi</i>	40
<i>Psi</i>	38
Total	1659
Estadística de las restricciones^c	
Violaciones de distancia > 0.0 Å	0
Violaciones de ángulos diedros $> 0^\circ$	0
Función blanco (Å²)	
Media/Valor mínimo	0.066/0.046
RMSD^d por pares de estructuras de los aminoácidos 7-77 en Å	
Cadena principal N,C α ,C	0.34 ± 0.06
Átomos pesados	0.78 ± 0.09
Representación de Ramachandran^e	
Regiones más favorables(%)	73.5
Regiones permitidas adicionalmente (%)	24.4
Regiones permitidas generosamente(%)	0.5
Regiones no permitidas (%)	1.6

^aLas estadísticas se obtienen del conjunto de las 20 estructuras de menor energía de SPCp41.

^bLas estadísticas de las restricciones hacen referencia a los NOEs asignados.

^cLas violaciones solamente están descritas cuando aparecen en 6 o más estructuras.

^dEl RMSD se calcula como un promedio por pares de estructuras.

^eEstos valores proceden de el análisis con PROCHECK-NMR que ha sido realizado excluyendo los 6 primeros aminoácidos que son los que presentan mayor flexibilidad.

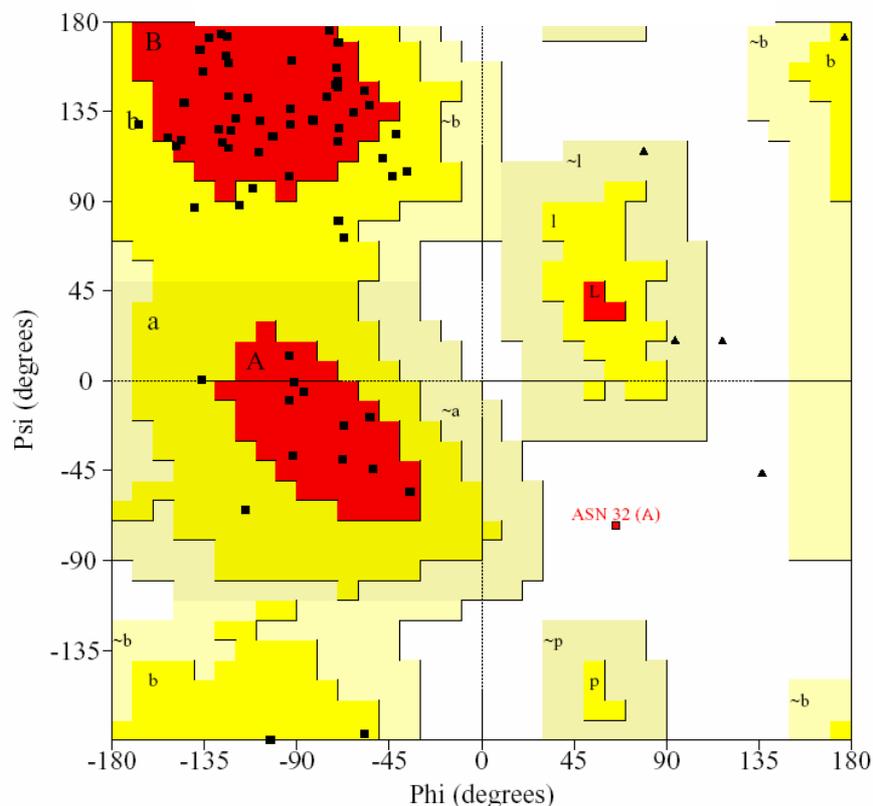


Figura 3.4.1. Representación de Ramachandran para la estructura de SPCp41 de menor energía excluyendo los 6 primeros aminoácidos. Se representan con triángulos negros las glicinas y con un cuadrado negro los aminoácidos no prolina ni glicinas, ni los aminoácidos terminales. En color rojo aparece representado el aminoácido que se encuentra en una región no permitida.

Otra forma de comprobar la calidad de la estructura es la representación de Ramachandran, en la que se representa el ángulo ψ frente al ϕ para cada aminoácido. En la Figura 3.4.1 podemos ver el mapa de Ramachandran para la estructura de SPCp41 de menor energía, en el que se han distinguido distintas zonas. Las zonas A, B y C corresponden a las regiones más favorables para las cadenas polipeptídicas. Las zonas a, b, l y p corresponden a las regiones permitidas adicionalmente, las zonas ~a, ~b, ~l y ~p corresponden a zonas permitidas generosamente, y el resto de las zonas son regiones no permitidas. En la Figura observamos que todos los restos se encuentran en regiones permitidas o permitidas adicionalmente, excepto las glicinas (representadas con triángulos negros) y la Asn32. Tal como se observa en la Tabla 3.4.1 en algunas de las 20 estructuras existen aminoácidos (1.6%) en zonas desfavorecidas. Es destacable sin embargo que el resto Asn32, equivalente a Asn47 en el dominio Spc-SH3 natural, se encuentra en ambos ejemplos localizado en la misma región no permitida de la representación de Ramachandran (Vega *et al.*, 2000).

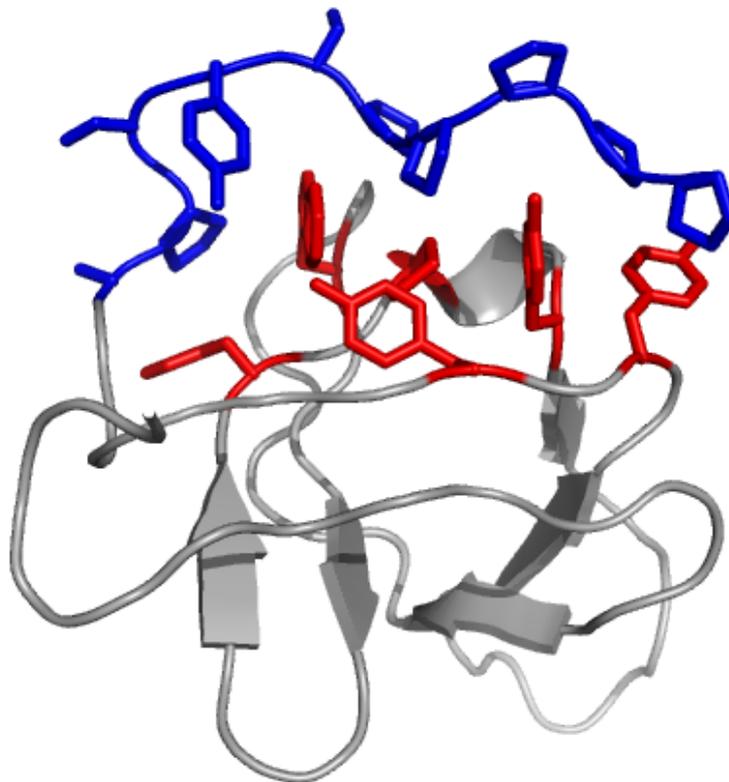


Figura 3.4.2. Estructura tridimensional de la quimera SPCp41 a pH 3.5. En la figura se indican en rojo los aminoácidos implicados en la unión al ligando p41 que está representado en color azul.

La estructura del dominio SH3 en la quimera SPCp41 es análoga a la que se observa en el resto de estructuras determinadas de dominios SH3. Como podemos observar en la Figura 3.4.2, está constituida por cinco hebras β antiparalelas que hemos denominado β_a , β_b , β_c , β_d y β_e que comprenden los aminoácidos 54-57, 15-17, 29-31, 34-35 y 43-45 respectivamente, y una pequeña hélice 3_{10} desde 39-41. Estas hebras β forman dos láminas β antiparalelas y ortogonales entre sí, denominadas β_I y β_{II} . La lámina más pequeña, la β_I , está constituida por las hebras β_3 - β_4 , y la mayor de ellas, la forman las hebras β_e - β_a - β_b . Otro elemento estructural típico de este dominio es un lazo que conecta las dos primeras hebras β_b y β_c y que se conoce como lazo *n-Src*. Entre las hebras β_c y β_d existe un giro β de tipo II' llamado giro *distal*. El largo lazo *RT*, característico de los dominios SH3, se encuentra abierto en la organización topológica de quimera SPCp41 como se describió anteriormente. Finalmente, la región C-terminal de la cadena correspondiente al péptido p41 adopta una conformación en hélice de poliprolina II.

El sitio de unión de los dominios SH3 para ligandos con secuencias ricas en prolina se encuentra fielmente conservado en la estructura de SPCp41 y consiste en una superficie hidrofóbica con tres bolsillos definidos por las cadenas laterales de los aminoácidos aromáticos Trp26, Phe37, Tyr42, Tyr59 y Tyr61 y por Pro39, que están representados en la Figura 3.4.2 en color rojo. El ligando p41, representado en la misma figura en color azul, se dispone en la estructura interaccionando con el sitio de unión y adoptando una conformación en hélice de poliprolina II, típica de los ligandos ricos en prolina en los complejos con dominios SH3.

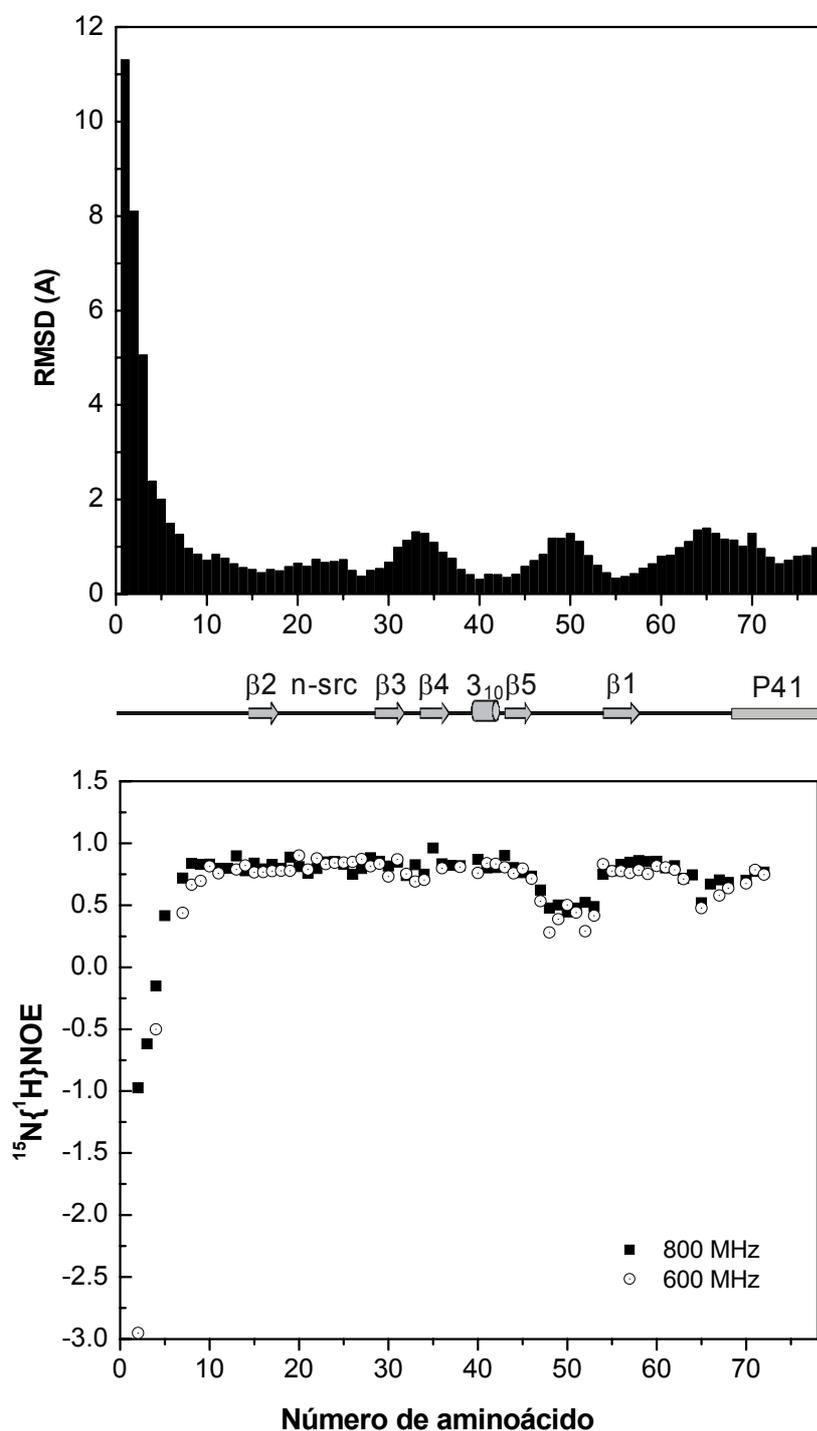


Figura 3.5.1. Comparación del RMSD calculado para el conjunto de las 20 estructuras representadas en la Figura 3.4.1.A con la dinámica de la cadena principal de SPCp41. **A.** Diagrama de barras del RMSD de la cadena principal calculado con MOLMOL. **B.** ^1H - ^{15}N NOE heteronuclear ($^{15}\text{N}\{^1\text{H}\}\text{NOE}$) medido a 600 y 800 MHz. Los errores han sido estimados del orden del 2%. Los diferentes elementos de estructura secundaria se muestran en el centro de los dos paneles con las hebras β numeradas de acuerdo con la numeración que corresponde a la forma natural del dominio SH3 de α -espectrina.

3.5. Dinámica del esqueleto de la quimera SPCp41

Para obtener información preliminar sobre el grado de flexibilidad conformacional de la proteína quimérica a lo largo de su esqueleto, hemos medido los valores del $^{15}\text{N}\{^1\text{H}\}$ NOE a 600 MHz a 25°C y a 800 MHz a 30°C en glicina-D5 20 mM, pH 3.5 (Farrow *et al.*, 1994). En la Figura 3.5.1 hemos comparado para cada residuo los valores del NOE heteronuclear obtenidos para SPCp41 con el RMSD del esqueleto calculado para las 20 estructuras de menor energía obtenidas de SPCp41. Para la mayoría de los residuos del dominio SH3 se observan altos valores del NOE heteronuclear (≈ 0.8) y bajos valores del RMSD lo que indica que estas regiones se encuentran altamente ordenadas en la quimera, con un bajo grado de flexibilidad conformacional. Por otra parte, existe un mayor grado de desorden en algunas regiones de la cadena, muy acusado en el extremo N-terminal, y algo menor en el giro que forman los aminoácidos 48-52 y en la región de la secuencia de conexión (aminoácidos 65-67) que une el dominio SH3 al ligando p41 (aminoácidos 68-77). Este desorden se puede atribuir a cierta flexibilidad conformacional de la proteína en estas regiones, como indican los bajos valores del NOE heteronuclear en las mismas zonas, que reflejan la existencia de movimientos de alta frecuencia (ps a ns) (Jarymowycz y Stone, 2006).

Finalmente, los valores del NOE heteronuclear de los cinco primeros aminoácidos correspondientes al péptido p41 son similares a los de las regiones más rígidas de la proteína, lo que confirma la fuerte unión intramolecular entre el ligando p41 y el dominio SH3 en la estructura terciaria de la quimera SPCp41.

3.6. Comparación de la estructura de SPCp41 con otras estructuras

La Figura 3.6.1 muestra la superposición de la estructura de la quimera SPCp41 con la del permutante circular S19P20s (PDB:1TUC) (Viguera *et al.*, 1996), que se usó como molde para el diseño de la quimera. La región correspondiente al dominio SH3 en la estructura de SPCp41 es muy similar a la estructura de S19P20s con un RMSD de 0.87 Å. En la Figura 3.6.1 también se muestra la superposición de la estructura del mutante R21A de Spc-SH3 (PDB:2JMA), que fue determinada recientemente por RMN (Casares *et al.*, 2007). También podemos observar que las estructuras son muy parecidas, con un RMSD de 0.83 Å. Además, el ángulo ϕ del aminoácido Asn47, que se encuentra en una zona “no permitida” del mapa de Ramachandran en las estructuras determinadas por difracción de rayos X de la forma natural de Spc-SH3, (PDB:1SHG) (Musacchio *et al.*, 1992b) y del permutante circular S19P20s, así como también en la del mutante R21A de Spc-SH3, se halla también en la misma zona para el resto equivalente (Asn32) en todo el conjunto de estructuras obtenidas para la quimera SPCp41. El valor positivo del ángulo ϕ para este aminoácido se confirmó experimentalmente mediante el análisis de un espectro $^1\text{H}, ^{15}\text{N}$ -HSQC. Así, para el ángulo ϕ recientemente se ha propuesto un procedimiento sencillo que permite, en proteínas marcadas en ^{15}N , identificar aquellos aminoácidos que poseen un ángulo ϕ positivo (Crowley *et al.*, 2000). El procedimiento está basado en los efectos de correlación cruzada dipolar originados por los espines $^1\text{H}_\alpha$ y ^{15}N sobre el ^1HN , que quedan reflejados en la forma de múltiplete de la línea de este último. El experimento consiste básicamente en un espectro $^1\text{H}, ^{15}\text{N}$ -HSQC en el que se permite el acoplamiento del protón amida con el nitrógeno, lo que da lugar a una señal del ^1HN en forma de doblete de dobletes, debido al acoplamiento fuerte entre el $^1\text{H}_i$ y $^{15}\text{N}_i$ y al acoplamiento débil entre el $^1\text{HN}_i$ y $^1\text{H}_{\alpha_i}$. La información acerca del diedro ϕ se obtiene de la intensidad relativa de las líneas del múltiplete. Cuando las

intensidades de los picos interiores del multiplete son menores que las de los exteriores, implican un ángulo ϕ positivo. Por el contrario, cuando las intensidades de los picos exteriores son menores, el signo del ángulo diedro ϕ no va a poder ser determinado, si bien ϕ queda restringido a la zona “negativa” de la representación de Ramachandran ($150^\circ < \phi < 360^\circ$).

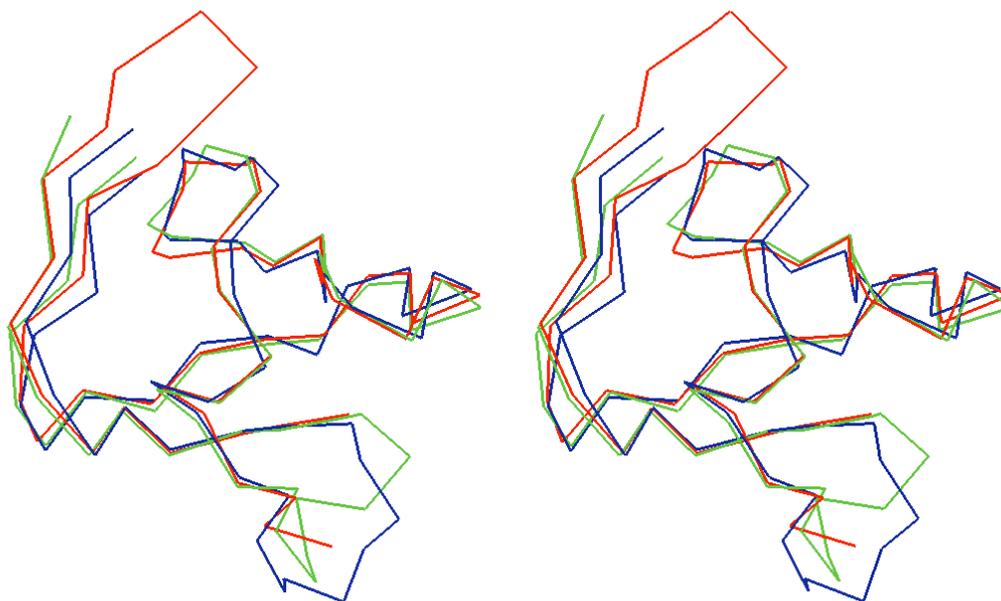


Figura 3.6.1. Superposición en estereo de la estructura de más baja energía obtenida para la quimera SPCp41 en color azul, de la estructura obtenida por cristalografía de rayos X del permutante circular S19P20s en color verde (1TUC) y la estructura obtenida por RMN del complejo formado por el mutante R21A Spc-SH3:p41 (2JMA) en color rojo (solamente se muestra la parte correspondiente al dominio SH3). La superposición se ha hecho con los átomos de los aminoácidos de la cadena principal (N, C_α y C) 8-46 y 53-62 para SPCp41 y los correspondientes aminoácidos, 23-61 y 7-16, de S19P20s y R21A Spc-SH3 respectivamente.

Hemos comparado también la estructura obtenida para SPCp41 con la del complejo formado entre el mutante R21A de Spc-SH3 y el péptido p41, determinada por RMN (Casares *et al.*, 2007). Esta es la estructura más parecida a SPCp41 ya que el péptido en el complejo es el mismo que se encuentra covalentemente unido al dominio SH3 en SPCp41. Además, el dominio SH3 en SPCp41 es una variante del dominio Spc-SH3. En la Figura 3.6.2 se muestra la superficie de unión de SPCp41 y la del complejo R21A Spc-SH3 (Casares *et al.*, 2007). El ligando p41 adopta una conformación muy similar en ambas estructuras como podemos observar en el panel A de la Figura 3.6.2. Las pequeñas diferencias estructurales se observan principalmente en la Tyr4 del péptido p41, probablemente debido a que en la estructura de la quimera SPCp41 el lazo RT está cortado, mientras que en el complejo no, y la Lys18 de éste último, que no está presente en la quimera SPCp41, forma un puente de hidrógeno con la Tyr4 de p41 en el complejo. En el panel B de la Figura 3.6.2 están representadas las cadenas laterales de los aminoácidos que forman parte del sitio de unión de ambas estructuras. Puede apreciarse que los aminoácidos aromáticos y la Pro39 en SPCp41 (Pro54 en R21A Spc-SH3), que se encuentran altamente conservados entre los dominios SH3 y que son cruciales para la unión de motivos poliprolina, están dispuestos de igual forma en ambas estructuras.

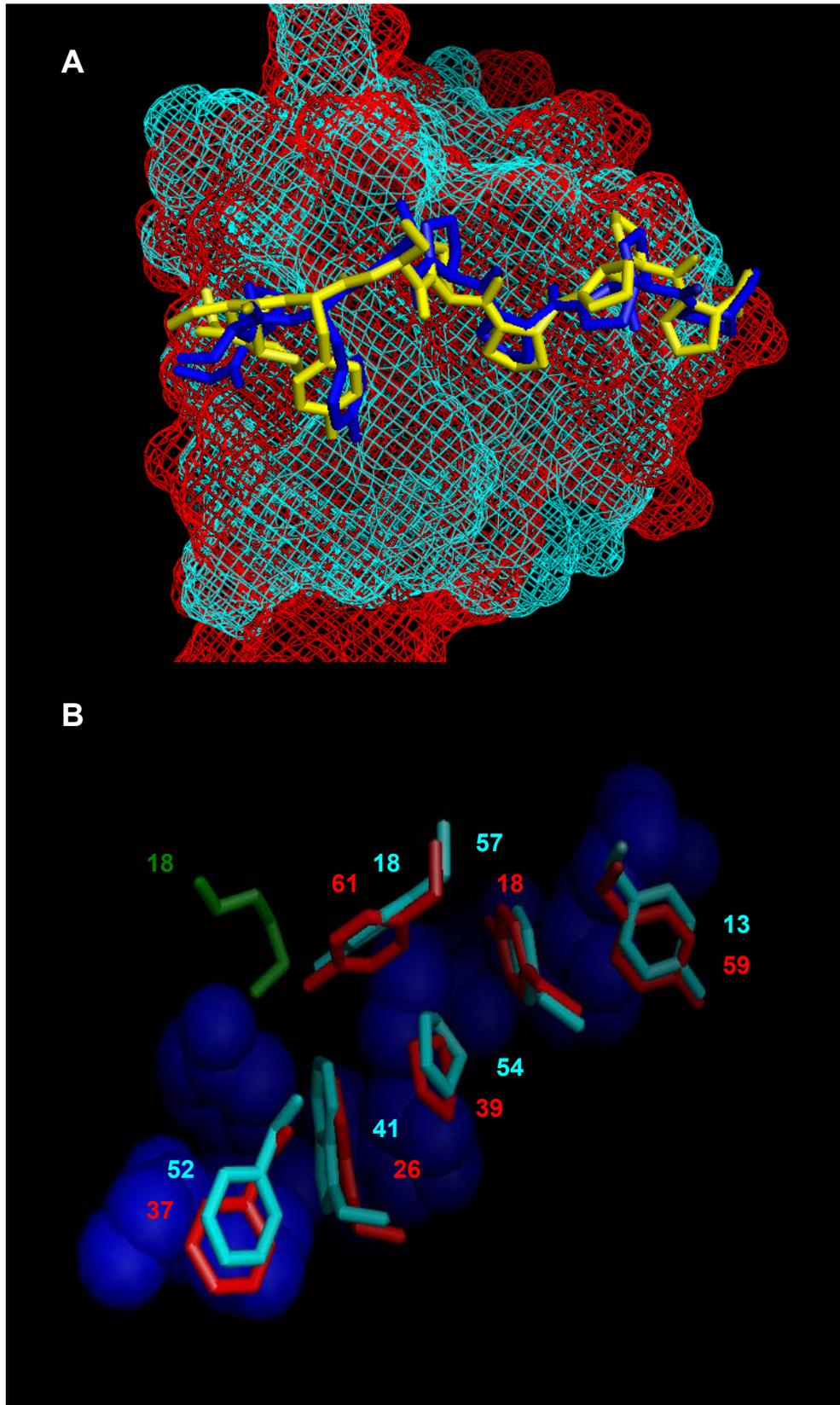


Figura 3.6.2. Comparación entre las interfases de unión de SPCp41 y el complejo R21A Spc-SH3:p41. **A.** Representación de la disposición que adopta el péptido p41 en la quimera SPCp41 (p41 en azul, dominio SH3 en rojo) y en el complejo con el mutante R21A Spc-SH3 (p41 en amarillo, dominio SH3 en turquesa). Las diferencias estructurales se observan principalmente en la Tyr4 (del péptido) debido a la ausencia en SPCp41 de un puente de hidrógeno con la Lys18 que se observa en R21A Spc-SH3 (representada en color verde en el panel B). **B.** Cadenas laterales de los aminoácidos del dominio SH3 que forman el sitio de unión. En color rojo para SPCp41 y en color turquesa para R21A Spc-SH3. El péptido p41 aparece en color azul.

Tabla 3.6.1. Estadística de los contactos entre el dominio SH3 y el segmento p41^a de las 20 estructuras de SPCp41.

Residuos que interaccionan		Enlaces de hidrógeno			Contactos SH3-péptido	Frecuencia	Residuos que interaccionan en R21A Spc-SH3
p41	Quimera	P-P	LL	P-L		%	
Ala1	Phe37	0	0	0	1	35	Phe52 (3, 95%)
Pro2	Trp26	0	0	0	5	100	Trp41 (7, 100%)
Pro2	Phe37	0	0	0	4	100	Phe52 (2, 50%)
Tyr4	Trp26	0	0	0	3	100	Trp41 (1, 95%)
Tyr4	Trp26	0	0	1	0	100	Trp41 (0, 100%)
Pro6	Pro39	0	0	0	1	100	Pro54 (1, 70%)
Pro6	Asp25	0	0	0	2	25	Asp40 (2, 100%)
Pro6	Trp26	0	0	0	2	100	Trp41 (3, 100%)
Pro7	Tyr61	0	0	0	1	95	Tyr15 (1, 80%)
Pro9	Tyr59	0	0	0	3	100	Tyr13 (2, 100%)
Pro9	Tyr42	0	0	0	2	100	Tyr57 (3, 100%)
Pro10	Tyr59	0	0	0	8	100	Tyr13 (2, 65%)
Pro2	Asn23	0	0	0	1	90	-
Ser5	Trp26	0	0	0	1	85	-
Pro7	Trp26	0	0	0	2	100	-
Pro7	Pro39	0	0	0	3	100	-
Tyr4	-	0	1	0	-	-	Lys18 (0, 100%)
Pro7	-	0	0	1	-	-	Tyr57 (0, 95%)

^aContactos analizados con LIGPLOT (Wallace *et al.*, 1995) usando los parámetros que por defecto nos da el programa. Los contactos se describen si están presentes en al menos 10 de las 20 estructuras o en al menos una de las 20 estructuras cuando la interacción entre R21A Spc-SH3:p41 está presente en 10 o más de los 20 modelos del PDB (2JMA) (Casares *et al.*, 2007). Los enlaces de hidrógeno se describen para cadena principal-cadena principal (P-P), para cadena lateral-cadena lateral (L-L) y cadena principal-cadena lateral (P-L). El valor medio de los contactos no-unidos se describe como el % de frecuencia de contacto. El residuo correspondiente a la interacción en R21A Spc-SH3 para un cierto contacto y el número de contactos y frecuencia en las 20 estructuras (entre paréntesis, respectivamente), se representan en la última columna. Las últimas dos filas corresponden a contactos presentes en R21A Spc-SH3:p41, que no están presentes en ninguna de las 20 estructuras de la quimera SPCp41. Para comparar, la numeración de los residuos del p41 en la quimera, se ha hecho del 1-10, que corresponde a los residuos 68-77 en SPCp41.

Hemos usado el programa LIGPLOT (Wallace *et al.*, 1995) para analizar los contactos hidrofóbicos y los puentes de hidrógeno que se encuentran respectivamente en la quimera SPCp41 y en el complejo R21A Spc-SH3:p41; el resultado está recogido en la Tabla 3.6.1. Generalmente encontramos los mismos contactos en los dos complejos proteína-ligando, aunque se observan pequeñas diferencias. Como hemos comentado anteriormente, éstas se deben principalmente a la ausencia en SPCp41 de un puente de hidrógeno entre las cadenas laterales de la Tyr4 de p41 y Lys18 de R21A Spc-SH3 en el lazo RT.

Considerando por separado en la quimera el segmento p41 (aminoácidos 68-77) y la parte del dominio SH3 (aminoácidos 1-67), hemos calculado el área superficial expuesta al disolvente que queda enterrada con la interacción entre ambas partes utilizando en programa NACCESS (Hubbard y Thornton, 1993). La unión intramolecular entierra un área de $648 \pm 8 \text{ \AA}^2$ de superficie no polar y de $227 \pm 28 \text{ \AA}^2$ de superficie polar, ligeramente menor que la que se entierra en el complejo R21A Spc-SH3:p41 ($686 \pm 11 \text{ \AA}^2$ y $241 \pm 11 \text{ \AA}^2$ respectivamente) (Casares *et al.*, 2007)

En definitiva, podemos concluir que la estructura tridimensional en disolución de SPCp41 obtenida mediante RMN demuestra que esta quimera reproduce correctamente en su estructura terciaria las interacciones típicas que se observan en los complejos entre dominios SH3 y péptidos ricos en prolinas. Así, disponemos ahora de una herramienta que nos permitirá analizar y optimizar de forma simple y eficiente las interacciones en el sitio de unión SH3-ligando mediante el estudio de las magnitudes termodinámicas de su desplegamiento combinado con técnicas de mutagénesis dirigida. Esta metodología nos facilitará la comprensión de las interacciones SH3-péptido desde un punto de vista termodinámico, estructural y dinámico, que son esenciales para la racionalización del reconocimiento molecular y el diseño eficaz de ligandos.

CAPÍTULO 4:
ANÁLISIS DE LA TERMODINÁMICA DE UNIÓN DE
DOMINIOS SH3 A PÉPTIDOS RICOS EN PROLINA
USANDO LA QUIMERA SPCp41

CAPÍTULO 4: ANÁLISIS DE LA TERMODINÁMICA DE UNIÓN DE DOMINIOS SH3 A PÉPTIDOS RICOS EN PROLINA USANDO LA QUIMERA SPCp41

En los Capítulos anteriores de esta Memoria demostramos que es posible construir una proteína quimérica que mimetiza perfectamente en su estructura terciaria las interacciones típicas entre los dominios SH3 y sus ligandos peptídicos. A partir de este punto, nuestro objetivo fue demostrar que es posible utilizar estas quimeras para evaluar la termodinámica de esta interacción mediante una combinación de mutagénesis dirigida y experimentos de desplegamiento térmico.

Concretamente analizamos el efecto sobre las magnitudes termodinámicas de la interacción SH3:p41 producido por la sustitución en SPCp41 de cada una de las prolinas de la secuencia del ligando p41 por alanina. Para cada mutante Pro-Ala se estudió su desplegamiento térmico mediante CDB, se ha caracterizó su conformación mediante espectroscopias de CD y fluorescencia y se realizaron estudios estructurales y dinámicos mediante RMN con el mutante P74A. Los resultados nos permitieron evaluar la importancia de cada uno de los residuos prolina del ligando p41 en la afinidad de unión.

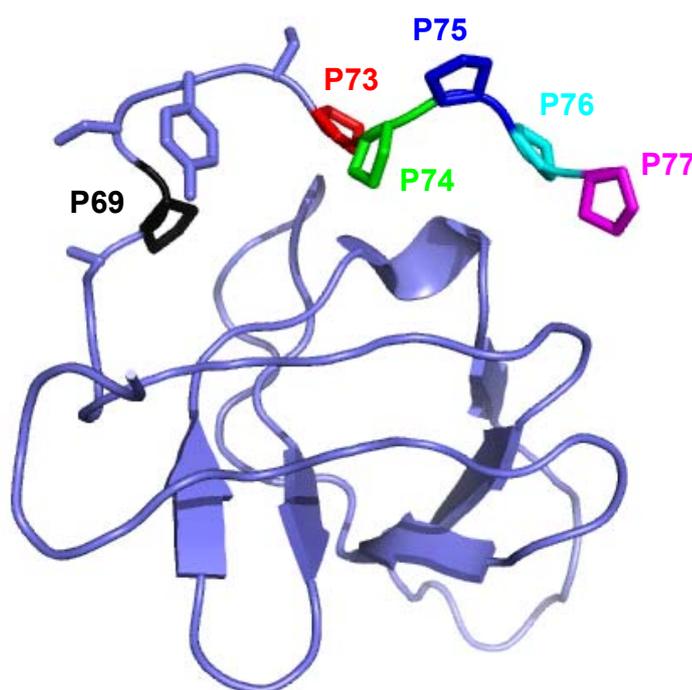


Figura 4.1.1. Representación en cintas de la estructura tridimensional de la quimera SPCp41 determinada por RMN (Candel, A. M. et al., 2007) en la que están representadas las cadenas laterales de los residuos sobre los que se han realizado las mutaciones. El dominio SH3 aparece en azul claro, la secuencia de conexión en amarillo y las distintas prolinas mutadas del péptido siguen un código de colores y están marcadas con el número que ocupan en la secuencia de la quimera.

4.1. Descripción de los mutantes Pro-Ala de la quimera SPCp41

En la Figura 4.1.1 se muestra esquemáticamente la estructura tridimensional de la quimera SPCp41 en la que se indican las posiciones de las prolinas que se sustituyeron por alanina. La obtención de todos estos mutantes de la quimera SPCp41 se llevó a cabo mediante técnicas de mutagénesis dirigida realizadas sobre la quimera SPCp41 original, tal y como se describe en la Sección 6.2 de Materiales y Métodos. La purificación de las quimeras mutantes se realizó de la misma forma que la de SPCp41, como se encuentra recogida en la Sección 6.4 de Materiales y Métodos.

4.2. Caracterización estructural de la quimera SPCp41 y los mutantes Pro-Ala por espectroscopia de fluorescencia y dicroísmo circular

En primer lugar analizamos los efectos que provocan las mutaciones Pro-Ala en la secuencia del péptido p41 sobre la conformación de la quimera utilizando espectroscopias de fluorescencia y dicroísmo circular. Las condiciones experimentales utilizadas fueron pH 2.5, 3.0 y 3.5, en tampón glicina 20 mM y pH 4.0 en acetato sódico 20 mM, idénticas a las de los experimentos de desplegamiento térmico por calorimetría diferencial de barrido, que se describen más adelante en este Capítulo.

Se registraron espectros de emisión de fluorescencia de triptófano a 25°C entre 305 y 400 nm para cada mutante y cada pH, excitando selectivamente a 298 nm y las rendijas de excitación y emisión se dispusieron a 2.5 y 5 nm respectivamente, para evitar problemas de saturación de la señal. Estos espectros se compararon con los realizados en las mismas condiciones con la quimera SPCp41 y el permutante circular S19P20s (Figura 4.2.1). Se observa un aumento en la intensidad de fluorescencia para todas las quimeras en comparación con S19P20s y, al mismo tiempo, se produce un desplazamiento del máximo de fluorescencia hacia longitudes de onda menores en relación al máximo de fluorescencia que presenta el permutante circular. Como se demostró en la Sección 2.3.1 del Capítulo 2, estos efectos son similares a los que se observaron en distintos estudios previos de la unión del péptido p41 al permutante circular S19P20s (Martin-Sierra *et al.*, 2003), al dominio natural Spc-SH3 (Casares *et al.*, 2007) y al dominio Abl-SH3 (Pisabarro y Serrano, 1996) y se deben a la interacción entre la secuencia del péptido p41 que se encuentra en el extremo C-terminal y el sitio de unión del dominio SH3 de las quimeras que altera el entorno químico y la superficie expuesta al disolvente del grupo indol del Trp26 que se encuentra en la superficie de del sitio de unión (Viguera *et al.*, 1994a; Candel *et al.*, 2007; Casares *et al.*, 2007).

Si observamos la Figura 4.2.2 vemos que la intensidad de fluorescencia para algunos mutantes es mayor que para otros, lo que nos indica que existen diferentes grados de interacción entre la secuencia de p41 y el sitio de unión del dominio SH3 entre las diferentes variantes, que sigue el orden aproximado: SPCp41 > P75A ≥ P76A ≈ P73A > P69A ≈ P74A ≈ P77A.

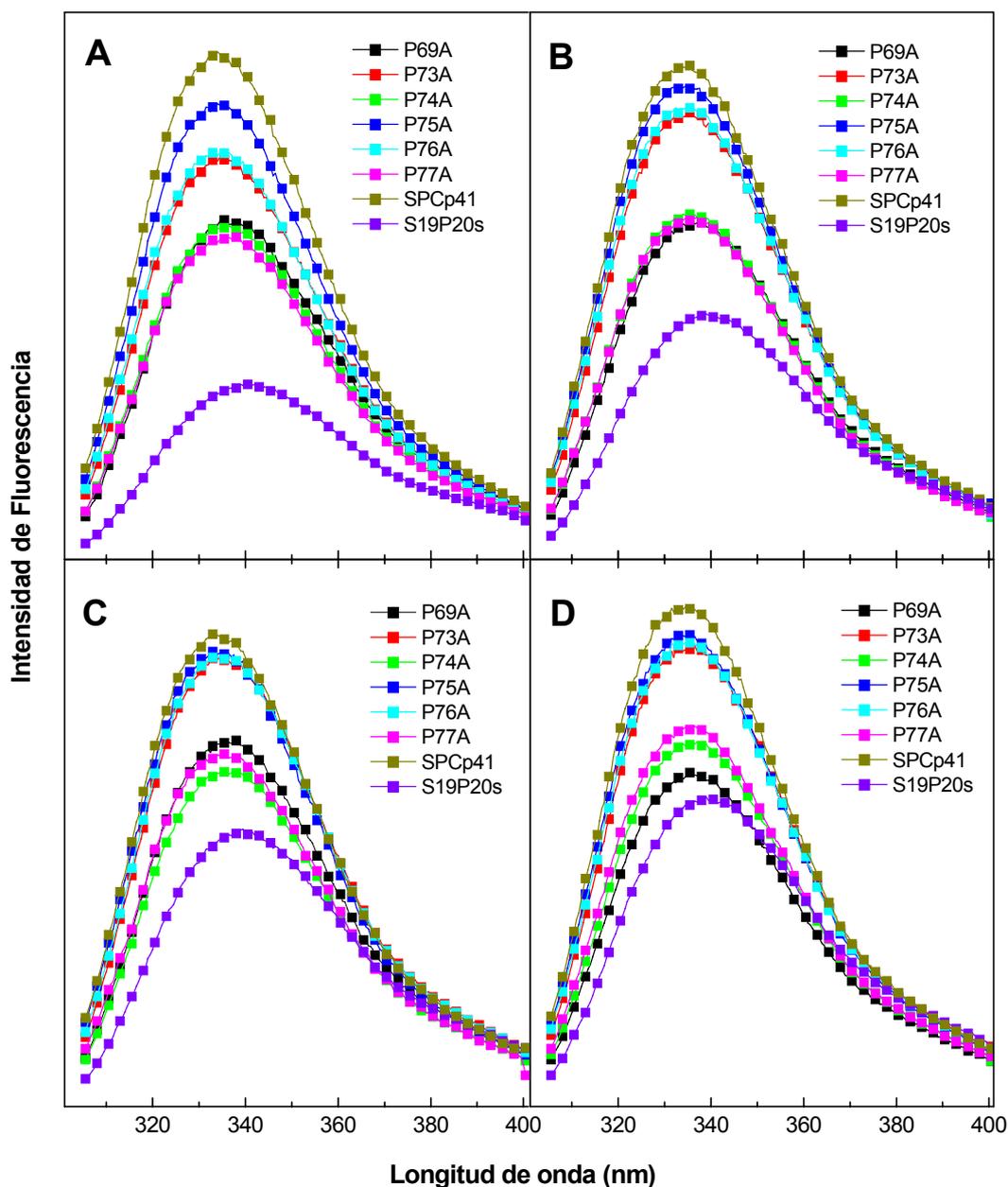


Figura 4.2.1. Espectros de emisión de fluorescencia del permutante circular S19P20s, de la quimera SPCp41 y de los distintos mutantes Pro-Ala, a 25°C y a diferentes pHs: **A.** pH 2.5; **B.** pH 3.0; **C.** pH 3.5; **D.** pH 4.0.

Los espectros de dicroísmo circular (DC) en el UV-lejano y cercano se registraron a 25°C en las mismas condiciones experimentales de los espectros de fluorescencia. En la Figura 4.2.3 se comparan los espectros de DC en el UV-lejano de la quimera SPCp41 y de los distintos mutantes Pro-Ala con el espectro del permutante circular S19P20s. Se observa un cambio considerable en la banda de elipticidad entre 223 y 240 nm que se hace más positiva, con los cambios más acusados a 228 nm. Hemos calculado la diferencia entre los espectros de elipticidad molar de cada quimera y el del permutante circular S19P20s. En la Figura 4.2.4.A hemos representado los espectros de diferencia para las distintas quimeras a uno de los pH, concretamente a pH 3.5. Estos espectros tienen la forma característica de los espectros de DC de los péptidos ricos en prolina que adoptan una conformación en hélice de poliprolina II (Rabanal *et al.*, 1993; Kelly *et al.*, 2001). Esto indica que el

ligando adopta esta conformación en todos los mutantes, al igual que lo hace en la quimera, como ya vimos en el Capítulo 2 y como lo demuestra la estructura de SPCp41 resuelta por RMN (Capítulo 3). No obstante, la alteración en el entorno de las cadenas laterales de los residuos aromáticos también puede contribuir a cambios en la banda positiva de elipticidad que aparece en los espectros de DC a 230 nm (Woody, 1994), fenómeno que ocurre en la unión de ligandos ricos en prolina a los dominios SH3. La magnitud de las diferencias de elipticidad a 228 nm entre cada mutante y el permutante circular se representa en la Figura 4.2.4.B, que muestra una buena correlación con los cambios en la fluorescencia del triptófano para cada mutante Pro-Ala. Estos resultados apoyan la hipótesis de que cada una de las mutaciones Pro-Ala producen un diferente grado de debilitamiento de la interacción SH3:p41 en el sitio de unión de las quimeras dependiendo de su posición, lo que se traduce para algunas quimeras en un mayor desorden en el segmento de cadena del ligando, con pérdida parcial de la conformación en hélice de PPII.

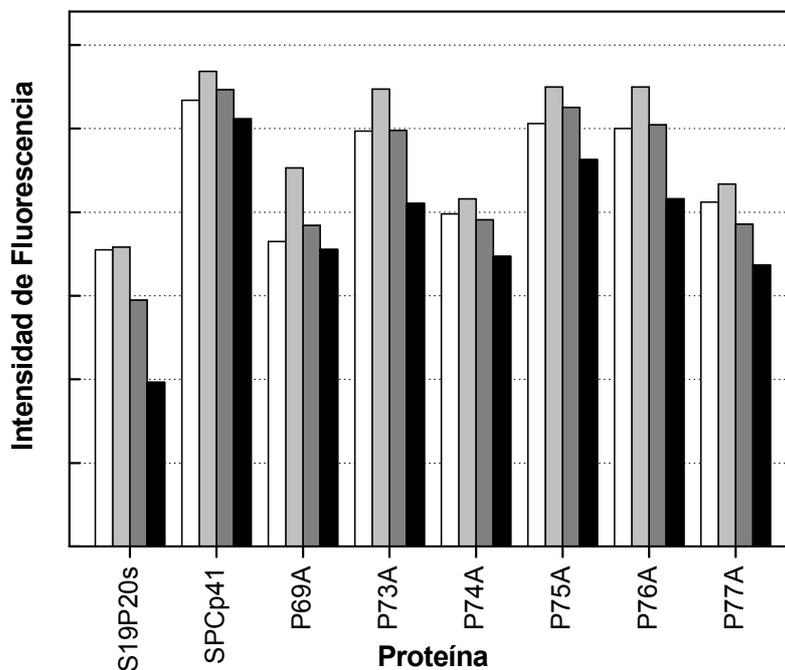


Figura 4.2.2. Intensidad relativa de fluorescencia para cada proteína quimérica y para S19P20s a 25°C y a diferentes pHs: pH 4.0 en blanco; pH 3.5 en gris claro; pH 3.0 en gris; pH 2.5 en negro.

Por otro lado, en los espectros de DC en el UV-cercano (Figura 4.2.5), la quimera SPCp41 y todos sus mutantes muestran un incremento de elipticidad en una banda positiva centrada a 278 nm en comparación con el espectro de S19P20s en las mismas condiciones. Estos cambios pueden atribuirse a la inmovilización de las cadenas laterales de los restos aromáticos de la superficie del sitio de unión por su interacción con el segmento p41. La mutación de cada prolina de p41 produce una variación considerable en la forma y en la intensidad de las bandas de elipticidad, que depende de la posición de la mutación, lo que nos sugiere que hay diferencias en el grado de interacción en la interfase del sitio de unión, que siguen cualitativamente la misma tendencia observada en los espectros de fluorescencia y DC en el UV-lejano. No obstante, es muy difícil dar una interpretación cuantitativa a estos cambios espectrales, dada la diversa naturaleza y magnitud de los efectos que puede provocar cada mutación Pro-Ala en el entorno químico y en el grado de inmovilización de los restos aromáticos del sitio de unión.

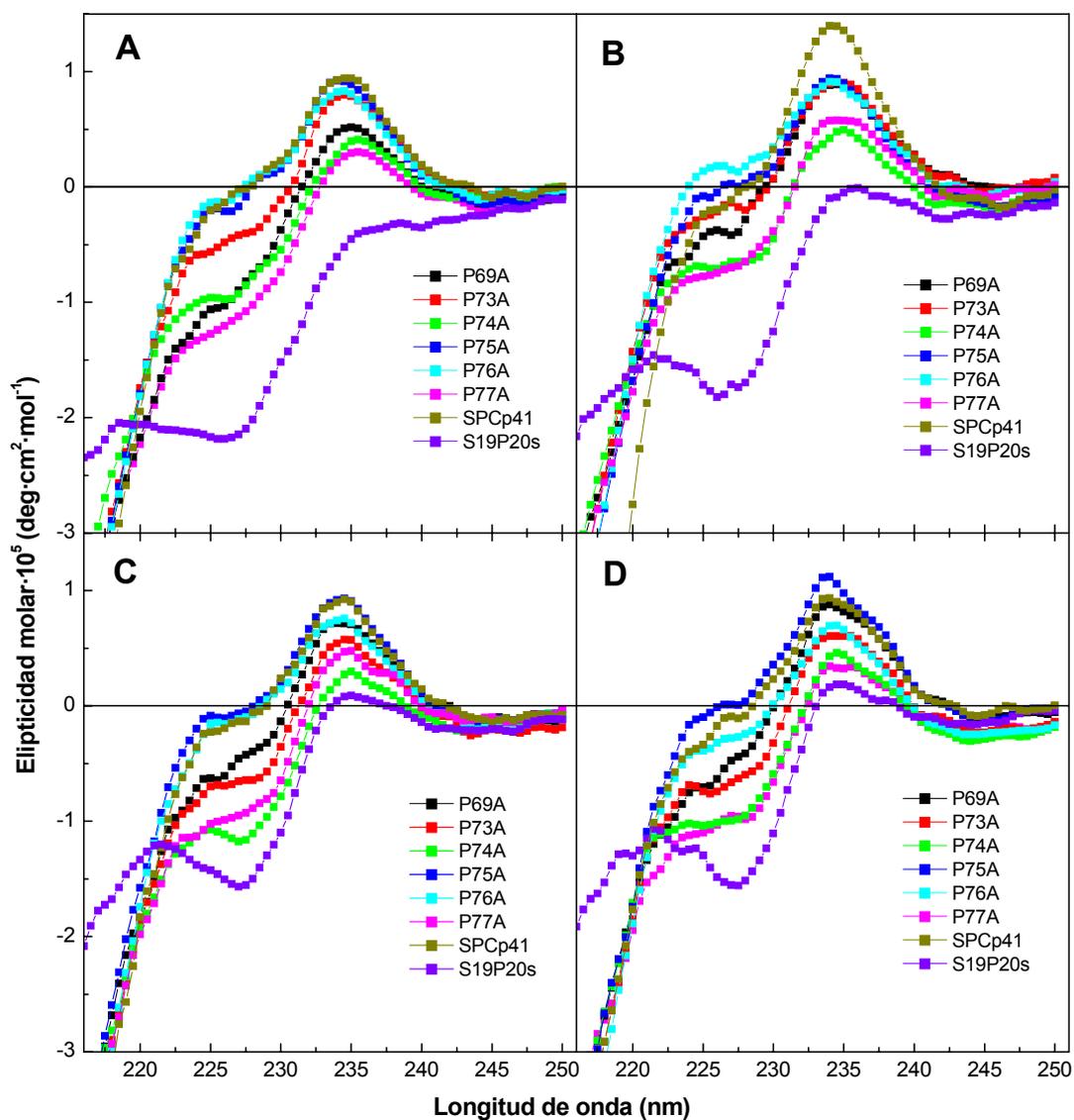


Figura 4.2.3. Espectros de DC de la región del UV-lejano del permutante circular S19P20s, de la quimera SPCp41 y de los distintos mutantes Pro-Ala a 25°C y a diferentes pHs: **A.** pH 2.5; **B.** pH 3.0; **C.** pH 3.5; **D.** pH 4.0.

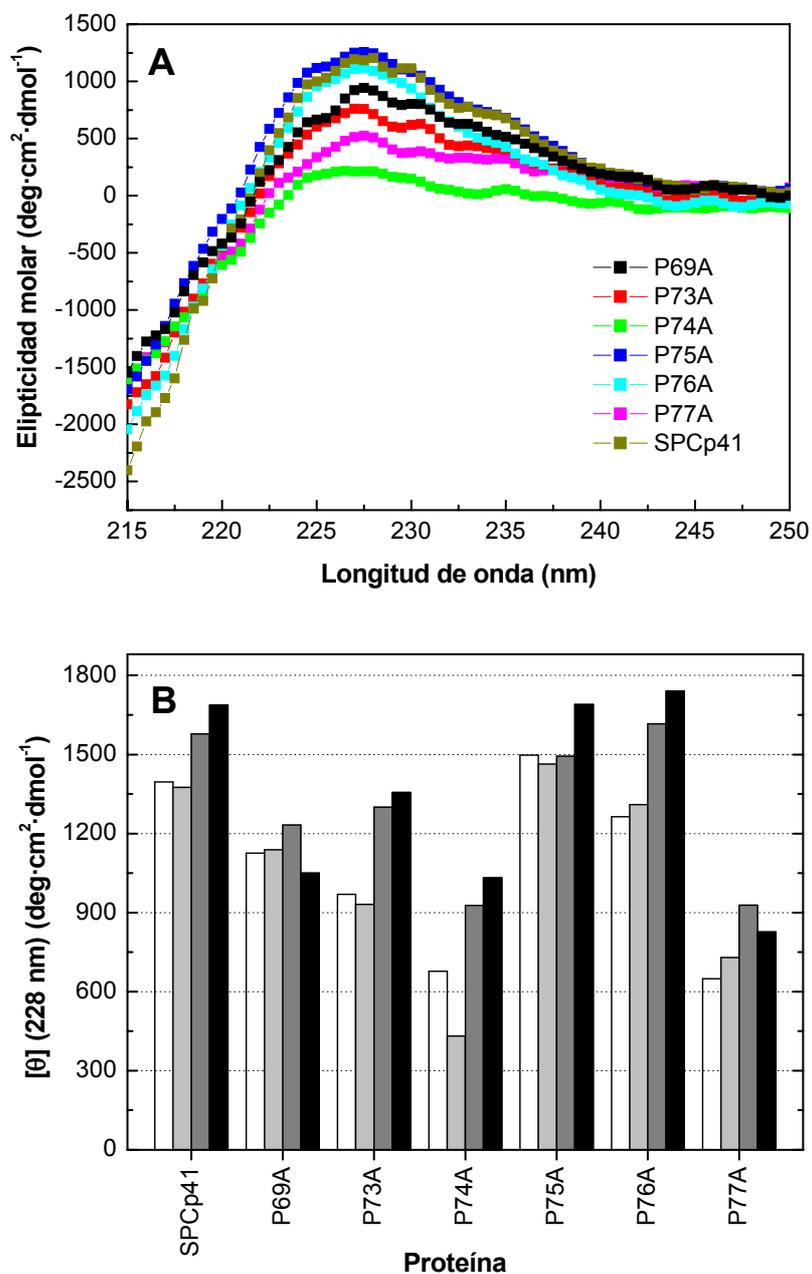


Figura 4.2.4. A. Espectros de diferencia de DC de la región del UV-lejano entre cada proteína quimérica y S19P20s, a 25°C y pH 3.5. **B.** Diferencias de elipticidad molar por residuo, considerando los residuos del péptido más los de la secuencia de conexión, a 228 nm entre cada proteína quimérica y S19P20s, a 25°C y a diferentes pHs: pH 4.0 en blanco; pH 3.5 en gris claro; pH 3.0 en gris; pH 2.5 en negro.

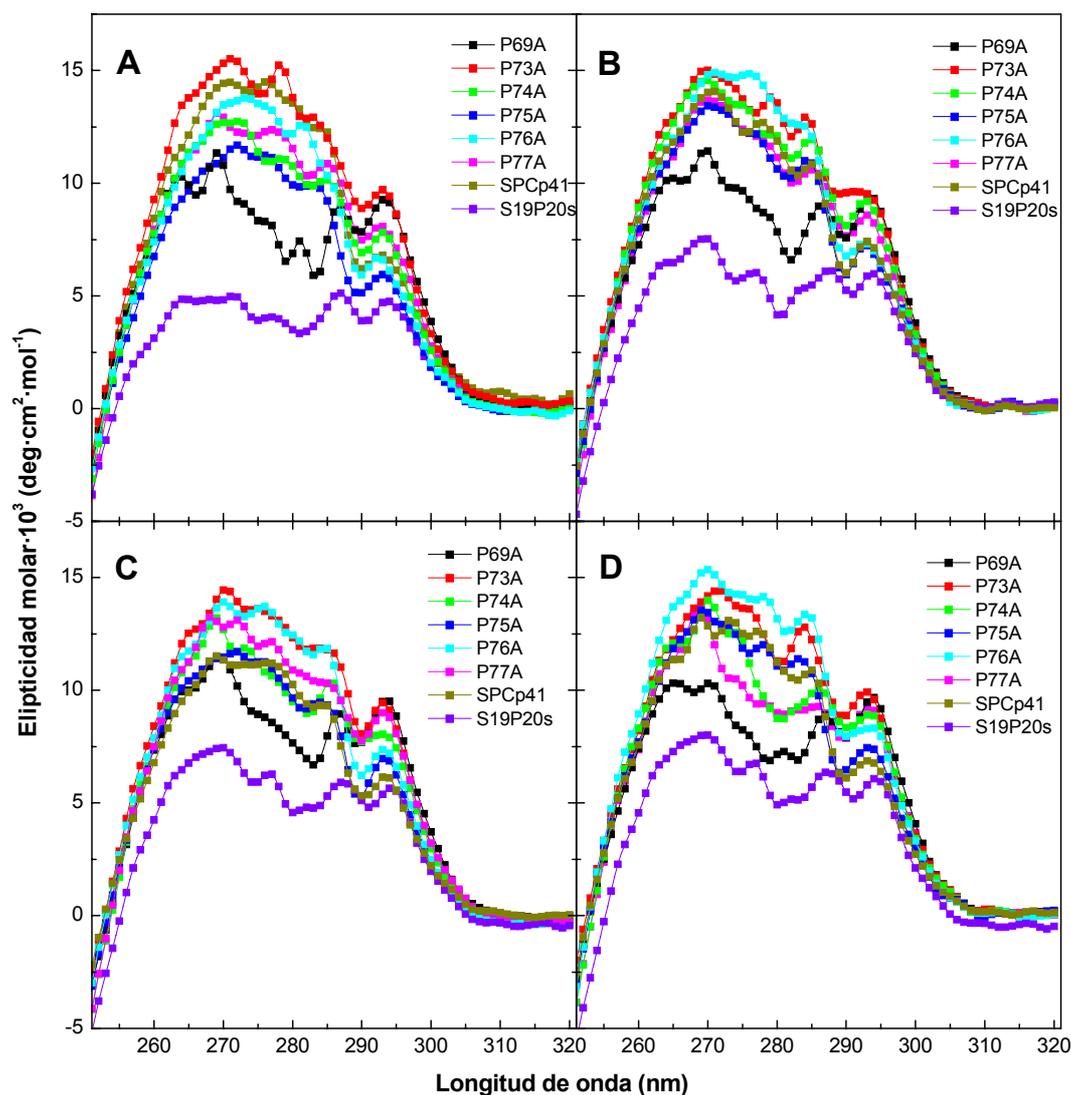


Figura 4.2.5. Espectros de DC de la región del UV-cercano del permutante circular S19P20s, de la quimera SPCp41 y de los distintos mutantes Pro-Ala a 25°C y a diferentes pHs: **A.** pH 2.5; **B.** pH 3.0; **C.** pH 3.5; **D.** pH 4.0.

4.3. Estudio del desplegamiento térmico de la quimera SPCp41 y de sus mutantes Pro-Ala por calorimetría diferencial de barrido (CDB)

Para caracterizar en profundidad la termodinámica del proceso de desplegamiento térmico de la quimera SPCp41 se realizaron experimentos de CDB en tampón glicina 20 mM, a pH 2.0, 2.5, 3.0 y 3.5, en tampón acetato sódico 20 mM, a pH 4.0 y en tampón fosfato sódico 20 mM, a pH 7.0, tal y como se describe en la Sección 6.6 de Materiales y Métodos. Todos los experimentos de CDB se llevaron a cabo a una concentración de proteína en torno a 1-1.5 mg·mL⁻¹. El desplegamiento inducido por temperatura resultó ser muy reversible en todas las condiciones investigadas por lo que podemos asumir que transcurre en equilibrio. En todos los casos, la velocidad de barrido utilizada fue de 1.5 K·min⁻¹ y los termogramas se recogieron en un intervalo de temperaturas comprendido entre 5 y 110°C.

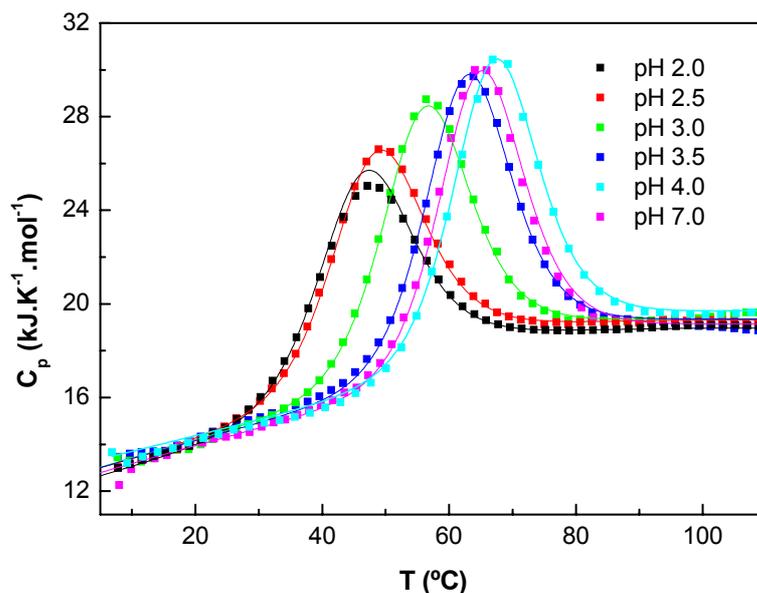


Figura 4.3.1. Dependencia de la capacidad calorífica molar parcial, C_p , con la temperatura de la quimera SPCp41 a distintos valores de pH. Los datos experimentales están representados con símbolos y las líneas representan los resultados del análisis global aplicando el modelo de dos estados. Los termogramas se han realizado a una velocidad de barrido de $1.5 \text{ K}\cdot\text{min}^{-1}$ entre 5 y $110 \text{ }^\circ\text{C}$.

En la Figura 4.3.1 se representan los termogramas experimentales de la dependencia de la capacidad calorífica molar parcial de la quimera SPCp41, C_p , con la temperatura para cada valor de pH. Las líneas continuas corresponden al ajuste global de todos los termogramas obtenidos a diferentes valores de pH utilizando el modelo de dos estados (Sección 6.6.2.2.1 de Materiales y Métodos). En estos ajustes se definió un juego de parámetros único para representar las dependencias con la temperatura de las capacidades caloríficas de los estados nativo y desplegado, $C_{p,N}$ y $C_{p,D}$, así como el cambio de entalpía de desplegamiento, $\Delta H_D(T)$, lo cual es una buena aproximación para proteínas globulares (Privalov, 1979). En todos los casos el modelo de dos estados describe aparentemente bien el desplegamiento térmico de la proteína.

También se realizaron los experimentos de CDB para el permutante circular S19P20s en las mismas condiciones utilizadas para SPCp41, obteniendo valores para los parámetros termodinámicos de desplegamiento similares a los obtenidos en trabajos previos (Martinez *et al.*, 1999). El desplegamiento inducido por la temperatura resultó ser también muy reversible en todas las condiciones investigadas y, por tanto, podemos considerar que transcurre en equilibrio. En la Figura 4.3.2 se representan los termogramas experimentales de la dependencia de la capacidad calorífica molar parcial de permutante circular S19P20s, C_p , con la temperatura para cada valor de pH. Las líneas continuas mostradas en la Figura 4.3.2 corresponden al ajuste individual de cada termograma obtenido a diferentes valores de pH utilizando el modelo de dos estados (Sección 6.6.2.2.1 de Materiales y Métodos). En todos los casos el modelo de dos estados describe bien el desplegamiento térmico de la proteína. En la Tabla 4.3.1 están recogidos los parámetros del desplegamiento del permutante obtenidos a partir del ajuste para cada pH.

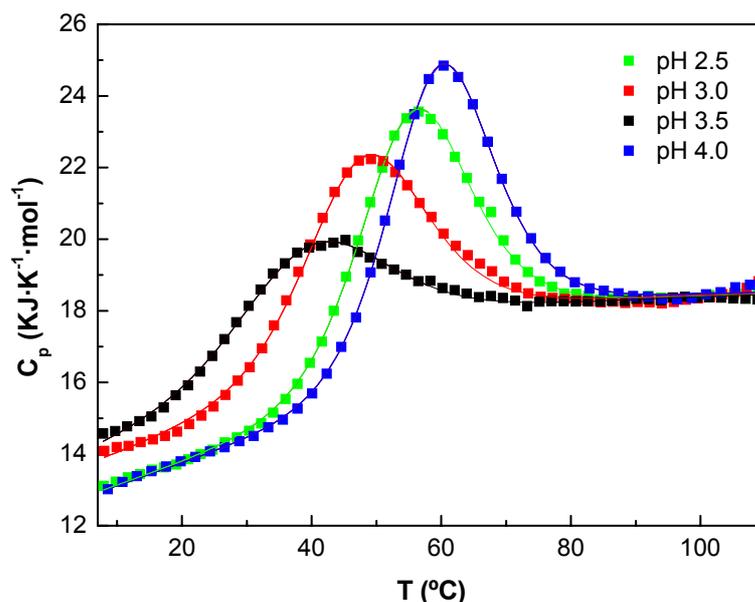


Figura 4.3.2. Dependencias de la capacidad calorífica molar parcial, C_p , con la temperatura, del permutante circular S19P20s a distintos valores de pH. Los datos experimentales están representados con símbolos y las líneas representan los resultados del análisis individual de cada termograma aplicando el modelo de dos estados. Los termogramas se han realizado a una velocidad de barrido de $1.5 \text{ K}\cdot\text{min}^{-1}$ entre 5 y $110 \text{ }^\circ\text{C}$.

Una vez caracterizado el proceso de desplegamiento térmico de la quimera SPCp41 y del permutante S19P20s, se realizaron idénticos experimentos de CDB para los mutantes Pro-Ala de SPCp41 descritos anteriormente (ver Figura 4.6.1). Todos los mutantes se estudiaron en las mismas condiciones experimentales que la quimera SPCp41, de tal manera que la única variable fuese la mutación realizada sobre la quimera en cada caso y así poder comparar los resultados obtenidos.

Todas las curvas de desplegamiento de las distintas formas mutantes presentaron un alto grado de reversibilidad en la mayoría de las condiciones estudiadas, lo que indica que los procesos transcurren en equilibrio. Sin embargo, a pH 7.0 en fosfato sódico los mutantes mostraron tendencia a agregarse y su desplegamiento térmico fue irreversible lo que imposibilitó el análisis comparativo con la quimera SPCp41 a este pH cercano al fisiológico. No obstante, el pH no parece afectar significativamente a la termodinámica de unión de p41 a los dominios de Abl-SH3 y Spc-SH3, según se ha mostrado en estudios previos (Palencia *et al.*, 2004; Casares *et al.*, 2007). Por tanto, las magnitudes termodinámicas medidas a pH ácido se pueden considerar representativas de la unión entre dominios SH3 y sus ligandos.

En una primera aproximación analizamos las curvas de capacidad calorífica molar parcial, C_p , de los mutantes Pro-Ala de SPCp41 utilizando el modelo de desplegamiento de dos estados y comparamos los parámetros termodinámicos obtenidos con los de SPCp41 y S19P20s. Todas las curvas de C_p se describen muy bien mediante el modelo de dos estados, igual que ocurría con SPCp41. En la Figura 4.3.3 se muestra una representación de las correlaciones existentes entre las entalpías de desplegamiento, ΔH_m , y las temperaturas de la transición, T_m , cuyos valores están recogidos en la Tabla 4.3.1, para todos los mutantes comparadas con las de SPCp41 y S19P20s en las mismas condiciones experimentales (Martinez *et al.*, 1999; Martin-Sierra *et al.*, 2003). Si asumimos la validez del modelo de dos estados para el desplegamiento y de acuerdo con la interpretación habitual de este tipo de

representaciones, la pendiente de la dependencia lineal de ΔH_m con T_m correspondería al cambio de capacidad calorífica de desplegamiento, ΔC_p , que se supone aproximadamente constante en el intervalo de temperaturas de los valores de T_m . De manera sorprendente, el mayor valor de esta pendiente se obtiene para S19P20s ($3.57 \pm 0.20 \text{ kJ}\cdot\text{K}^{-1}\cdot\text{mol}^{-1}$), el cual no presenta interacciones en el sitio de unión, ya que es el dominio SH3 aislado. Para todas las quimeras se obtienen valores de pendiente considerablemente menores, lo que indicaría valores más bajos de ΔC_p para sus desplegamientos. Como se ha descrito en el Capítulo 3, la introducción de la interacción SH3-p41 en la quimera SPCp41 produce un enterramiento de superficie hidrofóbica con un área de 875 \AA^2 (Candel *et al.*, 2007), que debería traducirse en un aumento en el ΔC_p de desplegamiento en relación al del dominio S19P20s, ya que existe una correlación directa bien establecida entre el ΔC_p de desplegamiento de las proteínas y el incremento de área superficial expuesta al disolvente con el desplegamiento (Spolar y Record, 1994; Makhatadze y Privalov, 1995). Además, estos resultados no serían coherentes con el valor negativo del cambio en capacidad calorífica que se ha observado para la unión entre el péptido p41 y los dominios Spc-SH3 (Casares *et al.*, 2007) y Abl-SH3 (Palencia *et al.*, 2004). Por estas razones, nos planteamos la hipótesis de que debe de existir alguna contribución adicional a la aparente dependencia con la temperatura de la entalpía de desplegamiento mostrada en la Figura 4.3.3 que no es posible justificar bajo la suposición del modelo de equilibrio de desplegamiento de dos estados.

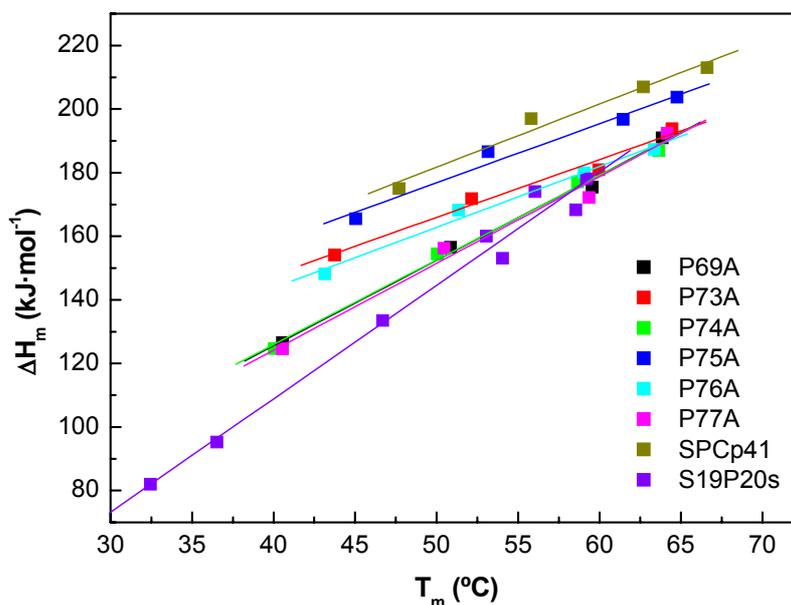


Figura 4.3.3. Dependencia con la temperatura de los cambios de entalpía de desplegamiento de S19P20s, de SPCp41 y de todos los mutantes Pro-Ala estudiados a distintos valores de pH, obtenidos del ajuste global de los datos de CDB mediante el modelo de dos estados. Las líneas continuas corresponden a la regresión lineal de los datos obtenidos para cada proteína.

Tabla 4.3.1. *Parámetros termodinámicos del desplegamiento de la quimera SPCp41, de sus mutantes Pro-Ala y del permutante circular S19P20s obtenidos del análisis de las curvas de CDB mediante el modelo de dos estados.*

Proteína	pH	T _m (°C)	ΔH _m (kJ·mol ⁻¹)
P69A	2.5	40.5 ± 0.1	126.5 ± 0.8
	3.0	50.8 ± 0.1	156.5 ± 0.7
	3.5	59.5 ± 0.1	175.4 ± 0.6
	4.0	63.8 ± 0.1	190.9 ± 0.6
P73A	2.5	43.7 ± 0.1	154.1 ± 0.8
	3.0	52.1 ± 0.1	171.8 ± 0.7
	3.5	59.9 ± 0.1	180.9 ± 0.8
	4.0	64.4 ± 0.1	193.8 ± 0.7
P74A	2.5	40.0 ± 0.2	124.7 ± 1.0
	3.0	50.0 ± 0.1	154.5 ± 0.8
	3.5	58.6 ± 0.1	177.1 ± 0.8
	4.0	63.6 ± 0.1	186.9 ± 0.8
P75A	2.5	45.0 ± 0.1	165.5 ± 0.3
	3.0	53.1 ± 0.1	186.6 ± 0.3
	3.5	61.4 ± 0.1	196.8 ± 0.3
	4.0	64.7 ± 0.1	203.7 ± 0.3
P76A	2.5	43.1 ± 0.1	148.1 ± 0.5
	3.0	51.3 ± 0.1	168.2 ± 0.5
	3.5	59.0 ± 0.1	180.0 ± 0.5
	4.0	63.3 ± 0.1	187.3 ± 0.5
P77A	2.5	40.5 ± 0.1	124.6 ± 0.7
	3.0	50.4 ± 0.1	156.2 ± 0.5
	3.5	59.3 ± 0.1	172.2 ± 0.5
	4.0	64.1 ± 0.1	192.4 ± 0.5
SPCp41	2.5	47.7 ± 0.1	175.0 ± 1.0
	3.0	55.8 ± 0.1	197.0 ± 1.0
	3.5	62.7 ± 0.1	207.0 ± 1.0
	4.0	66.6 ± 0.1	213.0 ± 0.4
S19P20s*	2.5	36.5 ± 0.1	95.2 ± 0.3
	3.0	46.7 ± 0.1	133.5 ± 0.3
	3.5	54.0 ± 0.1	153.1 ± 0.2
	4.0	58.5 ± 0.1	168.3 ± 0.3

*En la Figura 4.3.3 de ΔH_m frente a T_m además de estos valores hemos representado los valores obtenidos previamente para el permutante S19P20s (Martinez *et al.*, 1999).

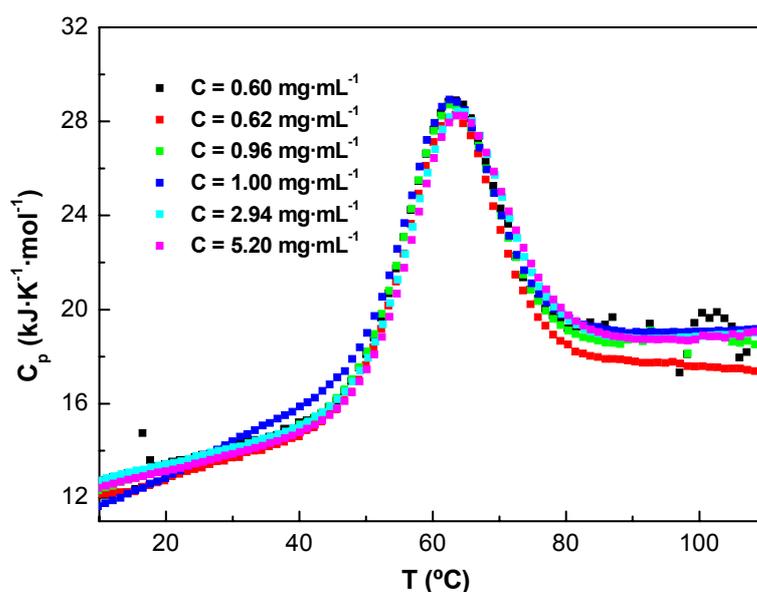


Figura 4.3.4. Dependencias de la capacidad calorífica molar parcial, C_p , con la temperatura, de la quimera SPCp41 a distintas concentraciones de muestra en tampón glicina 20 mM, pH 3.5. Los datos experimentales están representados con símbolos. Los termogramas se han realizado a una velocidad de barrido de $1.5 \text{ K}\cdot\text{min}^{-1}$ entre 5 y $110 \text{ }^\circ\text{C}$.

El que una proteína no se despliegue siguiendo un modelo de dos estados suele deberse bien a que se estén produciendo fenómenos de asociación o bien a que en el proceso de desplegamiento existan estados intermedios, diferentes del nativo y el desplegado, significativamente poblados. Para descartar la posibilidad de que se estuvieran produciendo fenómenos de asociación que pudieran haber pasado inadvertidos, estudiamos el efecto de la concentración de proteína sobre el desplegamiento. Así, realizamos experimentos de CDB con SPCp41 a distintas concentraciones de muestra, cuyos termogramas están representados en la Figura 4.3.4. Como podemos observar, no se aprecia ningún efecto importante sobre los termogramas al aumentar la concentración de la muestra, lo que demuestra que no se producen fenómenos de asociación significativos durante el desplegamiento en nuestras condiciones experimentales.

La otra posibilidad que nos habíamos planteado consistía en la existencia de algún estado intermedio de desplegamiento significativamente poblado. Como ya se ha descrito anteriormente, los experimentos de DC y fluorescencia indicaron un diverso grado de interacción entre el dominio SH3 y la secuencia p41 para SPCp41 y los diferentes mutantes Pro-Ala. Estos resultados sugirieron la existencia de un posible equilibrio conformacional que afectaría a la interfase del sitio de unión en las quimeras y que se encontraría modulado considerablemente por la posición de la mutación Pro-Ala en la secuencia p41. Así, era posible que este equilibrio conformacional tuviera una influencia importante en la termodinámica de desplegamiento de las proteínas quiméricas. Para apoyar estas hipótesis realizamos un análisis comparativo mediante métodos de RMN de la estructura y la dinámica conformacional entre SPCp41 y uno de los mutantes, concretamente en el mutante P74A, como describimos a continuación.

4.4. Análisis por RMN de la quimera SPCp41 y el mutante P74A

Para obtener más detalles estructurales sobre los cambios que se producen en la superficie de unión debidos a las mutaciones Pro-Ala en la secuencia del péptido p41, realizamos experimentos por RMN con muestras enriquecidas en ^{15}N de la quimera SPCp41 y del mutante P74A. Se eligió este mutante por que, según los resultados anteriores, era uno de los que presentaba cambios más importantes en la interacción SH3:p41, lo que se explica por que la prolina mutada se aloja en uno de los bolsillos hidrofóbicos del sitio de unión. Los espectros registrados para ambas proteínas fueron: ^{15}N -HSQC, 3D ^{15}N -NOESY-HSQC y 3D ^{15}N -TOCSY-HSQC. La asignación de los picos de cruce ^1H - ^{15}N del espectro ^{15}N -HSQC del mutante P74A se realizó tomando tiras del espectro 3D ^{15}N -NOESY-HSQC centradas en los distintos pares de frecuencias ^1H - ^{15}N y comparándolas con las tiras correspondientes a la quimera SPCp41, previamente asignadas (Candel *et al.*, 2007) (ver Capítulo 3). En la Figura 4.4.1 se muestra la superposición de los espectros ^{15}N -HSQC asignados de la quimera y del mutante P74A, donde podemos observar los cambios que se producen en los desplazamientos químicos ^1H - ^{15}N de la cadena principal de SPCp41 debidos a la mutación. Los cambios más importantes afectan a los residuos que se encuentran dentro de o cercanos a la superficie de unión SH3:p41, aunque también se afectan unos pocos restos del dominio SH3 más alejados. Estos cambios podrían ser debidos a que el grado de interacción entre el p41 y el sitio de unión del dominio SH3 en el mutante P74A es menor como habíamos postulado. Por el contrario, gran parte de los desplazamientos químicos del resto del dominio SH3 se ven poco afectados, lo que indica que la estructura del dominio SH3 no se altera de manera importante por la mutación.

A continuación comparamos las señales NOE que presentan los dos protones indólicos H ϵ 1 de los dos restos de triptófano de SPCp41 con las que se observan para el mutante P74A. Estos dos restos Trp están muy conservados entre los diferentes dominios SH3. El Trp26 interacciona directamente con el péptido p41, no sólo en SPCp41 sino también en otros complejos que forman distintos dominios SH3 con el péptido p41 libre (Candel *et al.*, 2007; Casares *et al.*, 2007). En cambio, el Trp27 contribuye de manera importante a la estabilidad de los dominios SH3 y presenta muchos contactos con restos del núcleo de la estructura del dominio SH3. En los espectros 3D ^{15}N -NOESY-HSQC de ambas quimeras encontramos numerosos NOEs entre los protones H ϵ 1 de ambos Trp y los protones de otros restos, como podemos ver en la Figura 4.4.2. En ambos espectros se observan todos los NOEs que afectan al H ϵ 1 de Trp27, excepto uno de baja intensidad que no se aprecia en el espectro del mutante P74A. Esto indica que la mutación P74A no produce cambios significativos en la estructura del dominio SH3. Por el contrario, la mayoría de los NOEs que presenta el H ϵ 1 de Trp26 en SPCp41 no se observan en el mutante P74A, lo que indica que la distancia media entre el protón H ϵ 1 de Trp26 y el p41 aumenta considerablemente con la mutación, ya que la intensidad del NOE es inversamente proporcional a la sexta potencia de la distancia protón-protón. Estos resultados ponen claramente de manifiesto que las interacciones en la superficie de unión se debilitan considerablemente al realizar la mutación P74A, mientras que la estructura del dominio SH3 permanece inalterada en gran parte, lo que está de acuerdo con las similitudes encontradas entre las estructuras del dominio Spc-SH3 libre y cuando forma el complejo con p41 (Casares *et al.*, 2007).

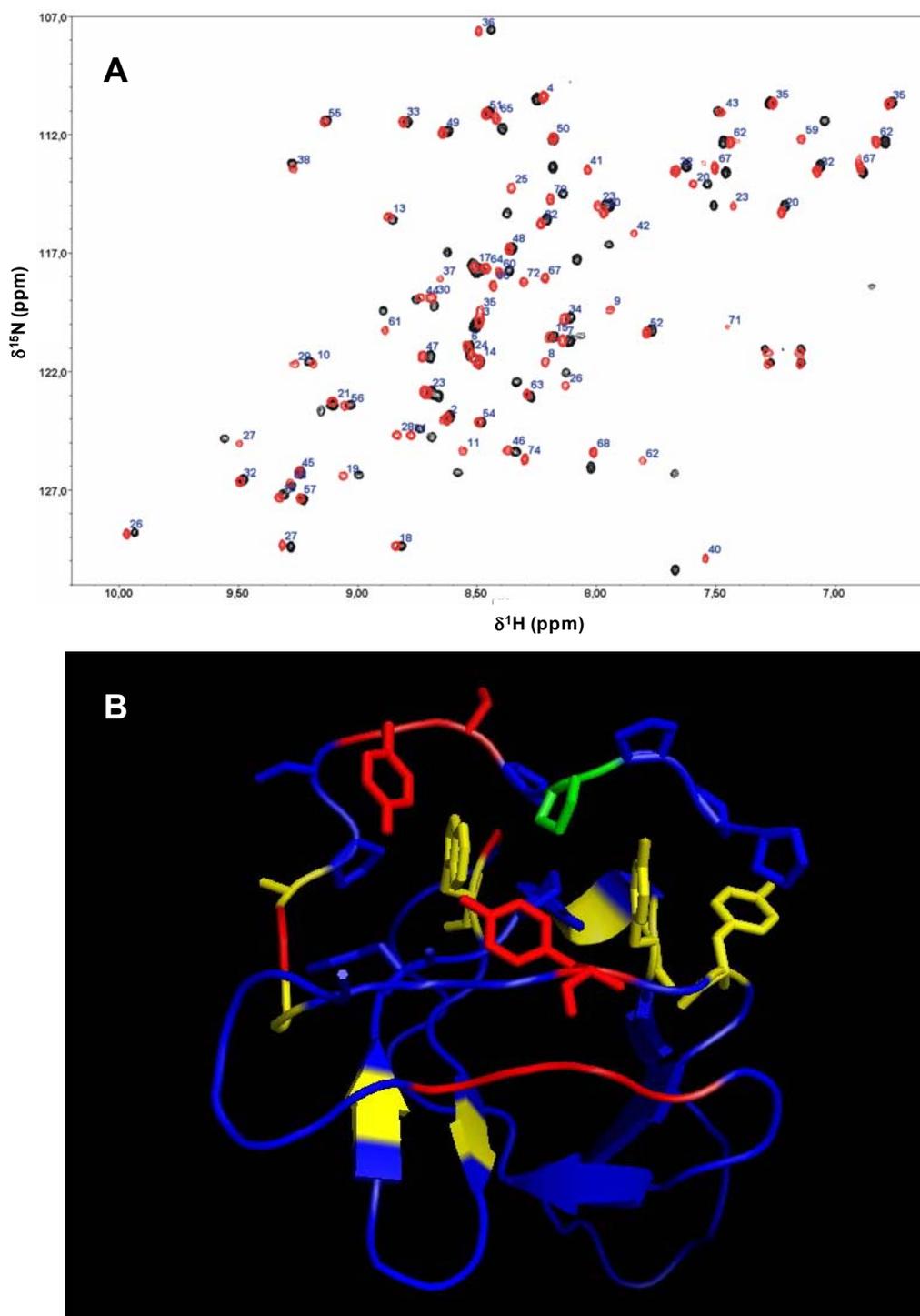


Figura 4.4.1. Cambios en los desplazamientos químicos de la quimera SPCp41 producidos por la mutación P74A. **A.** Superposición de los espectros asignados ^{15}N -HSQC de SPCp41 (negro) y P74A (rojo), registrados a 600 MHz. Cada pico de cruce ^1H - ^{15}N de la cadena principal se ha etiquetado con el número de residuo al que corresponde. **B.** Estructura de menor energía de la quimera SPCp41 (PDB:2JMC) coloreada según la magnitud de los cambios en los desplazamientos químicos ($\Delta\delta$) producidos por la mutación y calculados con la ecuación:

$$\Delta\delta_{ppm} = \sqrt{(\Delta\delta_{HN})^2 + \left(\frac{\Delta\delta_N}{6.51}\right)^2}$$

El código de colores es: $\Delta\delta \leq 0.25$ ppm azul, $0.25 \text{ ppm} \leq \Delta\delta \leq 0.75$ ppm amarillo y $\Delta\delta \geq 0.75$ ppm. Las cadenas laterales de los residuos del péptido p41 y las cadenas laterales de los residuos aromáticos del sitio de unión del dominio SH3 se han representado en varillas. La prolina mutada (P74) está representada en color verde.

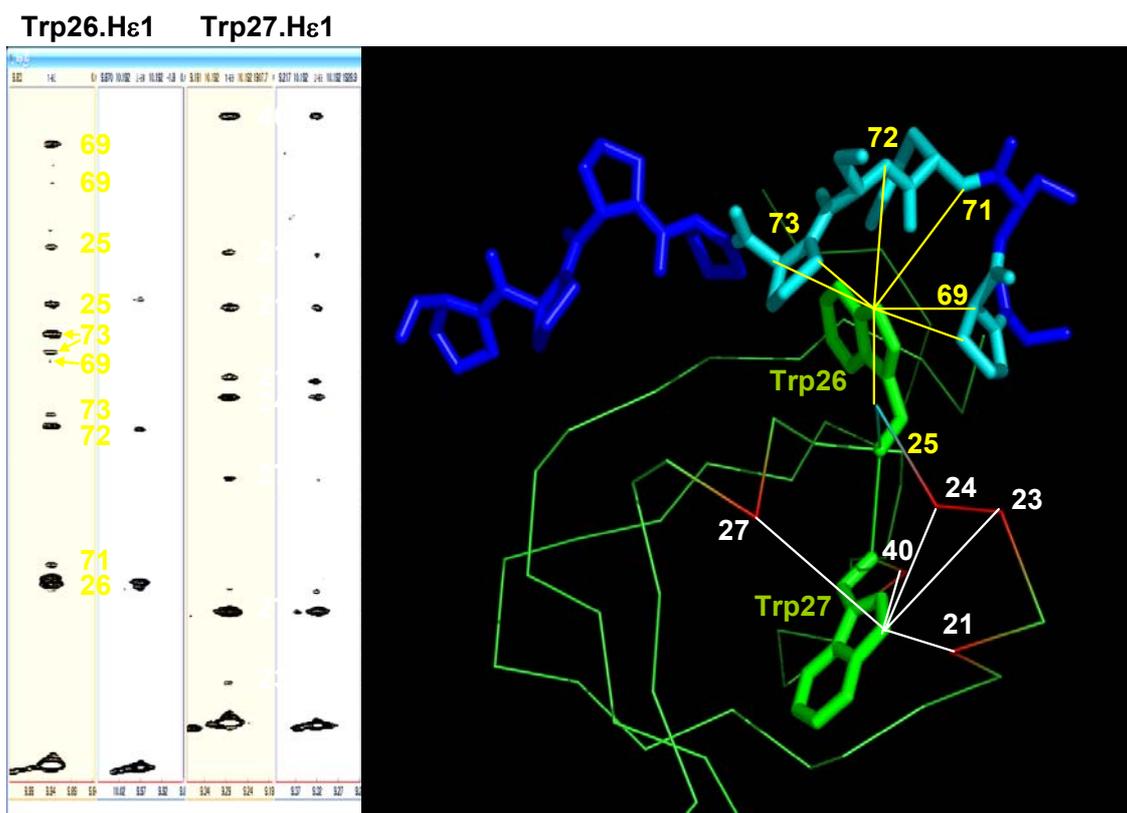


Figura 4.4.2. Diferencias entre los patrones de NOEs que presentan los protones indol $H\epsilon 1$ de Trp 26 y de Trp27 en SPCp41 y P74A. El panel de la izquierda muestra las tiras de los espectros 3D ^{15}N -HSQC-NOESY tomadas a las frecuencias $^1H\epsilon 1$ y $^{15}N\epsilon 1$ de Trp26 y Trp27, para la quimera SPCp41 (fondo naranja) y el mutante P74A (fondo azul). Hemos marcado cada NOE con el número que le ha sido asignado. El panel de la derecha muestra la estructura de la quimera SPCp41 con los mismos NOEs representados con líneas. Los NOEs de Trp26- $H\epsilon 1$ se representan en amarillo y los de Trp27- $H\epsilon 1$ en blanco. El segmento p41 y las cadenas laterales de los dos Trp se representan con varillas, mientras que el resto de la quimera se representa con líneas.

4.5. Análisis de la dinámica de la cadena principal de las proteínas quiméricas mediante experimentos de relajación de RMN ^{15}N

Las medidas de la relajación de RMN del ^{15}N en NH amídicos son una sonda experimental muy sensible para el estudio de la difusión rotacional molecular y de la dinámica global del esqueleto de las proteínas. Así, para analizar el efecto de la mutación P74A en la flexibilidad conformacional de la proteína quimérica, medimos los parámetros de relajación de RMN ($^{15}N R_1$, R_2 y $^{15}N \{^1H\}NOE$) a 600 MHz y a 25°C para SPCp41 y P74A. Los valores del NOE heteronuclear en la región del dominio SH3 (residuos 2-64) y en la secuencia de conexión entre el dominio SH3 y el péptido p41 (65-67) son casi idénticos para ambas proteínas (Figura 4.5.1.A). Para la mayor parte de la cadena del dominio SH3 se observan valores elevados del NOE (≈ 0.8). Valores del NOE menores, que indican la existencia de movimientos de alta frecuencia (ps a ns), se observan en ambas proteínas para los restos 2 a 7 del extremo N-terminal, para los restos 46 a 53 que corresponden a la secuencia de conexión entre los extremos originales N- y C-terminal en el permutante S19P20s y para los restos 65 a 67 de la secuencia de conexión entre S19P20s y el ligando. Por el contrario, los valores del NOE en la región el ligando (restos 68, 70, 71 y 72 en SPCp41 más el resto 74 en el mutante) son significativamente menores en el mutante P74A que en SPCp41, lo que refleja un mayor desorden en esta región de la cadena provocado por

un incremento de los movimientos de alta frecuencia producidos por la mutación P74A (Jarymowycz y Stone, 2006).

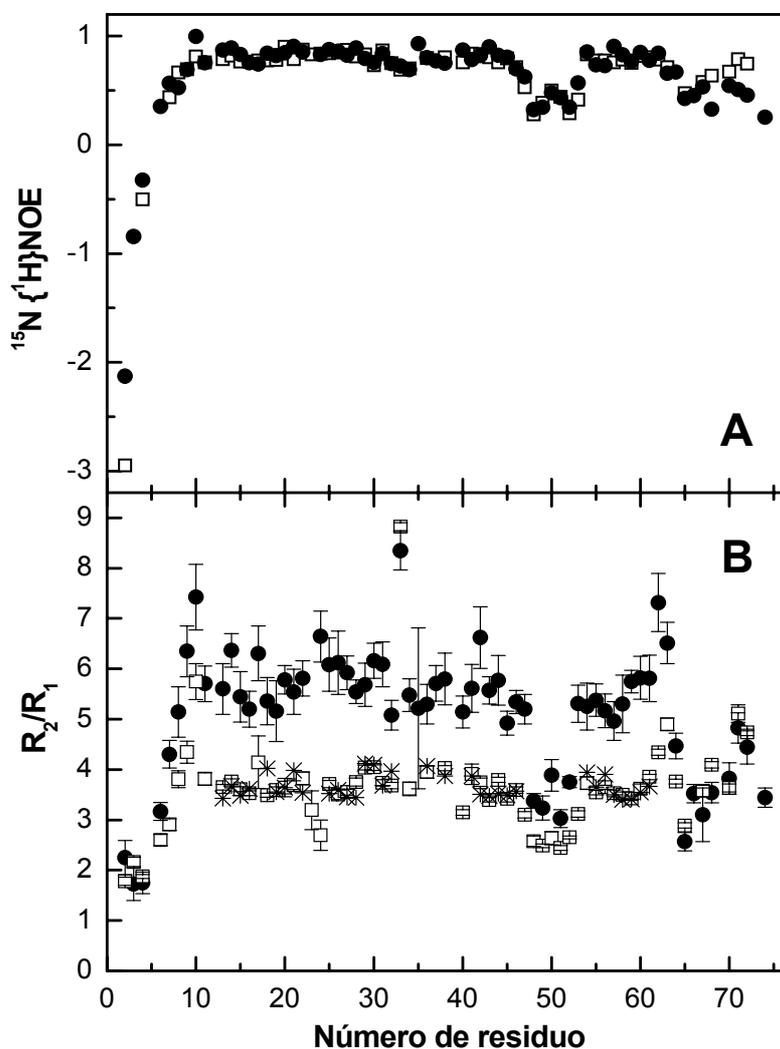


Figura 4.5.1. Dinámica del esqueleto de la quimera SPCp41 y del mutante P74A a pH 3.5 y 25°C, medida a 600 MHz. **A.** $^{15}\text{N}\{^1\text{H}\}\text{NOE}$ heteronuclear de SPCp41 (cuadros blancos) y de P74A (círculos negros). Los errores se han estimado del orden de un 2% para la quimera y de un 4% para el mutante. **B.** Relación entre las constantes de relajación R_2/R_1 para la quimera SPCp41 (cuadros blancos) y el mutante P74A (círculos negros). Los errores los hemos estimado como se describe en Materiales y Métodos. Los valores de la relación R_2/R_1 que se muestran con asteriscos se han calculado para los 33 restos de mayor orden en la quimera utilizando el programa TENSOR2.

También determinamos las constantes velocidad de relajación de ^{15}N , R_1 y R_2 . Desafortunadamente, durante las medidas con el mutante P74A se observó una agregación lenta de la proteína, que se manifiesta en un aumento general de los valores resultantes de R_2 en comparación con los obtenidos para SPCp41. Esto complicó nuestro análisis comparativo y requirió gran precaución a la hora de interpretar los datos de relajación (Korzhnev *et al.*, 2001). No obstante, las diferencias locales en la dinámica entre SPCp41 y P74A pudieron obtenerse analizando la relación R_2/R_1 (Figura 4.5.1.B). El análisis con el programa TENSOR2 (Dosset *et al.*, 2000) de los datos de relajación de los 33 restos correspondientes a la región más ordenada de SPCp41 nos indicó una relación R_2/R_1 promedio que corresponde a un tiempo de correlación rotacional aparente (τ_m) de 5.15 ± 0.04 ns. Este valor se

encuentra dentro del intervalo esperado para una proteína monomérica de unos 8 kDa. Un valor muy parecido, de 5.1 ns, se obtuvo también con el programa Fast-HYDRONMR (Ortega y Garcia de la Torre, 2005) a partir de cálculos hidrodinámicos utilizando la estructura de menor energía de SPCp41. En la Figura 4.5.1.B se observa una disminución de la relación R_2/R_1 para SPCp41 en las zonas en las que también aparecen valores de NOE menores, lo que confirma la presencia en estas regiones de movimientos locales rápidos, con tiempos entre ps y ns. Por el contrario, un aumento en la relación R_2/R_1 indica movimientos más lentos y/o un intercambio conformacional en la escala tiempos de μs -ms, que produce mayores valores R_2 . Esto resulta más evidente para los restos 68, 71 y 72 de SPCp41 que se encuentran en la secuencia del ligando p41 y en los restos 9, 10, 62 y 63, que se encuentran cercanos al ligando en la región del lazo RT del dominio SH3. Estos efectos sugieren un cierto grado de intercambio conformacional lento en la superficie de unión de la quimera SPCp41. Se observan también aumentos en la relación R_2/R_1 para el resto 17 y de manera más prominente para el resto 33, en el giro distal, ambos más alejados del sitio de unión.

En la Figura 4.5.1.B se aprecia también un aumento de la relación R_2/R_1 para el mutante P74A debido al efecto de la agregación y que afecta a la mayor parte de los restos del dominio SH3, excepto para aquellos que presentan movimientos de alta frecuencia que compensan el efecto de agregación sobre los valores de R_2 . En el mutante P74A el resto 33 todavía presenta una elevada relación R_2/R_1 que indica que presenta un fenómeno de intercambio conformacional en el giro distal similar al de la quimera SPCp41. En la secuencia de conexión y en la región del ligando se percibe una clara disminución de la relación R_2/R_1 para el mutante P74A en comparación con SPCp41, que indica que la relajación en estas zonas está gobernada por movimientos de alta frecuencia. Por tanto, la mutación P74A hace que se desplace la escala de tiempo de los movimientos conformacionales de la cadena principal en la región del ligando p41, siendo relativamente lentos en el caso de la quimera SPCp41 y mas rápidos en el mutante P74A.

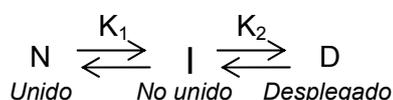
Los resultados de la caracterización estructural y dinámica del mutante P74A nos permiten, por tanto, concluir que las mutaciones Pro-Ala realizadas en la secuencia del p41 en la quimera SPCp41 producen una debilitación de las interacciones en la superficie de unión del dominio SH3 que se traduce en un aumento de los movimientos de alta frecuencia del péptido. Los datos de DC y fluorescencia de los mutantes descritos anteriormente indican que los restos aromáticos que forman parte del sitio de unión sufren un aumento del área de la superficie expuesta al disolvente y además se produce una reducción de la fracción de hélice de PPII del ligando. En definitiva, todos estos resultados confirman la existencia de un equilibrio conformacional dinámico en las quimeras que afecta a la interacción del ligando con el sitio de unión y que debe ser considerado en nuestro análisis termodinámico.

4.6. Análisis termodinámico de las curvas de capacidad calorífica molar parcial de desplegamiento utilizando un modelo de tres estados

Los resultados del análisis estructural y de relajación de RMN indican cambios significativos sobre la estructura, la dinámica conformacional y el grado de exposición al disolvente que afectan a la interfase del sitio de unión, provocados por las diferentes mutantes Pro-Ala en SPCp41. Estos cambios son más acusados para P69A, P74A y P77A, moderados para P73A and P76A y menores para P75A, y están bastante bien correlacionados con los cambios en estabilidad térmica medidos por CDB. Por el contrario, los cambios de capacidad calorífica de desplegamiento determinados

utilizando el modelo de dos estados varían en el orden opuesto con las mutaciones, lo que es incongruente con los cambios estructurales y en estabilidad.

Aunque los cambios de capacidad calorífica de desplegamiento de las proteínas han sido interpretados frecuentemente en términos de efectos en la hidratación (Spolar y Record, 1994; Makhatadze y Privalov, 1995), se han propuesto otras causas posibles (Prabhu y Sharp, 2005), aunque su magnitud y su signo no han sido aún esclarecidos. El hecho de que se produzcan otros procesos de equilibrio además del propio plegamiento puede tener una contribución importante a la capacidad calorífica de las proteínas (Fisher *et al.*, 1981; Eftink *et al.*, 1983). En el caso de nuestra quimera SPCp41 hemos demostrado que las mutaciones Pro-Ala producen una debilitación de las interacciones y un aumentando de las fluctuaciones conformacionales en la superficie de interacción, lo que pone de manifiesto la existencia de un equilibrio dinámico entre un estado “unido”, en el que el ligando interacciona intramolecularmente con el sitio de unión del dominio SH3, y un estado “no unido”, en el que el ligando se encuentra desordenado y el sitio de unión expuesto al disolvente. Este equilibrio conformacional también está presente, aunque de forma menos apreciable, en la quimera SPCp41. Si estos dos estados conformacionales tienen distintas entalpías y capacidades caloríficas, el equilibrio se desplazará con la temperatura cambiando su población relativa. El calor implicado en este desplazamiento del equilibrio debería tener influencia en las curvas de CDB y este efecto podría explicar las anomalías descritas anteriormente en las dependencias de la entalpía de desplegamiento con la temperatura (Figura 4.3.2). Este escenario se podría simplificar considerando el siguiente modelo de desplegamiento de tres estados:



donde N e I son los estados donde la interfase de unión SH3-péptido presenta una conformación “unido” y “no unido”, respectivamente, y D es el estado desnaturalizado de la proteína. K_1 y K_2 son las constantes de equilibrio de cada transición. Con este modelo de equilibrio de tres estados llevado a cabo el ajuste no lineal de todas las curvas de C_p medidas a distintos valores de pH para cada proteína. Las ecuaciones de este modelo y los detalles de los ajustes están explicados en la Sección 6.6.2.2 de Materiales y Métodos.

En el análisis consideramos que en el estado I los 12 últimos restos del extremo C-terminal, que incluyen al ligando, se encuentran completamente desordenados y expuestos al disolvente. También hemos considerado que la presencia de este segmento de cadena desordenada no influye sobre la termodinámica del desplegamiento del dominio SH3, por lo que hemos supuesto que la transición entre los estados I y D tiene las mismas magnitudes termodinámicas que el desplegamiento de dos estados de S19P20s. Esto nos permite incorporarlas como parámetros fijos en el ajuste de las curvas de las quimeras cada pH. Asimismo, las funciones de C_p de los estados I y D se han tomado de los análisis de las curvas de CDB de S19P20s y se han corregido teniendo en cuenta la contribución de los 12 últimos restos desordenados del extremo C-terminal, de acuerdo con Makhatadze y Privalov (Makhatadze y Privalov, 1990). Con estas suposiciones, los parámetros que se optimizaron en el ajuste global de las curvas de CDB de cada quimera son: la constante de equilibrio, K_1 a cada pH, el cambio de entalpía, ΔH_1 , y el cambio de

capacidad calorífica, ΔC_{p1} , para la transición entre los estados N e I a una cierta temperatura de referencia, además de la dependencia con la temperatura de la C_p para el estado N, que hemos supuesto lineal. Para simplificar aún más el análisis, supusimos que el cambio de entalpía, ΔH_1 , es aproximadamente independiente del pH en el intervalo estudiado, ya que no existen restos cargados en la secuencia del extremo C-terminal. Esta aproximación está de acuerdo con la baja dependencia con el pH de la entalpía de unión del decapeptido p41 al dominio Abl-SH3 (Palencia *et al.*, 2004) o al dominio Spc-SH3 (Casares *et al.*, 2007).

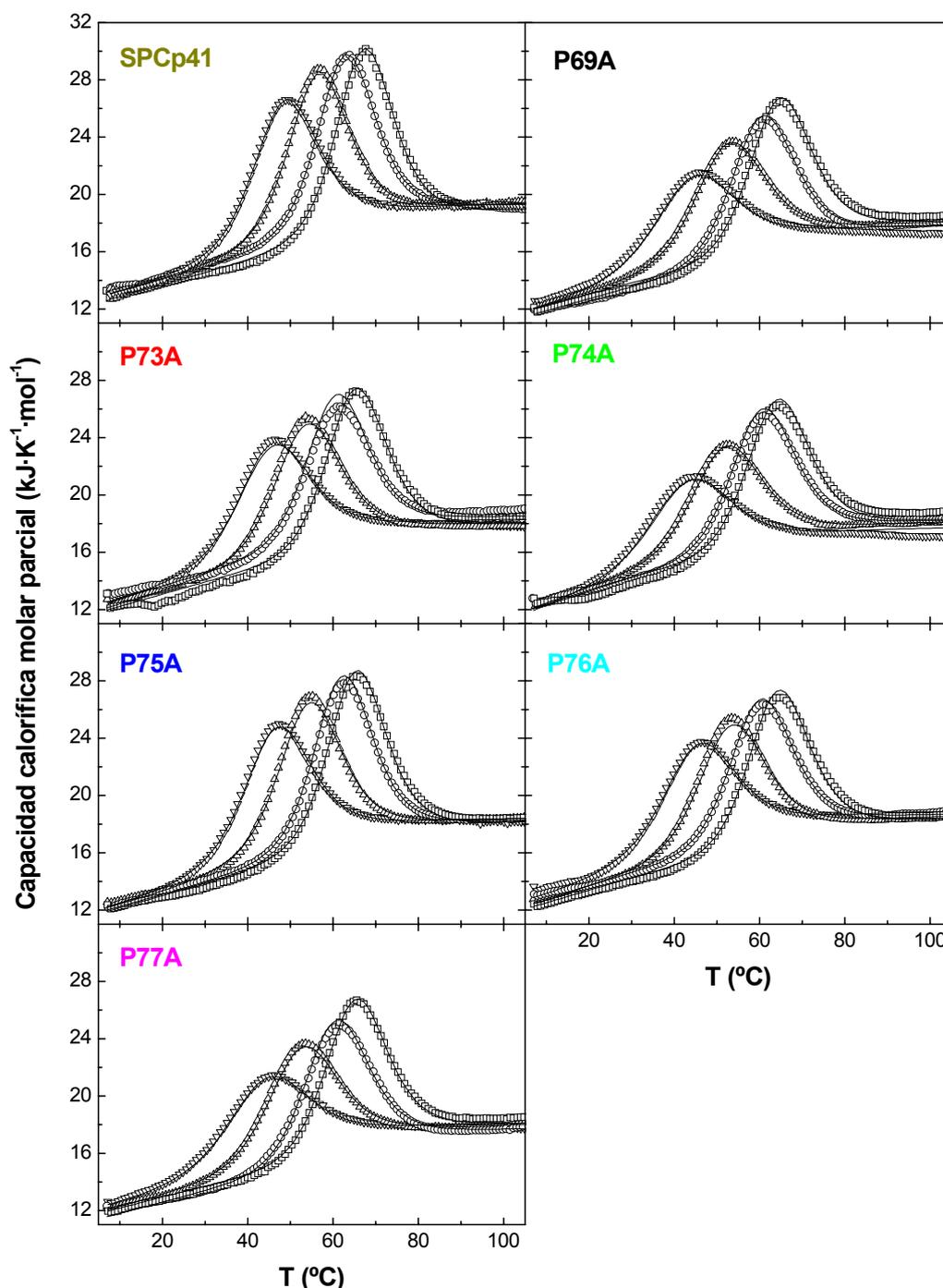


Figura 4.6.1. Dependencias de la capacidad calorífica molar parcial, C_p , con la temperatura de la quimera SPCp41 y de los mutantes Pro-Ala a distintos valores de pH. Los datos experimentales están representados con símbolos abiertos para cada pH: triángulos hacia abajo pH 2.5, triángulos hacia arriba pH 3.0, círculos pH 3.5 y cuadrados pH 4.0; y las líneas representan los resultados del análisis global aplicando el modelo de tres estados.

Tabla 4.6.1. Magnitudes termodinámicas* a 25°C obtenidas del ajuste de las curvas de CDB al modelo de tres estados para la interacción SH3-péptido evaluada a partir de la quimera SPCp41 y los mutantes Pro-Ala.

Proteína	pH	K_1	ΔG_1 (kJ·mol ⁻¹)	ΔH_1 (kJ·mol ⁻¹)	$T \cdot \Delta S_1$ (kJ·mol ⁻¹)	ΔC_{p1} (kJ·mol ⁻¹ K ⁻¹)
SPCp41	4.0	0.0234 ± 0.0008	9.31 ± 0.08	56.6 ± 0.9	47.3 ± 1.0	0.242 ± 0.011
	3.5	0.0288 ± 0.0010	8.80 ± 0.09		47.8 ± 1.0	
	3.0	0.0389 ± 0.0013	8.05 ± 0.08		48.6 ± 1.0	
	2.5	0.0541 ± 0.0017	7.22 ± 0.08		49.4 ± 1.0	
P69A	4.0	0.163 ± 0.011	4.49 ± 0.17	37.8 ± 1.4	33.3 ± 1.6	0.02 ± 0.04
	3.5	0.171 ± 0.011	4.37 ± 0.16		33.4 ± 1.6	
	3.0	0.35 ± 0.03	2.60 ± 0.21		35.2 ± 1.6	
	2.5	0.51 ± 0.04	1.69 ± 0.20		36.1 ± 1.6	
P73A	4.0	0.079 ± 0.006	6.28 ± 0.19	49.0 ± 1.8	42.8 ± 2.0	-0.01 ± 0.05
	3.5	0.098 ± 0.008	5.75 ± 0.20		43.3 ± 2.0	
	3.0	0.142 ± 0.011	4.83 ± 0.19		44.2 ± 2.0	
	2.5	0.170 ± 0.012	4.39 ± 0.17		44.7 ± 2.0	
P74A	4.0	0.153 ± 0.012	4.65 ± 0.19	47.2 ± 1.6	42.5 ± 1.8	-0.12 ± 0.04
	3.5	0.129 ± 0.012	5.08 ± 0.21		42.1 ± 1.8	
	3.0	0.40 ± 0.04	2.27 ± 0.25		44.9 ± 1.8	
	2.5	0.54 ± 0.05	1.52 ± 0.23		45.7 ± 1.8	

P75A	4.0	0.0449 ± 0.0021	7.69 ± 0.12	56.0 ± 1.2	48.3 ± 1.3	0.19 ± 0.03
	3.5	0.0384 ± 0.0018	8.08 ± 0.12		47.9 ± 1.3	
	3.0	0.076 ± 0.003	6.39 ± 0.10		49.6 ± 1.3	
	2.5	0.096 ± 0.004	5.80 ± 0.10		50.2 ± 1.3	
P76A	4.0	0.117 ± 0.006	5.31 ± 0.13	48.4 ± 1.1	43.1 ± 1.2	-0.06 ± 0.03
	3.5	0.131 ± 0.007	5.04 ± 0.13		43.4 ± 1.2	
	3.0	0.179 ± 0.009	4.26 ± 0.12		44.1 ± 1.2	
	2.5	0.200 ± 0.009	3.99 ± 0.11		44.4 ± 1.2	
P77A	4.0	0.181 ± 0.009	4.24 ± 0.12	28.2 ± 1.2	24.0 ± 1.3	0.014 ± 0.018
	3.5	0.228 ± 0.012	3.66 ± 0.13		24.5 ± 1.3	
	3.0	0.49 ± 0.03	1.78 ± 0.15		26.4 ± 1.4	
	2.5	0.73 ± 0.05	0.79 ± 0.17		27.4 ± 1.4	

* Los errores en los parámetros corresponden a un intervalo de confianza del 95% en los ajustes.

En la Figura 4.6.1 se muestran los termogramas experimentales de la dependencia de la capacidad calorífica molar parcial, C_p , con la temperatura de la quimera SPCp41 y de todos los mutantes Pro-Ala, para cada valor de pH, además de los ajustes globales por mínimos cuadrados no lineales utilizando el modelo de desplegamiento de tres estados. En la Tabla 4.6.1 se resumen los parámetros termodinámicos obtenidos de este análisis. A pesar de las restricciones que se han impuesto en el análisis, las curvas de C_p son bien reproducidas por el modelo, tanto en la forma como en su dependencia con el pH, lo que apoya la validez de las suposiciones que se han descrito antes.

Estos resultados confirman que las propiedades estructurales aparentes de los distintos mutantes Pro-Ala observadas a 25°C mediante DC, fluorescencia y RMN, son propiedades promedio representativas del equilibrio conformacional entre de los estados “unido” y “no unido”. La posición de este equilibrio a cada temperatura está representada para cada quimera por la constante de equilibrio K_1 . Las magnitudes termodinámicas de la Tabla 4.6.1 nos expresan el grado de importancia en la interacción SH3-ligando que tiene cada uno de los restos Pro mutados. Además, en todos los casos, la interacción SH3-péptido es termodinámicamente favorable aunque, en general, y especialmente para algunos mutantes como P69A, P74A y P77A a bajo pH, la energía de Gibbs de esta interacción es bastante baja.

Los valores de ΔH_1 y $T \cdot \Delta S_1$ son fuertemente positivos, lo que indica que la interacción intramolecular entre el dominio SH3 y el ligando está dominada por una entalpía favorable grande, parcialmente compensada por un término entrópico desfavorable. Esto también se suele observar en la unión de péptidos ricos en Pro a otros dominios SH3 (Renzoni *et al.*, 1996; Arold, S. *et al.*, 1998; Ferreon y Hilser, 2003; Ferreon y Hilser, 2004; Palencia *et al.*, 2004; Casares *et al.*, 2007). La afinidad de la unión se reduce progresivamente para los distintos mutantes, produciéndose un considerable aumento en la entalpía y en la entropía de unión. Las diferencias son relativamente grandes entre los distintos mutantes. Así, por ejemplo, la diferencia de entalpía de unión entre SPCp41 y el mutante P77A es 28 kJ·mol⁻¹ a 25°C.

Las magnitudes termodinámicas de la Tabla 4.6.1 nos permitieron estimar la población relativa de cada estado en función de la temperatura para cada una de las proteínas quiméricas. En la Figura 4.6.2 se representan las poblaciones del estado “unido” para cada quimera a 25°C a cada valor de pH. Estas poblaciones se correlacionan muy bien con las variaciones observadas para las diferentes quimeras en la fluorescencia intrínseca del Trp y de la señal de DC en el UV-lejano que se muestran en las Figuras 4.2.2 y 4.2.4.B.

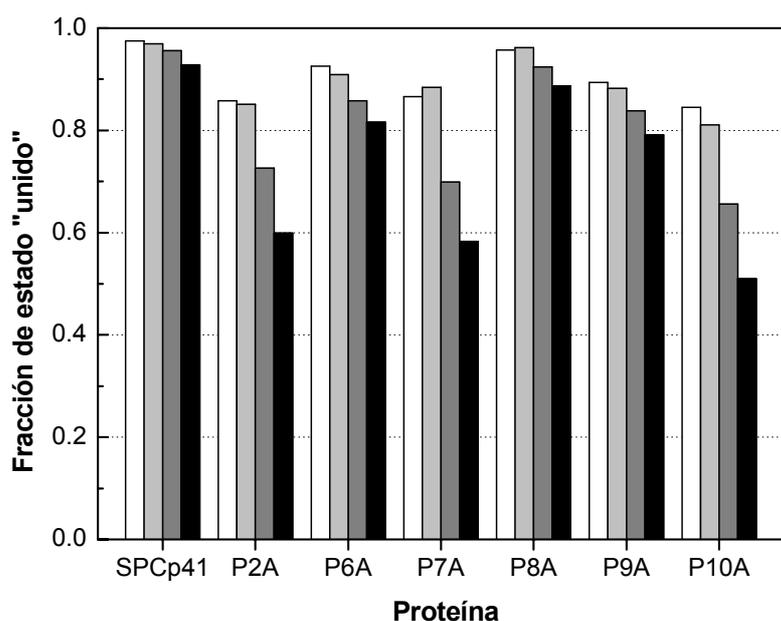


Figura 4.6.2. Fracción de estado "unido" de cada proteína quimérica a 25°C y a diferentes valores de pH: pH 4.0 en blanco; pH 3.5 en gris claro; pH 3.0 en gris; pH 2.5 en negro, que se han calculado a partir de las magnitudes termodinámicas de la Tabla 4.6.1.

4.7. Estudio termodinámico de la unión del permutante circular S19P20s al péptido p41 por calorimetría isotérmica de titulación (CIT)

Como una prueba adicional de la validez de nuestra metodología para el análisis de la termodinámica de interacción entre el dominio SH3 y los ligandos ricos en prolina, determinamos directamente las magnitudes termodinámicas de la unión del péptido aislado p41 al dominio S19P20s mediante calorimetría isotérmica de titulación (CIT), en glicina 20 mM, pH 3.5 y a 25°C. Los experimentos previos de titulación de S19P20s con p41 realizados mediante fluorescencia nos permitieron estimar la constante de asociación, K_a , con el resultado de $(6.3 \pm 1.3) \times 10^{-3} \text{ M}^{-1}$, que equivale a una constante de disociación, K_d , de $160 \pm 30 \mu\text{M}$ (Martin-Sierra *et al.*, 2003). Dado que esta afinidad es muy baja, para poder determinar de forma precisa los parámetros termodinámicos de unión diseñamos un experimento de CIT que tuviera un valor del parámetro c (que se define como el producto entre la constante de unión y la concentración de la proteína en la célula del calorímetro) dentro del intervalo adecuado, que está entre 1 y 100, aunque su valor óptimo está entre 10 y 100 (Wiseman *et al.*, 1989). Para realizar el experimento utilizamos una concentración de proteína en la célula del calorímetro de 0.44 mM y una concentración de péptido p41 en la jeringa del calorímetro de 10.83 mM, lo que nos dio un valor de 3 aproximadamente para el parámetro c , dentro del intervalo apropiado.

En la Figura 4.7.1 podemos ver los resultados experimentales obtenidos y en la Tabla 4.7.1 se resumen los parámetros termodinámicos de la interacción, derivados del ajuste por mínimos cuadrados a un modelo n sitios de unión iguales e independientes (Casares *et al.*, 2007). La constante de disociación obtenida por CIT, K_d , es de $260 \pm 20 \mu\text{M}$, que es ligeramente mayor que la determinada mediante fluorescencia, lo que confirma la baja afinidad entre p41 y el permutante S19P20s. La entalpía de unión es grande y negativa y se compensa con un valor de entropía muy desfavorable, lo cual se observa normalmente para este tipo de interacción. La energía de Gibbs estándar de unión, a pH 3.5 y 25°C, entre S19P20s y p41 es de $-20.44 \pm$

$0.12 \text{ kJ}\cdot\text{mol}^{-1}$. Estas magnitudes termodinámicas son similares a las determinadas para unión de p41 a la forma WT del dominio Spc-SH3 bajo condiciones similares (Casares *et al.*, 2007), aunque la afinidad del permutante circular S19P20s por p41 es ligeramente menor, lo que indica que la permutación de la secuencia del dominio SH3 y la apertura del lazo RT no tienen un efecto determinante en su interacción con el ligando p41.

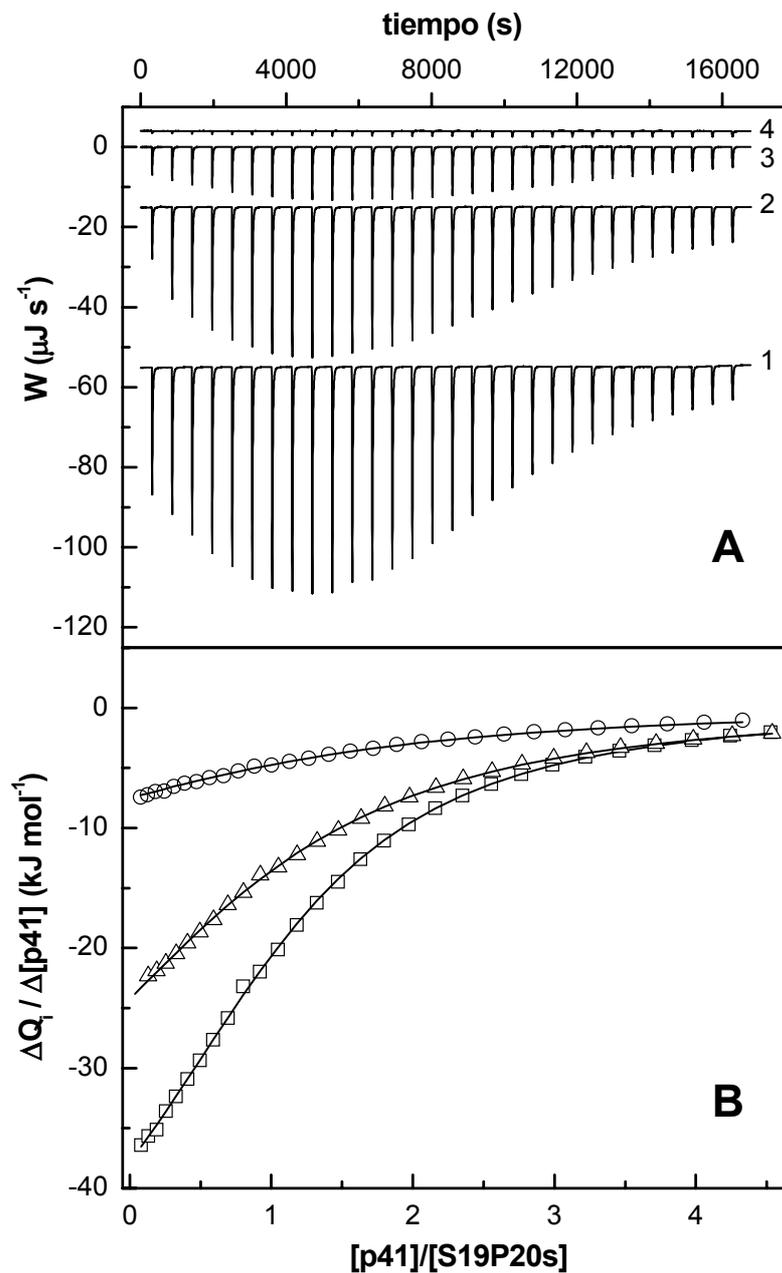


Figura 4.7.1. Experimentos de titulación del dominio S19P20s con el decapeptido p41 a 25°C en glicina 20 mM, pH 3.5, en ausencia y presencia del decapeptido P7Ap por calorimetría isotérmica de titulación (CIT). **A.** Termogramas obtenidos al inyectar volúmenes variables de una disolución de 10.83 mM de p41 en la célula del calorímetro conteniendo: (1) 0.44 mM de S19P20s; (2) 0.44 mM de S19P20s + 1.32 mM de P7Ap; (3) 0.46 mM de S19P20s + 5.17 mM de P7Ap; (4) sólo tampón. **B.** Curvas de titulación, normalizadas y corregidas, obtenidas de los termogramas que se muestran en el panel A. Los símbolos representan los datos experimentales: sin P7Ap cuadrados; 1.32 mM P7Ap triángulos y 5.17 mM P7Ap círculos. Las líneas representan los ajustes por mínimos cuadrados a un modelo de un sitio de unión, como se describe en Materiales y Métodos.

Tabla 4.7.1. Parámetros termodinámicos de la unión entre los decapeptidos p41 y P7Ap y el permutante circular S19P20s determinados por CIT a pH 3.5 y 25°C. Los errores corresponden a un intervalo de confianza del 95% en la determinación de los parámetros.

Péptido	K_a (M^{-1})	ΔH_a ($kJ \cdot mol^{-1}$)	n	ΔG_a ($kJ \cdot mol^{-1}$)
APSYSPPPPP (p41)	3800 ± 180	-58.1 ± 1.6	1.09 ± 0.02	-20.44 ± 0.12
APSYSPAPP (P7Ap) ^A	480 ± 100	-46 ± 13	1.09^B	-15.3 ± 0.5

^AAnalizado por el método de desplazamiento (Sigurskjold, 2000).

^BEl número de sitios se ha fijado en el ajuste de los datos experimentales al modelo correspondiente.

A partir de la energía de Gibbs estándar del proceso de unión entre S19P20s y p41 (interacción intermolecular), ΔG_{inter}^0 , que hemos calculado por CIT y que resulta igual a: $-20.44 \pm 0.12 \text{ kJ} \cdot \text{mol}^{-1}$, y de la energía de Gibbs estándar de la interacción intramolecular que hemos obtenido por CDB para la quimera SPCp41, ΔG_{intra}^0 , que resulta ser -8.80 ± 0.09 , podemos calcular el efecto que produce la ligadura covalente entre el ligando p41 y el dominio S19P20s sobre la energética de la unión a 25°C en el estado estándar (actividad 1 M) de la siguiente forma:

$$\Delta G_{lig}^0 = \Delta G_{intra}^0 - \Delta G_{inter}^0 \quad (4.7.1)$$

que resulta igual a $11.64 \pm 0.21 \text{ kJ} \cdot \text{mol}^{-1}$. Suponiendo que la ligadura no implica contribución entálpica importante, el valor de $-T\Delta S_{lig}^0$, resulta ser igual a $11.64 \pm 0.21 \text{ kJ} \cdot \text{mol}^{-1}$. Este valor aparentemente indicaría que el la unión covalente entre el péptido y el permutante contribuye a desestabilizar entrópicamente la interacción, lo que resulta sorprendente.

La contribución entrópica de introducir una ligadura covalente sobre la estabilidad de un dímero de dos hélices se ha calculado teóricamente por otros autores (Zaman *et al.*, 2002), resultando un valor para el término $-T\Delta S_{lig}^0$ de $13.4 \text{ kJ} \cdot \text{mol}^{-1}$ a 27°C, es decir, prácticamente igual a nuestro resultado. Esto indica por tanto que la ligadura covalente produce una desestabilización del sistema en condiciones estándar (actividad 1 M), de forma similar a nuestro sistema. Este resultado es, al menos, paradójico, puesto que no es esperable que se produzca una desestabilización. Sin embargo, la confusión procede del hecho de que para macromoléculas el estado estándar de actividad 1 M es irreal. Así, considerando un estado de referencia más coherente, como es 1 μM , los autores obtienen que la contribución entrópica derivada de la ligadura covalente pasa a ser fuertemente estabilizadora ($-21 \text{ kJ} \cdot \text{mol}^{-1}$), como cabía esperar inicialmente (Zaman *et al.*, 2002). En nuestro caso, si hacemos estos cálculos para un estado de referencia con una concentración de reactivos de 50 μM , el término $-T\Delta S_{lig}^0$ resulta ser igual a $-12.9 \pm 1.3 \text{ kJ} \cdot \text{mol}$ lo cual indica, de forma análoga, que la ligadura covalente añade estabilidad a las interacciones que se producen en la superficie de unión en SPCp41.

La secuencia de conexión entre S19P20s y el ligando aparece bastante desordenada en la estructura por RMN de SPCp41 y no presenta contactos con el resto de la estructura, tal y como se describió en el Capítulo anterior (Candel *et al.*, 2007). La alta flexibilidad conformacional de esta secuencia de conexión se manifiesta también en sus parámetros de relajación de RMN, según lo descrito anteriormente. Si admitimos, por tanto, que la contribución de la unión covalente entre el péptido y el dominio a la energía de la interacción es la misma para todas las proteínas quiméricas, utilizando los valores de la Tabla 4.6.1 se pueden estimar las constantes de disociación correspondientes a la interacción de S19P20s con cada péptido mutante aislado. Por ejemplo, K_d sería 430 μ M para P75A, 1.5 mM para P74A o 2.6 mM para P77A, a pH 3.5 y 25°C. Estas afinidades son extremadamente bajas, lo que hace que el estudio de estas interacciones mediante experimentos de titulación directa del permutante S19P20s con estos péptidos sea prácticamente inviable. Una posible solución a este problema consistiría en la utilización de un experimento de CIT por desplazamiento, en el cual el complejo de una proteína con un ligando de baja afinidad se titula con otro ligando de mayor afinidad (Sigurskjold, 2000). Ese ha sido el procedimiento para verificar por CIT las magnitudes termodinámicas de la unión obtenidas para el mutante P74A mediante nuestro análisis de los datos de CDB. Hemos hecho experimentos de CIT de desplazamiento en glicina 20 mM, pH 3.5 y 25°C para determinar directamente las magnitudes termodinámicas de la unión de mutante P7A del péptido p41 (P74A si seguimos la numeración de la quimera SPCp41), denominado P7Ap, al permutante circular S19P20s. Hemos elegido esta mutación porque la Pro que se sustituye es una de las más importantes para la unión. En estos experimentos hemos titulado una disolución 0.44 mM de S19P20s en presencia del péptido P7Ap en la célula calorimétrica con una disolución de 10.83 mM de p41 desde la jeringa. Concretamente, se han realizado dos experimentos independientes con concentraciones de P7Ap de 1.32 y 5.17 mM. Los detalles de estos experimentos se han descrito en la Sección 6.11 de Materiales y Métodos. En la Tabla 4.7.1 se recogen las magnitudes termodinámicas que resultan del análisis. Con este tipo de diseño experimental hemos podido determinar la constante de disociación del péptido P7Ap, que resulta ser 2.1 ± 0.4 mM y que esta de acuerdo con los resultados obtenidos por CDB para el mutante P74A de la quimera SPCp41. Aunque la precisión en la determinación de la entalpía de unión mediante este tipo de experimento es baja, coincide razonablemente bien con los datos obtenidos mediante CDB.

4.8. Proteínas quiméricas de fusión como herramientas útiles para el análisis termodinámico de las interacciones débiles proteína-ligando

Uno de los grandes problemas en el análisis termodinámico de las interacciones SH3-ligando es que sus afinidades normalmente están por debajo del intervalo accesible para ser estudiadas por CIT (Wiseman *et al.*, 1989). Para realizar estos experimentos de CIT se necesitan decenas de miligramos de péptidos sintéticos y de proteína, debido a la baja afinidad que presentan estas interacciones. Esto hace inviable el uso del CIT para un estudio extenso de la energética de la unión, a causa del elevado coste y a problemas colaterales como la limitada solubilidad y/o agregación de las muestras. Por otra parte, la CDB se ha utilizado frecuentemente con anterioridad para el análisis de la termodinámica de la interacción proteína-ligando, tanto en nuestro grupo de investigación como en otros, presentando ciertas ventajas con respecto a la CIT en muchos casos (Brandts y Lin, 1990; Martínez *et al.*, 1994; Martínez *et al.*, 1995; Casares *et al.*, 2007). En este trabajo hemos demostrado que los experimentos de CDB combinados con la mutagénesis dirigida son muy sensibles a

los cambios provocados por las mutaciones sobre la energética de la unión del ligando y por tanto, constituyen una herramienta muy útil en el análisis termodinámico de estas interacciones. Esta metodología termodinámica es completamente novedosa, aunque en otros estudios se han utilizado proteínas de fusión artificiales para analizar estructuralmente las interacciones entre dominios SH3 y péptidos ricos en prolina (Gmeiner *et al.*, 2001), o también en dominios GYF (Freund *et al.*, 2003). Además, nuestro abordaje se puede extender a otras interacciones biomoleculares cuando ambas moléculas se puedan unir covalentemente mediante ingeniería de proteínas o por métodos químicos.

4.9. Importancia de los restos Pro en la unión SH3-p41

El análisis mutacional que realizamos con las proteínas quiméricas nos permitió establecer claramente la importancia que tiene cada resto de Pro del péptido p41 en la interacción SH3-péptido. Ayudándonos de la Figura 4.1.1, podemos observar que la mutación P75A es la que parece tener menor importancia en la interacción, puesto que este resto está expuesto al disolvente y no presenta contactos entre su cadena lateral y el dominio SH3 (Candel *et al.*, 2007). La sustitución de Pro69, Pro74 y Pro77 tiene, por el contrario, el mayor impacto sobre la unión. Estas tres prolinas son esenciales para la unión de péptidos ricos en Pro al dominio Abl-SH3 (Ren *et al.*, 1993) y, además, la Pro74 y la Pro77 ocupan posiciones conservadas en la secuencia consenso de los ligandos de los dominios SH3 ($\Phi P x \Phi P$) (Kay, B. K. *et al.*, 2000). La Pro73 y la Pro76 son intermedias en el orden de importancia en la interacción y ocupan posiciones menos conservadas en la secuencia consenso. Estos resultados, están de acuerdo con el modelo de Lim para las interacciones SH3-péptido (Lim *et al.*, 1994; Nguyen, Jack T. *et al.*, 1998; Musacchio, 2003), con la Pro73 y la Pro76 en una posición de “*empaquetamiento interno*”, de acuerdo con la terminología del modelo, y la Pro74 y la Pro77 en una situación de “*empaquetamiento externo*”, que es la más apropiada para que un resto de Pro establezca interacciones óptimas con los bolsillos hidrofóbicos de la superficie del dominio SH3. La cadena lateral de la Pro69 también está orientada favorablemente para interactuar hidrofóticamente con el bolsillo “compás” del sitio de unión, lo que explica la fuerte desestabilización de la mutación P69A.

4.10. El impacto de las mutaciones Pro-Ala sobre la termodinámica de la interacción SH3-ligando

Los parámetros termodinámicos obtenidos para la unión SH3-p41 que se muestran en la Tabla 4.6.1, revelan que en todas la quimeras la unión esta dirigida por un cambio de entalpía negativo favorable, parcialmente compensado por una entropía de unión negativa desfavorable, como se observa normalmente para las interacciones SH3-péptido (Renzoni *et al.*, 1996; Arold, S. *et al.*, 1998; Ferreon y Hilser, 2003; Ferreon y Hilser, 2004; Palencia *et al.*, 2004; Casares *et al.*, 2007). A medida que en los diferentes mutantes la afinidad de la unión se reduce progresivamente, las magnitudes de la entalpía y la entropía disminuyen de manera considerable, es decir, tanto la entalpía como la entropía de unión se hacen menos negativas. Las diferencias que se observan entre los mutantes son relativamente grandes. Por ejemplo, la diferencia de entalpía entre la quimera SPCp41 y el mutante P77A es aproximadamente $28 \text{ kJ}\cdot\text{mol}^{-1}$ a 25°C . Estos resultados son difíciles de explicar simplemente por la pérdida de unos pocos contactos de van der Waals en el estado “*unido*” a causa de cada mutación Pro-Ala. Además, el efecto entálpico de la pérdida

de los contactos de van der Waals es parcialmente compensado por un cambio opuesto en el balance de entalpía de hidratación (Hilser *et al.*, 1996).

Los cambios de hidrofobicidad en la interfase SH3-péptido producidos por las mutaciones Pro-Ala no nos proporcionan tampoco una explicación adecuada para los cambios en la entropía de unión. El efecto hidrofóbico produce una fuerte contribución favorable (positiva) a la entropía de unión, debida a la liberación de moléculas de agua que se produce por la interacción. La sustitución por alanina de una prolina que interacciona con la superficie del dominio SH3 produce a lo sumo una pequeña disminución del área que queda enterrada con la unión, lo cual, en contraste con nuestras observaciones, debería implicar únicamente una ligera disminución de la entropía de unión con la mutación Pro-Ala.

El plegamiento del ligando para adquirir la conformación de PPII debería influir fuertemente en la termodinámica de este tipo de interacción (Ferreon y Hilser, 2004). En ausencia de otros efectos, se podría esperar que la sustitución en el ligando de una prolina por alanina produjera una entropía de unión menos favorable, es decir, más negativa, debido al aumento de la libertad conformacional del ligando libre. En nuestro caso, esta contribución se podría evaluar con la mutación de la Pro75, que tiene la cadena lateral expuesta al disolvente, y que por tanto, no está implicada directamente en la interfase de la interacción. Sin embargo, los resultados obtenidos para el mutante P75A indican una entropía de unión ligeramente más negativa, que está acompañada por una entalpía negativa ligeramente menor. Ambos cambios tienen una magnitud de $<1 \text{ kJ}\cdot\text{mol}^{-1}$, que prácticamente se puede considerar dentro del error experimental (ver Tabla 4.6.1). Este resultado es un tanto inesperado, dado que las mutaciones Pro-Ala deberían aumentar la degeneración conformacional del resto mutado y del resto que le precede (Daquino *et al.*, 1996). Un estudio de CIT previo mostró que la sustitución de una prolina similar a la P75 por Leu en el ligando p41 aislado, produjo un cambio de entalpía desfavorable, parcialmente compensado por una entropía de unión al dominio Abl-SH3 más favorable, lo cual era contrario a los resultados esperados inicialmente (Palencia *et al.*, 2004). Para explicar este comportamiento inesperado se propuso la contribución de otros efectos, como la capacidad aceptora acentuada del grupo carbonilo de la prolina para formar un enlace de hidrógeno o la existencia de efectos cooperativos que se propagan vía interacciones específicas mediadas por moléculas de agua a través del ligando y el dominio SH3. En contraste con estos resultados, Hilser y colaboradores estudiaron la unión del péptido Sos al extremo C-terminal del dominio SH3 de Sem-5, y describieron que la mutación de una prolina que se encuentra expuesta al disolvente (equivalente a la Pro75 en la quimera SPCp41), por glicina o alanina, produce una entalpía de unión más negativa (de aproximadamente $-4 \text{ kJ}\cdot\text{mol}^{-1}$), parcialmente compensada por una entropía de unión más negativa (de aproximadamente $-2 \text{ kJ}\cdot\text{mol}^{-1}$) (Ferreon y Hilser, 2003; Hamburger *et al.*, 2004). Una posible explicación para estas diferencias es que el ligando p41 y sus mutantes podrían tener una mayor predisposición a formar una hélice de PPII que los péptidos Sos. Cualquier resto que va seguido por una prolina tiene restringida la conformación de su cadena principal por efectos estéricos a ocupar principalmente la región de los ángulos ϕ y ψ correspondiente a láminas β (Macarthur y Thornton, 1991; Creamer, 1998; Rucker *et al.*, 2003). Por esta razón, el aumento de libertad conformacional del ligando “no unido” por causa de cada mutación Pro-Ala debe de ser relativamente pequeño para la mayoría de los mutantes de SPCp41, como se observa por ejemplo para P75A. Por razones similares, las diferencias en entalpía y entropía de unión entre los distintos mutantes Pro-Ala no se pueden considerar debidas a posibles diferencias en el coste energético de adoptar el ligando una conformación en PPII. Por tanto, nuestros resultados no se pueden explicar como causados por cambios en la libertad

conformacional del ligando con las mutaciones y tiene que haber otras contribuciones que expliquen este comportamiento.

En la bibliografía se han encontrado contribuciones adicionales a la termodinámica de la interacción SH3-péptido, que consideran la presencia de una red de enlaces de hidrógeno mediados por moléculas de agua en la interfase de unión y una reducción en la dinámica conformacional de la molécula como consecuencia de la interacción (Lim *et al.*, 1994; Wang, C. Y. *et al.*, 2001a; Palencia *et al.*, 2004; Casares *et al.*, 2007). El significado de estos efectos en la determinación de la energética de la unión todavía no se comprende bien. Es muy probable que su contribución dependa de los detalles específicos de la interacción entre cada dominio SH3 y su ligando respectivo. Por ejemplo, la gran similitud entre las estructuras de los complejos del péptido p41 con los dominios SH3 de Abl y Spc sugirió que las grandes diferencias observadas entre estos dos dominios en las magnitudes termodinámicas de la unión podrían estar relacionadas con la ausencia de algunas de las interacciones mediadas por moléculas de agua (Casares *et al.*, 2007). No obstante, si los efectos de las mutaciones observados en las proteínas quiméricas fueran ante todo debidos a la desestabilización por interacciones específicas de moléculas de agua, podríamos haber encontrado los efectos más grandes en posiciones específicas mientras que, en lugar de esto, se observa una aparente periodicidad de los efectos energéticos en la hélice de PPII. Como consecuencia, aunque nuestros resultados no descartan la influencia de interacciones mediadas por moléculas de agua, una contribución importante a los cambios termodinámicos producidos por las mutaciones Pro-Ala puede deberse a su efecto sobre la dinámica conformacional de los complejos SH3-péptido, como se discutirá más adelante.

4.11. La interacción intramolecular SH3-ligando muestra una fuerte compensación entalpía-entropía

Hemos observado una fuerte compensación entalpía-entropía en la termodinámica de la interacción SH3-péptido para las proteínas quiméricas como podemos observar en la Figura 4.11.1. La existencia de compensación entre entalpía y entropía (CEE) es una propiedad ubicua de los procesos biológicos (Lumry y Rajender, 1970) y ha sido interpretada como una propiedad natural de interacciones múltiples, débiles y no-covalentes, como las que existen en sistemas macromoleculares en agua (Dunitz, 1995; Cooper *et al.*, 2001). Un comportamiento de CEE, similar al que nosotros hemos encontrado, ha sido descrito recientemente para la unión entre la anhidrasa carbónica bovina y diferentes bencil-sulfonamidas para-sustituidas con cadenas de oligo-glicina, oligo-sarcosina o oligo-etilenglicol de diferentes longitudes (Krishnamurthy *et al.*, 2006). Estos autores han observado un aumento en la movilidad en la interfase de unión en los complejos con el incremento en la longitud del ligando, lo que origina una entropía de unión menos desfavorable y una entalpía de unión menos favorable. Otros ejemplos de CEE en procesos de unión se han descrito recientemente (Williams *et al.*, 2004). La fuerte CEE de nuestros resultados indica de forma similar que la movilidad conformacional en la conformación “unida” de las proteínas quiméricas disminuye conforme se produce un aumento en la afinidad como consecuencia de la optimización cooperativa de las interacciones en la interfase de unión. Esta conclusión está de acuerdo con los parámetros de relajación de RMN, que muestran un aumento significativo de los movimientos rápidos en la interfase de unión del mutante P74A en comparación con la quimera SPCp41.

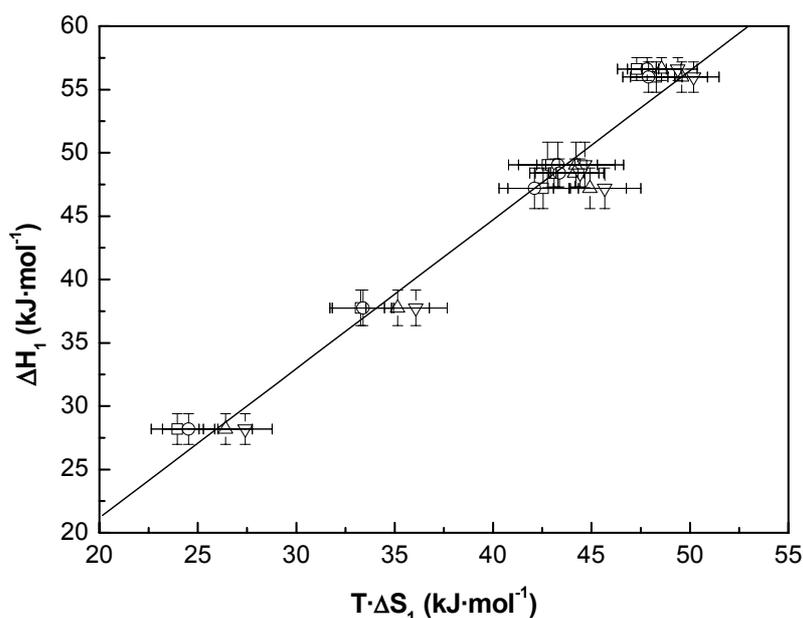


Figura 4.11.1. Compensación entalpía-entropía de la unión intramolecular SH3-péptido en las proteínas químicas. Los valores de ΔH_1 y $T \cdot \Delta S_1$ se muestran en la Tabla 4.6.1 y han sido obtenidos de los ajustes de los datos de CDB al modelo de tres estados que representan en la Figura 4.6.1, a distintos valores de pH: triángulos hacia abajo pH 2.5, triángulos hacia arriba pH 3.0, círculos pH 3.5 y cuadrados pH 4.0.

La disminución de la movilidad conformacional podría no estar restringida exclusivamente a la interfase de unión, como se ha mostrado en varios ejemplos de unión entre péptidos ricos en prolina y dominios SH3 o en otros complejos proteína-ligando (Mayne *et al.*, 1992; Wang, C. Y. *et al.*, 2001a; Wildes y Marqusee, 2005; Polshakov *et al.*, 2006; Casares *et al.*, 2007). La propagación de largo alcance de rigidez conformacional podría producir, por tanto, una amplificación de los cambios energéticos locales (tanto de entalpía como de entropía) producidos por las mutaciones puntuales. Sin embargo, el estudio comparativo de la relajación de RMN entre la quimera SPCp41 y el mutante P74A no proporciona, desafortunadamente, información acerca de efectos de largo alcance en el dominio SH3 debido al proceso de agregación que sufre el mutante, lo que impide un análisis de los datos en mayor profundidad.

4.12. Importancia biológica de la unión intramolecular frente a la unión intermolecular de secuencias de baja afinidad

A pesar de las bajas afinidades de unión entre los péptidos ricos en prolina aislados y el permutante S19P20s, hemos demostrado que, cuando los unimos covalentemente a través de una corta secuencia de conexión, la unión intramolecular ejerce un fuerte efecto termodinámico sobre la estabilidad de las quimeras. La unión intramolecular mediada por dominios SH3 juega un papel esencial en la regulación de muchas proteínas como, por ejemplo, en varias familias de las tirosinas quinasas y se sabe que los ligandos intramoleculares de estas proteínas son marcadamente diferentes a las secuencias de alta afinidad para los dominios SH3 (Andreotti *et al.*, 1997; Nguyen, J. T. y Lim, 1997; Boggon y Eck, 2004).

Zhou ha descubierto recientemente una relación cuantitativa entre las constantes de unión intramolecular e intermolecular entre dominios SH3 y sus ligandos ricos en prolina (Zhou, 2006): $K_i = K_b \cdot p$, donde K_i y K_b son las constantes de unión intramolecular e intermolecular respectivamente, y el parámetro p juega el papel de una concentración efectiva, que generalmente es del orden de mM. Esta concentración efectiva se determina considerando la longitud y la flexibilidad de la secuencia de conexión, y es el resultado de las restricciones de distancia impuestas entre las dos especies interaccionantes (Zhou, 2006). Nuestros resultados procedentes del análisis de CDB y CIT para la unión de los péptidos p41 y P7Ap al dominio S19P20s (Tablas 4.6.1 y 4.7.1) permiten estimar un valor para la concentración efectiva, p , de entre 9 y 16 mM, que están en buen acuerdo con los valores predichos por Zhou para el linker utilizado en nuestras quimeras (Zhou, 2006). El delicado balance, finamente ajustado, entre la concentración efectiva y la afinidad intrínseca entre las especies interaccionantes parece ser crucial para mantener un equilibrio correcto entre las formas activas e inactivas de las tirosinas quinasas, permitiendo alguna actividad basal y al mismo tiempo facilitando la activación eficiente de las quinasas por una amplia variedad de señales (Ayrapetov *et al.*, 2006). Se ha demostrado también que la longitud de la secuencia de conexión entre los dominios SH3 y las secuencias ricas en prolina es esencial para determinar la competición entre la unión intra- e intermolecular en la familia de las Tec-quinasas (Laederach *et al.*, 2003).

Es por tanto de gran importancia entender cuantitativamente los efectos energéticos provocados por la ligadura covalente entre los dominios SH3 y sus ligandos a través de secuencias de conexión flexibles y, para esto, la metodología que hemos empleado en este trabajo, que combina la mutagénesis dirigida con el análisis mediante CDB de proteínas quiméricas, puede servir como una poderosa herramienta para estudiar los efectos termodinámicos que se producen al variar la longitud y la composición de las secuencias de conexión, proporcionando además un método adecuado para validar el trabajo teórico.

CAPÍTULO 5:
ESTUDIO CINÉTICO DEL PLEGAMIENTO DE LA
QUIMERA SPC_{p41}

CAPÍTULO 5: ESTUDIO CINÉTICO DEL PLEGAMIENTO DE LA QUIMERA SPCp41

En este Capítulo se describe la caracterización de la cinética de plegamiento y desplegamiento de la quimera SPCp41 y sus mutantes Pro-Ala, en comparación con el permutante circular S19P20s, para analizar el grado de influencia del segmento adicional de cadena incluido en las quimeras sobre su mecanismo de plegamiento.

Este estudio cinético nos ha permitido obtener información energética acerca del estado de transición de plegamiento, que puede evidenciar aspectos relevantes del mecanismo de plegamiento de SPCp41, en comparación con la del dominio SH3 original.

5.1. Cinéticas de plegamiento de la quimera SPCp41 mediante el método del flujo detenido (*stopped-flow*)

Para evaluar la cinética de plegamiento de una proteína hemos de medir experimentalmente, por un lado el plegamiento y por otro el desplegamiento de la proteína, que se llevan a cabo de forma independiente. En nuestro caso, hemos estudiado las cinéticas de plegamiento y desplegamiento de la quimera SPCp41 en tampón fosfato sódico 50 mM, pH 7.0 y a 25°C. Cada una de las cinéticas experimentales obtenidas es el resultado de medir la intensidad de fluorescencia a medida que transcurre el proceso de plegamiento o desplegamiento de la proteína en un instrumento de flujo detenido. La cinética se registra inmediatamente después de la mezcla rápida de dos disoluciones de diferente volumen (en relación 1:10): una disolución de proteína y de agente desnaturizante y otra que sólo contiene el agente desnaturizante, en nuestro caso urea. Modificando adecuadamente la concentración de urea de cada disolución se consigue un cambio instantáneo en la concentración del agente desnaturizante lo que inicia el proceso de desplegamiento o plegamiento de la proteína (ver con más detalle en la Sección 6.12 de Materiales y Métodos).

En el caso de que los procesos de plegamiento y desplegamiento tengan lugar sin asociaciones intermoleculares, como ocurre con SPCp41 según hemos comprobado a partir del estudio del efecto de la concentración sobre su desplegamiento térmico mediante CDB (Sección 4.3 del Capítulo 4), las reacciones son de primer orden y la evolución temporal de la intensidad de fluorescencia se puede representar como una suma de n exponenciales, expresadas en función de las correspondientes constantes de velocidad (k_i) y amplitudes (A_i) a través de la siguiente ecuación (Bachmann, A.; y Kiefhaber, 2005):

$$I_t = I_\infty + \sum_{i=1}^n A_i \cdot \exp(-k_i \cdot t) \quad (5.1.1)$$

donde I_t es la intensidad de fluorescencia a tiempo t e I_∞ es la intensidad de fluorescencia una vez finalizado el proceso. Las constantes de velocidad, k_i , dependen de parámetros externos como la temperatura, la presión y la concentración de agente desnaturizante. Las amplitudes, A_i , dependen de las concentraciones iniciales y finales de las especies presentes durante el proceso y, como veremos después, reflejan las poblaciones de estas especies (Bachmann, A.; y Kiefhaber, 2005).

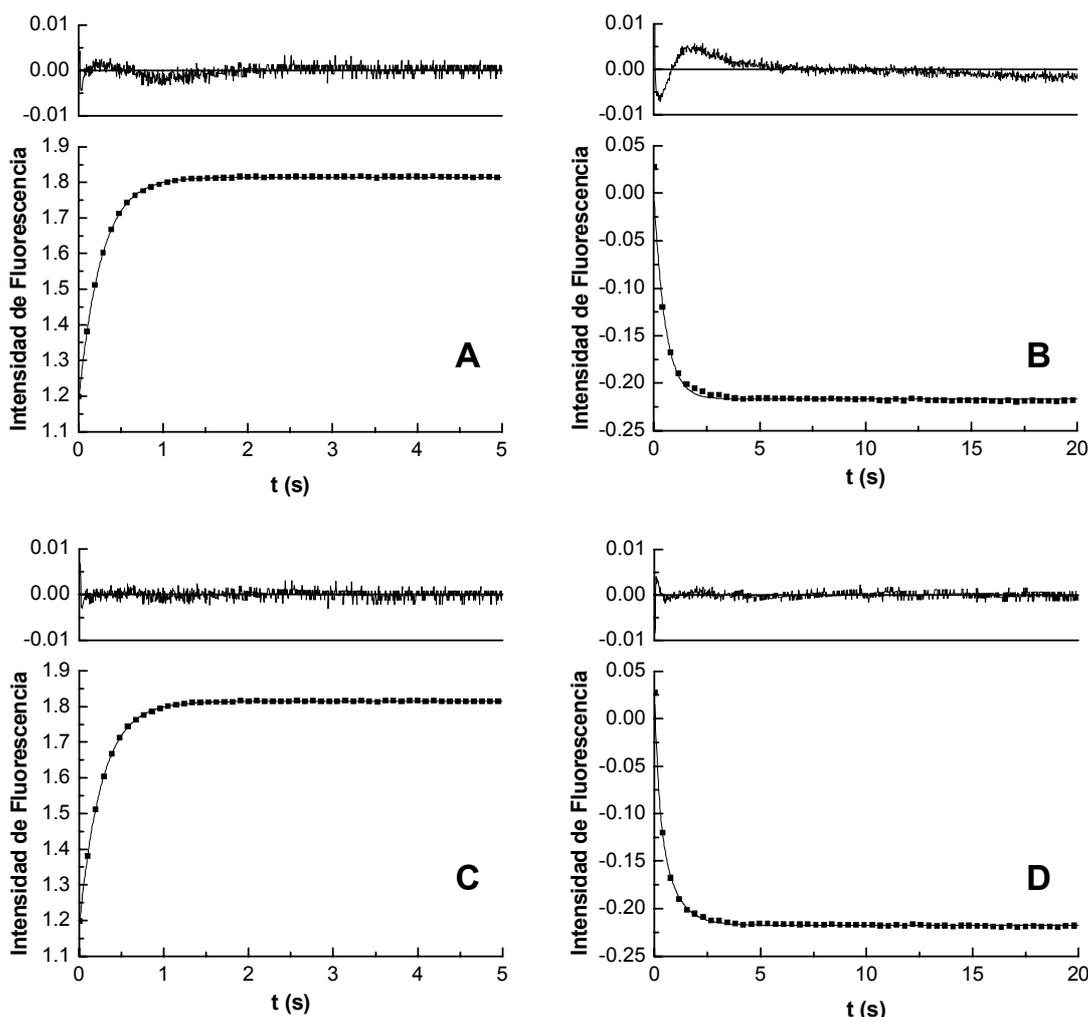


Figura 5.1.1. Perfiles cinéticos de la quimera SPCp41 registrados a 25°C con sus correspondientes gráficos de residuos. **A. y C.** Cinética de plegamiento, representada con símbolos, de una mezcla formada por la proteína a concentración 50 μM en tampón fosfato sódico 50 mM, pH 7.0 y urea 8 M con urea 2.0 M en el mismo tampón. En línea continua se representa el ajuste de los datos experimentales a una ecuación monoexponencial y a una ecuación biexponencial respectivamente. **B. y D.** Cinética de desplegamiento, representada con símbolos, de una mezcla formada por la proteína a concentración 50 μM en tampón fosfato sódico 50 mM, pH 7.0 con urea 7.0 M en el mismo tampón. En línea continua se representa el ajuste de los datos experimentales a una ecuación monoexponencial y a una ecuación biexponencial respectivamente.

En una primera aproximación, las curvas de plegamiento y de desplegamiento se analizaron mediante su ajuste no lineal utilizando una ecuación monoexponencial pero, como se observa en la Figura 5.1.1, el gráfico de los residuos del ajuste presenta desviaciones por encima del ruido experimental. Por esta razón, las cinéticas se analizaron realizando los ajustes no lineales con una ecuación biexponencial. En este caso, los ajustes de los datos son buenos y los gráficos de los residuos presentan una distribución adecuada.

De estos ajustes obtuvimos los valores de las constantes de velocidad, correspondientes a cada concentración de urea, para dos fases cinéticas aparentes, ambas presentes tanto en el proceso de plegamiento (p) como en el de desplegamiento (d). Así, los parámetros que se obtienen de cada ajuste son k_{p1} y k_{d1} , para la fase rápida que es la de mayor amplitud, y k_{p2} y k_{d2} , para la fase lenta de menor

amplitud. También se obtienen los valores de las amplitudes de cada fase, A_{p1} y A_{d1} , y A_{p2} y A_{d2} , y de la intensidad de fluorescencia al final de la reacción, I_{∞} .

A partir de las constantes de velocidad de las dos fases cinéticas se han construido los gráficos de Chevron para el plegamiento-desplegamiento de la quimera (como se explica en la Sección 6.12.4 de Materiales y Métodos), representados en la Figura 5.1.2. Los valores de las constantes cinéticas aparentes de la fase rápida describen un Chevron sin desviaciones apreciables de la linealidad en ninguna de las ramas de sus extremos, lo que *a priori* indicaría una distribución de sólo dos estados poblados durante esta fase cinética (Bachmann, A.; y Kiefhaber, 2005). Los valores de las constantes cinéticas aparentes de la fase lenta presentan una mayor dispersión, muy probablemente debida a una mayor incertidumbre en su determinación, aunque también parecen presentar la típica forma en V del gráfico de Chevron. En la Figura 5.1.3 hemos representado las amplitudes de las dos fases cinéticas y, como podemos observar, las amplitudes de la fase lenta son casi despreciables en comparación con las de la fase rápida.

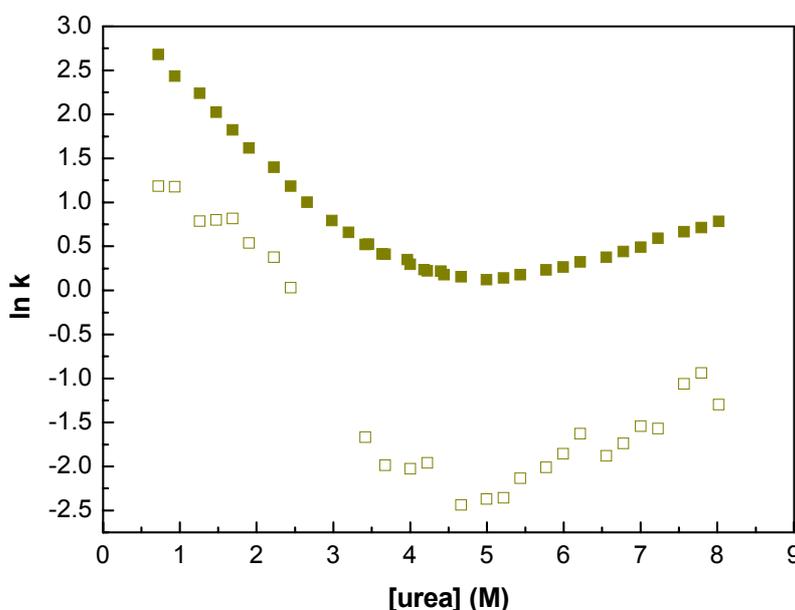


Figura 5.1.2. Gráfico Chevron de SPCp41 realizado a 25°C en tampón fosfato sódico 50 mM, pH 7.0. Los símbolos abiertos representan los datos obtenidos experimentalmente a partir de la realización de los experimentos cinéticos en las direcciones de plegamiento y desplegamiento de la constante de la fase lenta, y los símbolos cerrados de la constante de la fase rápida.

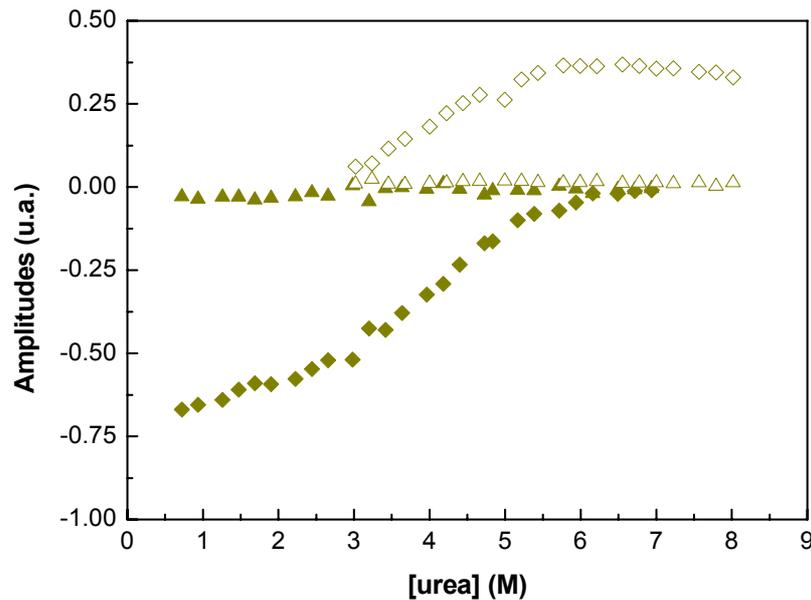


Figura 5.1.3. Amplitudes de las dos fases cinéticas de la quimera SPCp41. Los símbolos cerrados representan las amplitudes de plegamiento y los abiertos las de desplegamiento, con diamantes para la primera fase cinética, y con triángulos para la segunda.

5.2. Análisis cinético de la fase lenta del plegamiento de SPCp41

La presencia de dos fases en las cinéticas de plegamiento de una proteína puede ser debida a fenómenos de isomerización de prolinas y/o a la presencia de estados intermedios de plegamiento (Schmid, F. X., 2005). En una primera aproximación, la evidencia más significativa de la aparición de un intermedio de plegamiento suele ser una pérdida de la linealidad en alguna o ambas ramas del gráfico de Chevron. Aunque en nuestro caso esto no se ha observado, no deberíamos descartar esta opción sólo por este hecho. Así, otra de las posibles evidencias de la presencia de un intermedio consiste en la aparición de una segunda fase cinética que dibuje también un gráfico de Chevron, algo que en nuestro caso se podría intuir, como podemos observar en la Figura 5.1.2, aunque los datos presentan bastante dispersión.

Brandts *et al.* propusieron que la existencia de una fase lenta en el plegamiento de una proteína se debe generalmente a la influencia de la isomerización *cis-trans* en el estado desplegado de los enlaces peptídicos X-Pro, siendo X cualquier aminoácido, proceso que es relativamente lento, por lo que podría actuar como etapa limitante durante el plegamiento de una fracción pequeña de la proteína (Brandts *et al.*, 1975). Los enlaces peptídicos X-Pro pueden poblar la conformación *cis* mucho más fácilmente que cualquier otro enlace peptídico (Brandts *et al.*, 1975; Schmid, F. X. y Baldwin, 1978; Hurle y Matthews, 1987; Kiefhaber *et al.*, 1990), mientras que la relación *cis:trans* en los enlaces peptídicos de aminoácidos que no sean prolinas es muy desfavorable debido, en parte, a que el isómero *trans* presenta mayor entropía conformacional que el isómero *cis*, ya que éste presenta impedimentos estéricos debidos a la cadena lateral del aminoácido que le precede en la cadena polipeptídica. Para un polipéptido no ordenado, la relación *trans:cis* de los enlaces peptídicos X-Pro es aproximadamente 70:30 (Brandts *et al.*, 1975).

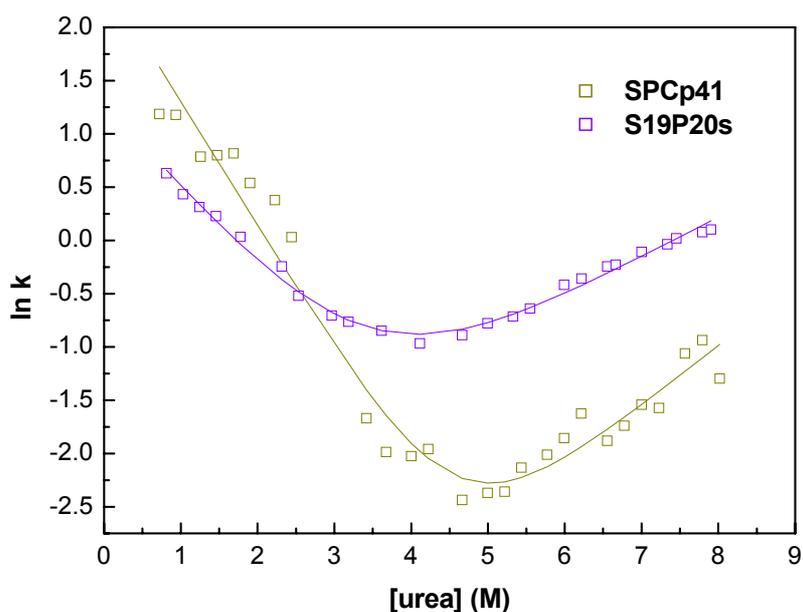


Figura 5.2.1. Gráficos de Chevron para la fase lenta del plegamiento de la quimera SPCp41 y del permutante circular S19P20s, realizados a 25°C en tampón fosfato sódico 50 mM, pH 7.0. Los símbolos abiertos representan los datos obtenidos experimentalmente para la fase lenta. La línea continua, en ambos casos, corresponde al análisis realizado mediante la ecuación 6.12.10.

En sus estudios de plegamiento con ribonucleasa A, Baldwin y colaboradores (Garel y Baldwin, 1973; Garel y Baldwin, 1975; Garel *et al.*, 1976) abordaron experimentalmente la hipótesis de la isomerización de prolina propuesta inicialmente por Brandts (Brandts *et al.*, 1975), en la que la fase lenta y la rápida diferirían en la configuración *cis-trans* de un enlace peptídico X-Pro. Posteriormente, se demostró que en la ribonucleasa A se forma un estado intermedio con cierta organización estructural que presenta una prolina en el isómero incorrecto (Nall *et al.*, 1978; Schmid, F. X. y Baldwin, 1979).

En el caso de la quimera SPCp41 y del permutante circular S19P20s, las constantes de la segunda fase cinética muestran una dependencia típica de un gráfico de Chevron, como podemos observar en la Figura 5.2.1. Así, estos datos experimentales se han analizado de igual forma que los datos de la primera fase cinética, como veremos en la Sección 5.4, y los parámetros cinéticos obtenidos se recogen en la Tabla 5.2.1. El que los datos dibujen un Chevron, tanto en la quimera como el permutante, indicaría que esta segunda fase no se debe exclusivamente a una isomerización de prolina, sino que podría haber un intermedio de plegamiento, análogamente a lo descrito para la Ribonucleasa A. A partir de los valores de los parámetros cinéticos obtenidos para el permutante circular y la quimera (Tabla 5.2.1), se observan diferencias entre la segunda fase cinética del permutante y la quimera, que podrían atribuirse a diferencias estructurales entre los posibles estados de ambas proteínas. Sin embargo, los valores de estos parámetros que se derivan del análisis llevan un alto error asociado, a pesar de la buena calidad aparente del ajuste de los datos experimentales al modelo usado, con lo que no se pueden obtener conclusiones definitivas a la vista de estos resultados, aunque sí podrían analizarse desde un punto de vista cualitativo. En el permutante circular S19P20s, si hay una contribución a la segunda fase debida a isomerización de prolina, tiene que ser por la prolina que se encuentra en el extremo N-terminal (P5) o, más probablemente, por la prolina que está formando parte del sitio de unión (P39), y que forma parte del núcleo de plegamiento de los dominios SH3.

Tabla 5.2.1. Parámetros cinéticos a 25°C en fosfato sódico 50 mM, pH 7.0*. Los datos obtenidos a partir de los valores de las constantes cinéticas resultantes del ajuste a una ecuación biexponencial corresponden a la segunda fase cinética.

Proteína	k_{p2} (s ⁻¹)	k_{d2} (s ⁻¹)	m_{p2} (M ⁻¹)	m_{d2} (M ⁻¹)
SPCp41	11.85 ± 6.72	0.004 ± 0.009	-1.17 ± 0.27	0.56 ± 0.30
S19P20s	3.57 ± 0.66	0.06 ± 0.27	-0.81 ± 0.12	0.38 ± 0.06

*Los errores han sido calculados como $3\sigma_x$ según se explica en la Sección 5.7 de este Capítulo, donde σ_x es el error de x obtenido del ajuste de los datos experimentales.

5.2.1. Catálisis por prolil isomerasas

Las prolil isomerasas son enzimas que catalizan la isomerización de los enlaces X-Pro y son excelentes herramientas para identificar la isomerización de prolinas en las cinéticas de plegamiento. La aceleración de la reacción de plegamiento por una prolil isomerasa es la prueba más concluyente para poner en evidencia la influencia de la isomerización de prolinas. Sin embargo, el que no se produzca catálisis no es una prueba de que no haya isomerización, ya que en algunas proteínas puede que las prolinas no estén accesibles a la isomerasa y la actividad de la enzima también depende de las condiciones experimentales (Schmid, F. X. *et al.*, 1993; Fischer, 1994). Además, según Golbik *et al.* la eficiencia de la prolil isomerasa depende de los restos que adyacentes a las prolinas (Fischer, 1994; Golbik *et al.*, 1999).

Para comprobar si la fase cinética lenta se debe a isomerización de prolinas, determinamos el efecto de la presencia de la *ciclofilina humana hCyp18* (una peptidil-prolil isomerasa) sobre la cinética de plegamiento de la quimera SPCp41. La catálisis de la fase lenta del plegamiento por esta prolil isomerasa se ha observado en el plegamiento de varias proteínas (Golbik *et al.*, 1999). La enzima se purificó en nuestro laboratorio como se describe en la Sección 6.4.4 de Materiales y Métodos. El experimento se realizó registrando cinéticas de plegamiento en el equipo de flujo detenido mediante la dilución de una disolución de SPCp41 en tampón fosfato sódico 50 mM, pH 7.0 y en presencia de urea 6.5 M con el mismo tampón en ausencia de urea, para que la concentración final de urea en el experimento fuera de 0.61 M (dilución 1:11), y se fueron variando las concentraciones de *hCyp18* en el tampón de plegamiento (Mello *et al.*, 2005). Se obtuvieron así las cinéticas de plegamiento de la quimera SPCp41 en presencia de distintas concentraciones de *hCyp18*. De cada una de las cinéticas se obtuvieron los valores de las constantes de velocidad de las dos fases, k_{p1} y k_{p2} , y se representaron en función de la concentración final de *hCyp18* presente en las distintas cinéticas (Figura 5.2.2). Como se observa, la constante de velocidad de la fase lenta, k_{p2} , aumenta con la concentración de *hCyp18* hasta duplicar aproximadamente su valor. Por el contrario, la fase rápida del plegamiento es prácticamente insensible a la presencia de *hCyp18* en el intervalo de concentraciones estudiadas. Estos resultados están de acuerdo con los resultados obtenidos en trabajos previos (Schmid, F. X. *et al.*, 1993; Fischer, 1994; Mello *et al.*, 2005; Schmid, F. X., 2005), y demuestran que la fase lenta está relacionada con efectos asociados al fenómeno de isomerización de prolinas.

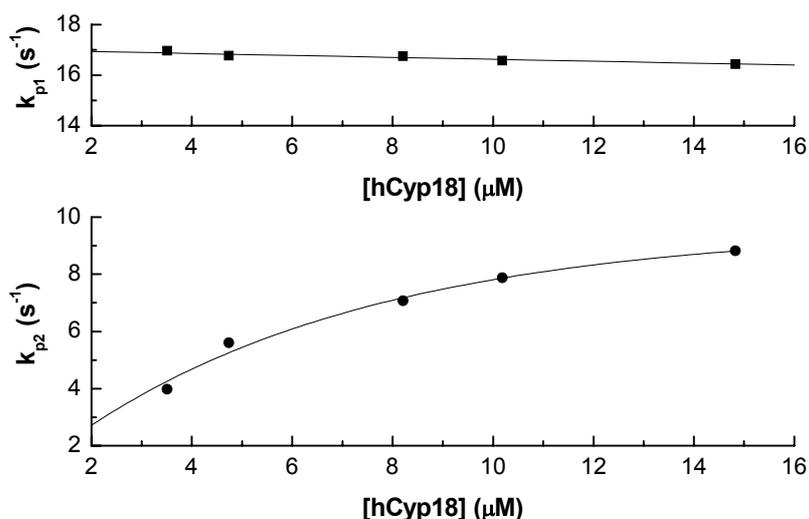


Figura 5.2.2. Efecto de la prolil isomerasa hCyp18 en las constantes de plegamiento de la quimera SPCp41. Se han representado las constantes de plegamiento obtenidas del ajuste de los perfiles cinéticos a una ecuación biexponencial (la fase rápida se ha representado con cuadrados, y la fase lenta con círculos) en función de la concentración final de ciclofilina.

5.2.2. Análisis mutacional de las prolinas del fragmento p41

La mutación de prolinas por otros restos puede ser una prueba directa para detectar y, a su vez, localizar los procesos de isomerización identificando aquellos restos de Pro más directamente implicados en el plegamiento. La velocidad a la que tiene lugar este proceso no debería cambiar con la sustitución de prolinas, si la isomerización de éstas no interviene en el plegamiento. Sin embargo, en el extremo opuesto, cuando una prolina se encuentra en conformación *cis* en el estado nativo, al ser sustituida se puede inducir la desestabilización de la proteína o, incluso, se podría impedir su plegamiento (Hurle *et al.*, 1991). No obstante, el que la fase lenta de plegamiento no varíe y permanezca cuando se sustituye una prolina, no es suficiente para descartar que se produzca un proceso de isomerización, ya que éste puede ser debido a la isomerización de más de una prolina (Kiefhaber *et al.*, 1992).

Con el fin de analizar el efecto de cada una de las prolinas del segmento p41 de la quimera SPCp41 en la segunda fase cinética observada, y comprobar si alguna de ellas es la responsable de la presencia de dicha fase, estudiamos las cinéticas de plegamiento y desplegamiento de cada uno de los mutantes de Pro-Ala que se utilizaron en el estudio termodinámico-estructural del Capítulo 4, en las mismas condiciones experimentales que las utilizadas para la quimera SPCp41, fosfato sódico 50 mM, pH 7.0.

Las cinéticas de plegamiento y desplegamiento de cada uno de los mutantes Pro-Ala se analizaron mediante ajustes tanto monoexponenciales como biexponenciales. En todos los casos, igual que en la quimera SPCp41 (Figura 5.1.1), las desviaciones de los residuos en los ajustes biexponenciales se mantienen dentro del ruido experimental mientras que los monoexponenciales presentan serias desviaciones, lo que quiere decir que las mutaciones puntuales de prolinas no evitan en ningún caso la fase cinética lenta. Esto podría deberse, como hemos comentado, a que el proceso de isomerización sea inducido por más de una prolina. Desafortunadamente, en nuestro caso no ha sido posible comprobar si esto es así, ya

que mutaciones múltiples de las prolinas inducen a la pérdida de estructura, según se describe en el Capítulo 4 de esta Memoria, en el que el estudio de CDB y el estudio de la dinámica conformacional de uno de los mutantes Pro-Ala (mutante P74A) ha puesto de manifiesto que se produce una debilitación de las interacciones en el sitio de unión de la quimera debida a mutaciones puntuales realizadas en el péptido. De hecho, en la Sección 4.7 del Capítulo 4 se ha medido la afinidad por CIT, que resulta ser pequeña, en el orden del mM. Por tanto, hemos descartado esta posibilidad, teniendo en cuenta, además, que en otros ejemplos en los que se ha intentado esta estrategia de mutagénesis múltiple los resultados tampoco han sido demasiado concluyentes (Kamen y Woody, 2002b; Bradley *et al.*, 2005).

Del análisis de las cinéticas se han obtenido de los gráficos Chevron correspondientes a las distintas proteínas quiméricas. Los valores de las constantes cinéticas de la fase rápida muestran una dependencia típica en V , lo que no se observa tan claramente para las constantes de la segunda fase, al igual que ocurre en la quimera SPCp41 (ver Figura 5.4.2).

5.3. La isomerización de prolinas como posible origen de la fase lenta

Los resultados obtenidos con la *ciclofilina humana hCyp18* nos indican que la fase cinética lenta observada en el plegamiento de SPCp41 está relacionada con procesos que implican isomería *cis-trans* de alguno de los restos de prolina, probablemente Pro39 según se desprende del análisis mutacional, que forma además parte del núcleo de plegamiento de los dominios SH3. Sin embargo, y en contra de esta afirmación, los valores obtenidos para la constante cinética de la fase lenta, k_2 (entre 0.08 y 3.5 s⁻¹), difieren apreciablemente de los medidos para otros procesos de isomería *cis-trans* en péptidos modelo (Brandts *et al.*, 1975) y en pequeñas proteínas globulares (Schmid, F. X., 2005), que suelen estar comprendidos entre 0.01 y 0.1 s⁻¹ (constantes de tiempo del orden de 10 a 100 s).

Brandts *et al.* propusieron que los isómeros con prolinas no nativas pueden bloquear el replegamiento correcto al principio (Brandts *et al.*, 1975), y que la lenta isomerización de las moléculas desplegadas es el primer paso y el limitante de la velocidad de plegamiento. El plegamiento conformacional puede empezar mientras algunas prolinas estén todavía en estados no nativos, e intermedios parcialmente plegados pueden tolerar prolinas incorrectas. Sin embargo, el paso final del plegamiento requiere prolinas en su isómero correcto y estas son las que limitan la velocidad de la isomerización de prolinas (Schmid, F. X. *et al.*, 1993). Los propios autores han medido en el citocromo C una constante de velocidad de isomerización 6 veces mayor que la que cabría esperar. Concluyen que la isomerización podría estar influenciada estéricamente por los aminoácidos adyacentes a las prolinas, los cuales podrían arrastrar la formación de alguna estructura, y proponen un factor que tiene en cuenta la naturaleza y el volumen de los restos vecinos que, en definitiva, son los que producen un aumento en la velocidad de la reacción. De hecho, no sólo los restos adyacentes parecen influir en la isomerización de las prolinas, incluso los restos más alejados parecen afectar a este proceso (Wuthrich y Grathwoh.C, 1974).

Por consiguiente, una de las razones de que las constantes cinéticas de la fase lenta sean mayores en SPCp41 y S19P20s que las que cabría esperar puede deberse, análogamente al caso del citocromo C, a la formación de cierto tipo de estructura secundaria o terciaria. En este sentido, aunque se ha demostrado que las constantes de velocidad de la isomerización de prolinas no muestran dependencia con la

concentración de agente desnaturalizante (Schmid, F.X., 1992), en nuestro caso se observa una clara dependencia de k_2 con la concentración de urea (Figuras 5.1.2 y 5.1.3). Estudios previos realizados con la proteína Barstar (Golbik *et al.*, 1999) y los anteriormente comentados con Ribonucleasa A (Nall *et al.*, 1978; Schmid, F. X. y Baldwin, 1979) demuestran que esta dependencia puede explicarse a partir de la presencia de estados intermedios de plegamiento que se organizan como consecuencia de la incorrecta isomería de algún resto de Pro. En el caso de Barstar, como en el nuestro, se observa una dependencia de k_2 con la urea, que podría dibujar la característica forma en V del gráfico Chevron.

Obviamente, nuestro análisis mutacional de las prolinas del segmento p41 es poco concluyente acerca de qué prolinas son las responsables de esa segunda fase cinética. Todos los mutantes persisten en su comportamiento bifásico, con valores para las constantes cinéticas k_2 similares a los de la quimera SPCp41. Esto indicaría que son prolinas tipo I, es decir, que afectan poco o nada al plegamiento de la quimera individualmente. Algo similar concluyen Kamen *et al.* (Kamen y Woody, 2002a), postulando que en el caso de varias prolinas que sufren isomerización con energías similares para su estado de transición, y por tanto, con velocidades similares, sus procesos cinéticos se acoplan y no pueden ser distinguidos unos de otros (Kamen y Woody, 2002b). Esto estaría de acuerdo con nuestros resultados, puesto que la hélice de poliprolina II presentaría sus restos prolina 69, 74 y 77, 76 y 73, y el 75, en condiciones energéticas similares. Todos los mutantes presentan el mismo comportamiento cinético, siendo similares sus constantes cinéticas, tanto de la fase rápida como de la fase lenta.

En nuestro caso, además, los valores casi nulos de las amplitudes nos permiten afirmar que, en las condiciones experimentales estándar (Maxwell *et al.*, 2005), este intermedio no se manifiesta en ningún momento de forma significativa. Según Brandts *et al.* (Brandts *et al.*, 1975), el valor relativo de las amplitudes y, por tanto, de las poblaciones de las distintas especies se podría modificar con las condiciones experimentales, lo que nos permitiría abordar un estudio más detallado de sus propiedades estudiando, por ejemplo, el efecto de algunas sales liotrópicas o del pH para intentar encontrar mejores condiciones experimentales que nos permitan un estudio de dicho estado en mayor profundidad.

Como ya hemos comentado, otra posibilidad sería que la fase lenta se debiera a la isomerización de Pro39, adyacente la hélice 3_{10} del dominio SH3. Puesto que esta prolina es esencial para el plegamiento de los dominios SH3, no es posible estudiar las cinéticas de mutantes con esta prolina sustituida. Por esta razón, para comprobar nuestra hipótesis hemos realizado el estudio cinético del plegamiento del permutante circular S19P20s. Las cinéticas presentan también dos fases, lo que nos sugiere que la isomerización del resto P39 podría ser responsable de la fase lenta, aunque la ausencia de esta fase cinética en el plegamiento del dominio original Spc-SH3 natural (Viguera *et al.*, 1996; Martínez *et al.*, 1999) no parece estar de acuerdo con este dato, puesto que la P39 está igualmente estructurada en ambas proteínas (Viguera *et al.*, 1996). No obstante, la diferente topología de la cadena entre S19P20s y el dominio natural Spc-SH3 podría estar favoreciendo en este último la ruta más productiva del plegamiento, reduciendo la influencia de la isomerización de prolinas y la aparición de fases cinéticas secundarias. Para comprobar esto, habría que realizar el experimento de catálisis de la *hCyp18* con el permutante circular, para ver si esta segunda fase desaparece, lo que sería una prueba evidente de que esta fase se debería a la P39.

En definitiva, los efectos observados en nuestros estudios podrían deberse a la presencia de uno o más intermedios de plegamiento, consecuencia de la presencia de prolinas en *cis*, que se encuentran muy poco poblados en nuestras condiciones de trabajo, dado el bajo valor de las amplitudes observadas. La formación de estos intermedios podría estar también favorecida por la topología particular del dominio SH3 en el permutante circular S19P20s y en la quimera SPCp41, que es diferente de la del dominio Spc-SH3 natural, para el que no se ha observado la fase lenta.

5.4. Análisis cinético de la fase rápida de plegamiento de SPCp41

La asignación de la fase rápida a la transición de plegamiento de una proteína es solamente posible en los casos donde el plegamiento es mucho más rápido que la isomerización de prolinas (Kiefhaber y Schmid, 1992; Bradley *et al.*, 2005). En estos casos, la dependencia de las constantes de velocidad de la fase rápida con la concentración de agente desnaturante se usa habitualmente para sacar conclusiones acerca del plegamiento. Sin embargo, cuando las velocidades del plegamiento y de la isomerización son similares se produce un acoplamiento de ambas fases, aunque la cinética aparentemente se pueda describir mediante un proceso de dos estados, porque las constantes de las distintas fases cinéticas se aproximan la una a la otra y porque las amplitudes de la fase rápida llegan a ser muy pequeñas (Kiefhaber *et al.*, 1992). En estos casos, la dependencia de la constante de velocidad de desplegamiento con la concentración de agente desnaturante (en una representación logarítmica) no suele ser lineal, porque el desplegamiento está dominado por la fase lenta cerca de la transición, y por la fase rápida bajo fuertes condiciones de desplegamiento (Kiefhaber y Schmid, 1992).

En nuestro caso, por tanto, podemos concluir que la fase rápida es representativa del plegamiento y del desplegamiento intrínseco de la quimera SPCp41, dada su velocidad netamente mayor que la de la fase lenta (Tabla 5.4.1) y es esta fase la que se ha utilizado en el análisis de los gráficos de Chevron correspondientes para caracterizar el proceso de plegamiento de SPCp41, como veremos a continuación.

5.4.1. Análisis de los gráficos de Chevron

El análisis del gráfico de Chevron para la fase rápida de la quimera SPCp41 se ha realizado utilizando el modelo de dos estados, tal y como se describe en la Sección 6.12.4 de Materiales y Métodos. En la Figura 5.4.1 hemos representado el resultado del análisis del Chevron del que se han obtenido los valores de las constantes de velocidad en ausencia de urea, k_{p1} y k_{d1} , y de las pendientes, m_{p1} y m_{d1} , de las ramas del plegamiento y desplegamiento respectivamente. A partir de estos parámetros se ha calculado el valor del cambio de la energía de Gibbs de equilibrio que tiene lugar durante el desplegamiento, ΔG_{D1} , y la concentración de urea en el punto medio de la transición, $[urea]_{1/2}$, además del valor de la pendiente de equilibrio, m_{D1} . Para comparar, hemos realizado el análisis del Chevron para la fase rápida del permutante circular S19P20s (Figura 5.4.1). Los valores de todas estas magnitudes están recogidos en la Tabla 5.4.1.

Tabla 5.4.1. Parámetros cinéticos a 25°C para el plegamiento y desplegamiento de las proteínas quiméricas en fosfato sódico 50 mM, pH 7.0*. Los datos obtenidos a partir de los valores de las constantes cinéticas resultantes del ajuste a una ecuación biexponencial corresponden a la primera fase cinética.

Proteína	k_{p1} (s ⁻¹)	k_{d1} (s ⁻¹)	m_{p1} (M ⁻¹)	m_{d1} (M ⁻¹)	ΔG_{D1} (kJ·mol ⁻¹)	m_{D1} (M ⁻¹)	[urea] _{1/2} (M)
SPCp41	28.12 ± 1.50	0.20 ± 0.03	-0.93 ± 0.03	0.29 ± 0.03	12.25 ± 0.81	1.22 ± 0.06	4.01
S19P20s	11.31 ± 1.77	0.44 ± 0.12	-0.93 ± 0.12	0.24 ± 0.03	8.04 ± 1.78	1.17 ± 0.15	2.77
P69A SPCp41	20.50 ± 2.10	0.19 ± 0.06	-0.85 ± 0.06	0.30 ± 0.06	11.60 ± 1.64	1.15 ± 0.12	4.07
P73A SPCp41	24.78 ± 2.67	0.26 ± 0.09	-0.91 ± 0.06	0.25 ± 0.06	11.29 ± 1.82	1.16 ± 0.12	3.93
P74A SPCp41	33.10 ± 4.77	0.27 ± 0.15	-1.02 ± 0.09	0.25 ± 0.09	11.91 ± 2.71	1.27 ± 0.18	3.79
P75A SPCp41	28.06 ± 4.23	0.30 ± 0.12	-1.02 ± 0.09	0.25 ± 0.06	11.24 ± 2.22	1.27 ± 0.15	3.57
P76A SPCp41	26.42 ± 2.91	0.33 ± 0.09	-0.98 ± 0.06	0.25 ± 0.03	10.86 ± 1.55	1.23 ± 0.09	3.56
P77A SPCp41	27.77 ± 4.05	0.32 ± 0.12	-0.97 ± 0.09	0.26 ± 0.06	11.06 ± 2.11	1.23 ± 0.15	3.63

*Los errores han sido calculados como $3\sigma_x$ según se explica en la Sección 5.7 de este Capítulo, donde σ_x es el error de x obtenido del ajuste de los datos experimentales

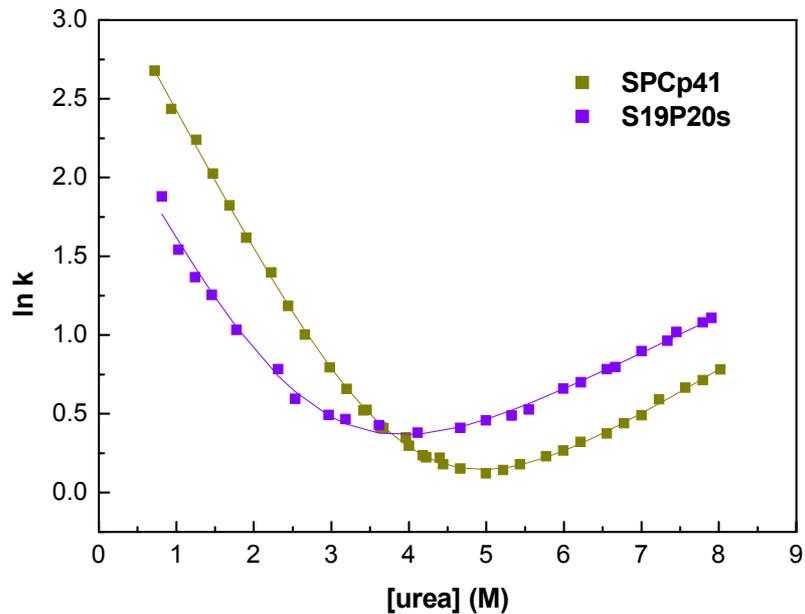


Figura 5.4.1. Gráficos de Chevron para la fase rápida del plegamiento de la quimera SPCp41 y del permutante circular S19P20s, realizados a 25°C en tampón fosfato sódico 50 mM, pH 7.0. Los símbolos cerrados representan los datos obtenidos experimentalmente para la fase rápida. La línea continua, en ambos casos, corresponde al análisis realizado mediante la ecuación 6.12.10.

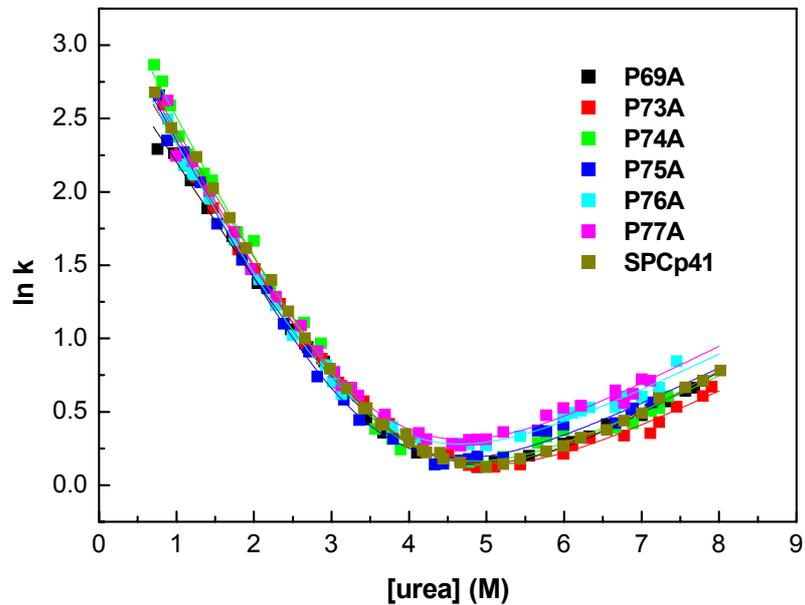


Figura 5.4.2. Gráficos de Chevron de la quimera SPCp41 y de sus mutantes Pro-Ala realizados a 25°C en tampón fosfato sódico 50 mM, pH 7. Los símbolos representan los datos obtenidos experimentalmente de la constante cinética de la fase rápida, analizados según una ecuación biexponencial. Las líneas continuas corresponden a los análisis realizados mediante la ecuación 6.12.10.

De la misma manera que para la quimera SPCp41, se han analizado los gráficos de Chevron para sus mutantes Pro-Ala (ver Figura 5.4.2) Los parámetros que resultan de dicho análisis están recogidos también en la Tabla 5.4.1, en la que se observa que los valores de las constantes de plegamiento y de desplegamiento de la quimera SPCp41 y de sus mutantes Pro-Ala muestran diferencias pequeñas que generalmente se encuentran dentro de los límites de error, lo que implica que las prolinas del segmento p41 afectan relativamente poco a la cinética de plegamiento y desplegamiento de SPCp41, tal y como se predice en otros estudios realizados (Prajapati *et al.*, 2007).

El valor de la pendiente de equilibrio, m_{D1} , al igual que el cambio en la capacidad calorífica de desplegamiento, se correlaciona con la cantidad de superficie expuesta al disolvente en el desplegamiento en el caso de proteínas que se despliegan siguiendo un modelo de dos estados (Myers *et al.*, 1995). A la vista de la similitud entre los valores de m_{D1} obtenidos para la quimera, sus mutantes de Pro-Ala y para el permutante circular, se deduce que en todos los casos el cambio del área accesible al solvente que tiene lugar durante el desplegamiento es muy parecida.

Las constantes cinéticas de plegamiento pueden obtenerse de forma empírica a partir del orden de contacto (OC) (Plaxco *et al.*, 1998b). Este parámetro refleja la importancia de los contactos entre residuos próximos y a larga distancia que tienen lugar en la estructura nativa de las proteínas. Se define como la distancia promedio a lo largo de la secuencia entre todos los pares de residuos que contactan entre sí, normalizada por el tamaño de la proteína, según la siguiente ecuación:

$$(\text{OC}) = 100 \cdot \frac{1}{L \cdot n} \sum_n \Delta Z_{i,j} \quad (5.4.1)$$

donde n es el número total de contactos en el estado nativo, $\Delta Z_{i,j}$ el número de residuos que separan los contactos i y j , y L es el número total de residuos en la proteína. Un valor del OC pequeño indica que, en promedio, las interacciones ocurren principalmente entre residuos próximos en la secuencia. Por el contrario, valores altos del OC implican que existe un mayor número de interacciones entre residuos que están separados en la secuencia. Existe una correlación empírica lineal entre el logaritmo de la constante de plegamiento, k_p , de un gran número de proteínas y su OC, en la que la primera disminuye al aumentar el OC (Plaxco *et al.*, 1998b; Fersht, 2000), y a partir de la cual es posible determinar k_p , calculando el valor del OC como se describe más adelante. Esta correlación se justifica argumentando que las proteínas con un mayor número de contactos entre residuos próximos en la secuencia (valor de OC pequeño) tienden a plegar más rápidamente que aquellas con un número mayor de contactos no locales (valores altos del OC). Es decir, que las diferencias en la topología de la estructura parecen explicar las diferencias entre las constantes de velocidad de plegamiento de diferentes proteínas.

El valor del OC se calcula haciendo uso de la estructura tridimensional de la proteína mediante su código PDB en la página web www.copredictor.ca. En nuestro caso, a partir de la estructura de la quimera SPCp41 determinada mediante RMN (Capítulo 3 de esta Memoria), cuyo código PDB es 2JMC, hemos obtenido un valor de OC de 13.62%. Este valor indica que los pares de residuos entre los que existen contactos en la estructura nativa se encuentran separados en promedio un 13.62% de la longitud total de la secuencia. A partir de este valor y de la correlación empírica entre $\ln k_p$ y el OC descrito por la ecuación: $\ln k_p = 16.1 - 0.71 \cdot (\% \text{OC})$

(www.depts.washington.edu/bakerpg/contact_order/), se obtiene un valor para $\ln k_p = 3.54$. La constante cinética de plegamiento obtenida, k_p , es 34.6 s^{-1} , valor que coincide aproximadamente con el obtenido experimentalmente en nuestro estudio (28.12 s^{-1}). La similitud entre las constantes cinéticas obtenidas empírica y experimentalmente implica que la unidad cooperativa está bastante bien definida en el diseño de esta proteína quimérica, puesto que las mayores discrepancias de la correlación la presentan proteínas que presentan regiones desordenadas en sus estructuras (Plaxco *et al.*, 1998b).

5.4.2. Análisis de las amplitudes de las cinéticas de plegamiento de la quimera SPCp41

A partir de los datos obtenidos del análisis del Chevron del plegamiento de la quimera SPCp41, se realizó el correspondiente análisis de los valores de las amplitudes de plegamiento y desplegamiento de la fase rápida, representados en la Figura 5.4.3. Para ello, se ha considerado el modelo de equilibrio de dos estados:



según el cual, la constante de equilibrio de desplegamiento, K , viene dada por:

$$K = \frac{[D]}{[N]} \quad (5.4.2)$$

y la fracción de proteína que se encuentra en el estado D, f_D es:

$$f_D = \frac{[D]}{[N] + [D]} = \frac{K}{1 + K} \quad (5.4.3)$$

Para analizar las amplitudes utilizando este modelo es necesario definir, en primer lugar, las líneas base de los estados nativo y desnaturalizado. Estas líneas base tienen en cuenta las dependencias de los cambios en la fluorescencia de ambos estados con la concentración de desnaturalizante, [urea], y se pueden describir mediante funciones lineales, A_N y A_D , para los estados nativo y desnaturalizado respectivamente, según las ecuaciones siguientes:

$$A_N = A_{N,0} + b_N \cdot [\text{urea}] \quad (5.4.4)$$

$$A_D = A_{D,0} + b_D \cdot [\text{urea}] \quad (5.4.5)$$

donde $A_{N,0}$ y b_N , por una parte, y $A_{D,0}$ y b_D por otra, representan las respectivas ordenadas en el origen y pendientes. La amplitud en función de la concentración de agente desnaturalizante viene dada por la ecuación:

$$A = A_N \cdot f_N + A_D \cdot f_D \quad (5.4.6)$$

en la que las fracciones f_N y f_D están determinadas por el valor de la constante de equilibrio entre los estados N y D, según la ecuación 5.4.2. Si expresamos f_D en función de K , podemos escribir la ecuación 5.4.6 en función de dicha constante:

$$A = \frac{(A_N + A_D \cdot K)}{1 + K} \quad (5.4.7)$$

A su vez, K guarda una relación con el cambio de energía de Gibbs de desplegamiento, ΔG , descrita por la siguiente ecuación :

$$K = \exp(-\Delta G / R \cdot T) \quad (5.4.8)$$

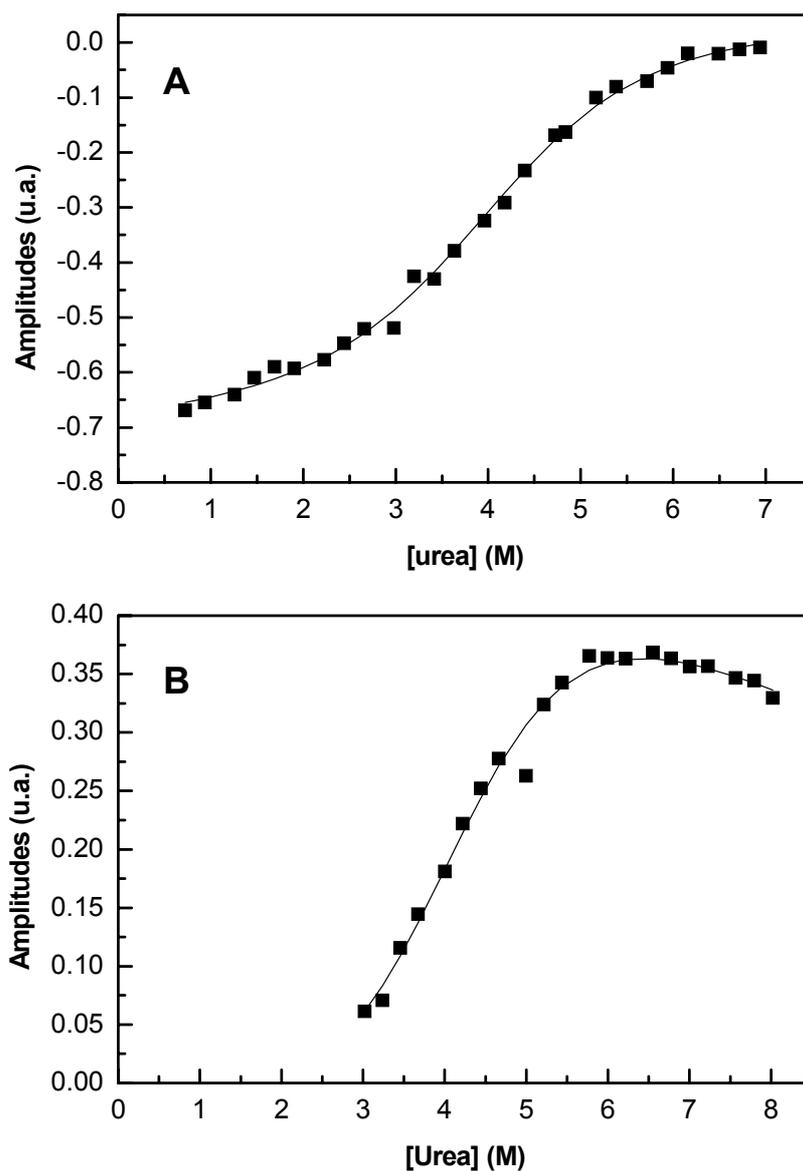


Figura 5.4.3. **A.** Amplitudes de la cinética de plegamiento de la quimera SPCp41 en función de la concentración de urea. **B.** Amplitudes de la cinética de desplegamiento de la quimera SPCp41 en función de la concentración de urea. En ambos casos, los símbolos representan los datos experimentales, y la línea el ajuste realizado asumiendo el modelo de equilibrio de dos estados.

La dependencia lineal del cambio de energía de Gibbs con la concentración de agente desnaturizante se representa con la siguiente ecuación:

$$\Delta G = \Delta G_D - m_D \cdot [\text{urea}] \quad (5.4.9)$$

donde $[\text{urea}]$ es la concentración molar de agente desnaturizante, m_D , es la pendiente de la dependencia y ΔG_D representa el cambio de energía de Gibbs en ausencia de desnaturizante. El punto medio de la curva de equilibrio representa aquella concentración de desnaturizante, $[\text{urea}]_{1/2}$, a la que el 50% de la población de la proteína total presente en la muestra se encuentra en estado nativo y el 50% restante en estado desnaturizado. En este punto se cumple que el valor de ΔG es cero. A partir de la ecuación anterior, $[\text{urea}]_{1/2}$ se define como:

$$[\text{urea}]_{1/2} = \frac{\Delta G_D}{m_D} \quad (5.4.10)$$

De esta manera, podremos expresar el cambio total de energía de Gibbs del desplegamiento en función de $[\text{urea}]_{1/2}$ despejando el valor de m y sustituyendo en la ecuación 5.4.9 para obtener la siguiente ecuación:

$$\Delta G = \Delta G_D \cdot (1 - [\text{urea}]/[\text{urea}]_{1/2}) \quad (5.4.11)$$

En la Figura 5.4.3 la línea continua representa el ajuste de las amplitudes del plegamiento y desplegamiento de la quimera SPCp41 utilizando la ecuación 5.4.6. Los parámetros ajustados son $A_{N,0}$, b_N , $A_{D,0}$ y b_D , fijando los valores de $[\text{urea}]_{1/2}$ y de ΔG_D obtenidos del análisis del Chevron correspondiente. El análisis del gráfico de Chevron y el de las amplitudes se han llevado a cabo de forma individual en lugar de global, con el fin de no afectar la alta correlación observada en el Chevron, con la comparativamente elevada dispersión que se observa en los datos de amplitud.

Una vez calculadas las líneas base del estado nativo y del desplegado, y con objeto de comparar los datos de plegamiento y desplegamiento, se normalizaron los valores de las amplitudes, A_{norm} , expresándolos en función de la fracción de la población de dicho estado, f_N . Esto es posible debido a que las amplitudes de las fases observadas guardan relación con las poblaciones de las diferentes especies (Hurle y Matthews, 1987). Teniendo en cuenta que las poblaciones de los estados N y D (f_N y f_D respectivamente) cumplen la siguiente relación:

$$f_N + f_D = 1 \quad (5.4.12)$$

y que el valor de la amplitud, A , a cualquier concentración de agente desnaturizante viene dado por la ecuación 5.4.6, combinando esta ecuación con la anterior, obtenemos que el valor de A_{norm} es:

$$A_{norm} = \frac{A - A_N}{A_N - A_D} \quad (5.4.13)$$

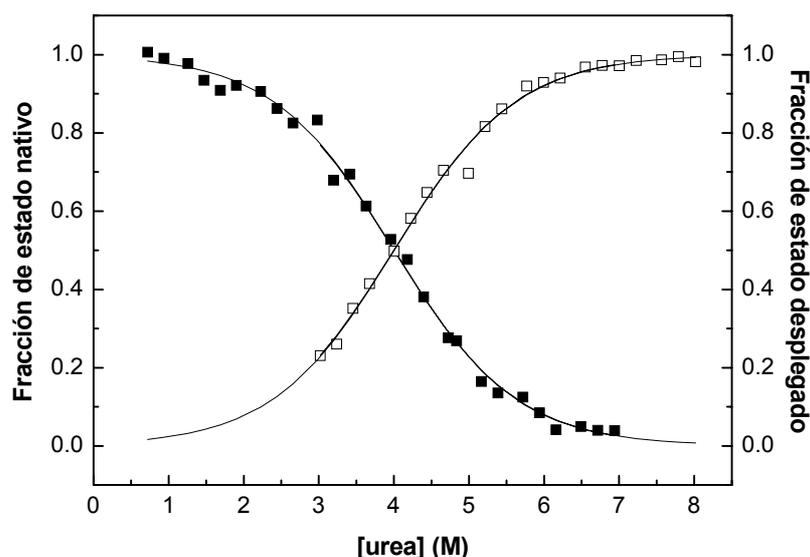


Figura 5.4.4. Fracción de estado nativo y de estado desplegado de la quimera SPCp41 en función de la concentración de urea, para las cinéticas de plegamiento (cuadrados cerrados) y desplegamiento (cuadrados abiertos). La línea continua representa el ajuste de los datos experimentales al modelo de equilibrio de dos estados.

De esta manera, podemos representar la fracción del estado nativo en función de la concentración de urea para las cinéticas de plegamiento y desplegamiento, tal como se observa en la Figura 5.4.4, en la que se representan mediante símbolos las fracciones de estado nativo y desplegado para la quimera SPCp41. La línea continua corresponde al ajuste de los datos experimentales utilizando el modelo de equilibrio de dos estados, fijando los valores de ΔG_D y $[urea]_{1/2}$.

A partir de esta representación parece ponerse de manifiesto que la fase rápida de plegamiento acontece por el “verdadero plegamiento-desplegamiento” de la quimera SPCp41. Sin embargo, como veremos a continuación, puede que no todo esté dicho en este sentido.

5.5. Fase ultra-rápida en el plegamiento de SPCp41

La formación de muchos intermedios cinéticos durante el plegamiento de las proteínas tiene lugar de forma muy rápida, dentro del tiempo muerto instrumental de los equipos de flujo detenido (2-5 ms), por lo que no es posible medirla mediante esta técnica. A pesar de esto, la presencia de dichos intermedios queda reflejada, bien por un cambio abrupto de la señal durante el tiempo muerto del aparato, definiendo lo que se conoce como fase ultra-rápida (“burst phase”) de la cinética de plegamiento, o bien por una desviación de la linealidad en una de las ramas, o en ambas, de los gráficos de Chevron. En algunas proteínas, como la ribonucleasa A, la fase ultra-rápida se ha observado cuando la proteína pasa de forma inmediata desde altas concentraciones de desnaturante a bajas concentraciones y, por tanto, el estado desnaturado se compacta (lo que no es un verdadero estado intermedio con una estructura termodinámicamente estable). En otras proteínas, sin embargo, se han encontrado ciertas evidencias a favor de una estructura estable que se forma cooperativamente durante las etapas tempranas del plegamiento. Sin embargo, la pérdida de la linealidad en el gráfico de Chevron y la presencia de fases ultra-rápidas no son suficientes evidencias para afirmar la presencia de intermedios estables de

plegamiento (Krantz y Sosnick, 2000; Takei *et al.*, 2000). El límite inferior de tiempo necesario para que se formen elementos de estructura secundaria depende del tipo de elementos concretos: algunas hélices α se forman en tiempos del orden de nanosegundos, mientras que otras hélices pueden, por el contrario, tardar varios microsegundos; los giros β tardan en plegarse alrededor de 10 μ s, mientras que los lazos alrededor de 1 μ s (Mateu, 2003).

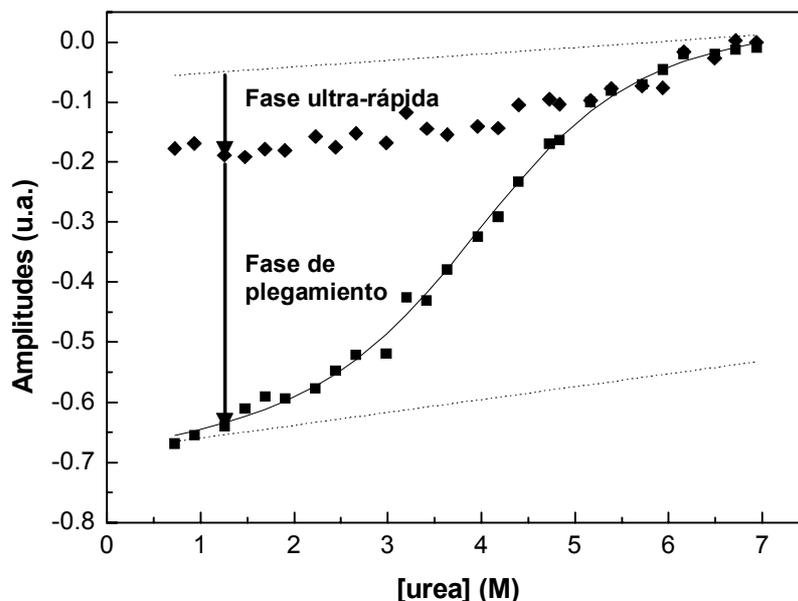


Figura 5.5.1. Amplitudes de la cinética de la fase rápida de la quimera SPCp41 en función de la concentración de urea representadas con cuadrados. La línea continua representa el ajuste de los datos experimentales al modelo de equilibrio de dos estados. Las líneas punteadas representan las líneas base de los estados nativo y desplegado obtenidas tal y como se describe en la Sección 5.4.2. Los diamantes representan la suma de las amplitudes de la fase rápida, de la fase lenta y del punto final de la fluorescencia a tiempo infinito. Como podemos ver, esta suma no coincide con la línea base del estado nativo, lo que indica que hay una fase ultra-rápida, que ocurre en el tiempo muerto del instrumento.

Para comparar las amplitudes de las cinéticas de plegamiento obtenidas experimentalmente con la amplitud teórica correspondiente al plegamiento de la proteína, hemos sumado la amplitud de la fase rápida, A_{p1} , la amplitud de la fase lenta, A_{p2} , y el valor final de la fluorescencia, I_{∞} , obtenidos de los ajustes biexponenciales para cada concentración de urea (Sauder *et al.*, 1996) (Mann y Matthews, 1993). La posición final de la señal de fluorescencia es proporcional a la fracción de estado nativo al final de cada cinética de plegamiento (Hurle y Matthews, 1987), por lo que si se le resta el valor final de la fluorescencia obtenido a altas concentraciones de agente desnaturante, donde la población del estado nativo formada es prácticamente cero, tendremos la diferencia entre la amplitud experimental y la teórica para cada cinética de plegamiento. Los valores así calculados están representados en la Figura 5.5.1 junto con el ajuste de las amplitudes utilizando el modelo de equilibrio de dos estados (Sección 5.4.2), y con las líneas base del estado desplegado y del estado nativo obtenidas del ajuste descrito en la Sección anterior. Si las amplitudes experimentales y las teóricas coincidieran los valores calculados deberían coincidir con la línea base del estado desplegado, ya que la diferencia entre las líneas base representa la diferencia total en la señal de fluorescencia que debería de observarse entre los estados N y D (amplitud teórica). En la Figura podemos observar que los datos no coinciden con la posición de la línea base del estado desplegado, lo que nos está indicando la

presencia de una fase ultra-rápida en el plegamiento que no se puede detectar, ya que ocurre en el tiempo muerto del aparato de medida que en nuestro caso es de 3 ms aproximadamente.

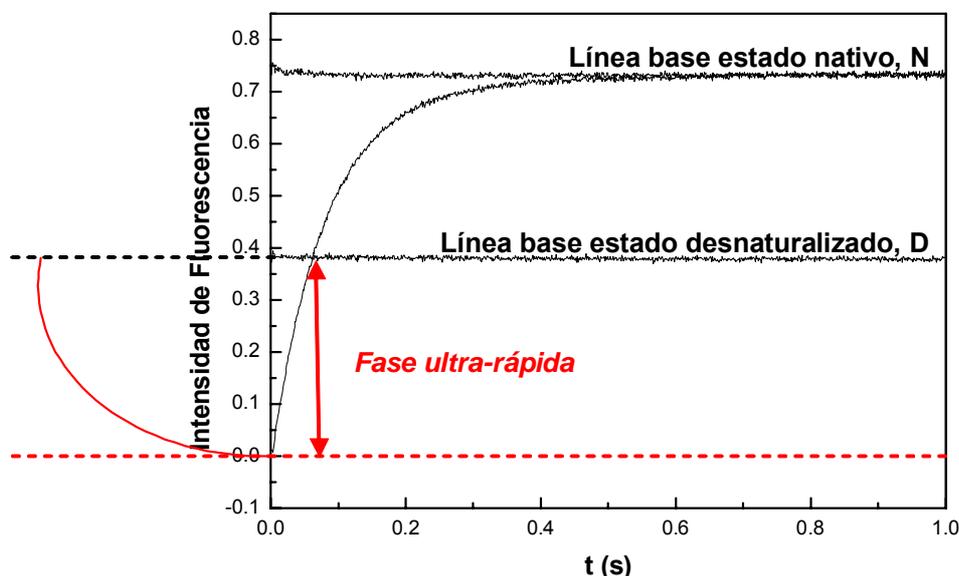


Figura 5.5.2. Fase ultra-rápida detectada en el experimento de plegamiento de la quimera SPCp41 mediante la técnica de flujo detenido. La flecha indica el cambio de amplitud que ocurre durante el tiempo muerto del instrumento, que en nuestro caso son 3 ms. Los valores de la fluorescencia de las líneas base del estado nativo y desnaturalizado aparecen indicadas. La línea discontinua roja inferior representa el valor de la fluorescencia del estado desnaturalizado obtenida realmente, mientras que la línea continua roja indica, sin escala temporal, la posible fase ultra-rápida en la cinética de plegamiento de la quimera SPCp41.

Tabla 5.5.1. Disoluciones necesarias para realizar los experimentos que se utilizan para detectar la existencia de una fase ultrarrápida en la cinética de plegamiento.

Disolución	[Proteína] (μM)	Tampón
A	50	Fosfato sódico 50 mM, pH 7.0 y urea 9 M
B	0	Fosfato sódico 50 mM, pH 7.0 y urea 9 M
C	0	Fosfato sódico 50 mM, pH 7.0
D	50	Fosfato sódico 50 mM, pH 7.0

Para corroborar la presencia de fases ultra-rápidas del plegamiento que transcurren dentro del tiempo muerto del instrumento de medida, es preciso comparar los valores de la fluorescencia inicial de la cinética con los de la fluorescencia de los estados desplegado y nativo a la misma concentración de urea (Mello *et al.*, 2005).

Con el fin de verificar la existencia de dicha fase ultra-rápida se realizaron una serie de experimentos de flujo detenido con la quimera SPCp41 (Figura 5.5.2). Para la realización de estos experimentos se prepararon 4 disoluciones: A, B, C y D, cuya composición se describe en la Tabla 5.5.1, que se dejaron incubar durante toda la noche a temperatura ambiente. Los experimentos consistieron en mezclar las disoluciones A y B para obtener la línea base de fluorescencia del estado desnaturalizado, y D y C para obtener la línea base de fluorescencia del estado nativo. También se obtuvo una cinética de plegamiento de la proteína, mezclando las

disoluciones A y C. A la vista de los resultados obtenidos, se observa la presencia de una fase ultra-rápida de plegamiento que tiene lugar dentro del tiempo muerto del instrumento, y cuyo cambio de fluorescencia presenta signo opuesto al de las cinéticas de plegamiento. En la Figura 5.5.2 se representa, obviamente sin escala temporal, esta posible fase ultra-rápida para la cinética de plegamiento de la quimera SPCp41.

Existen numerosas referencias en la bibliografía en las que se especula acerca del posible origen de estas fases ultra-rápidas que, como se ha mencionado previamente no pueden ser medidas haciendo uso de la técnica de flujo detenido y son muy pocos ejemplos en los que estas fases no aparecen, lo que parece indicar que representan una característica muy común en el proceso de plegamiento de pequeñas proteínas globulares (Jacob *et al.*, 2004). Por otra parte se ha comprobado que, en general, parece ser un proceso cooperativo, detectándose un cierto grado de estructuración en las especies que aparecen en estas fases (Bachmann, A. *et al.*, 2002). Se ha especulado sobre la posibilidad de que estas especies sean el resultado de un colapso hidrofóbico más o menos inespecífico (Wu, Y. *et al.*, 2005), o bien estados de glóbulo fundido formados por estructuras secundarias (Kuwajima *et al.*, 1991; Shastry y Roder, 1998). Otros autores postulan que estas fases no implican la presencia de especies intermedias en fases tempranas del plegamiento (Krantz *et al.*, 2002).

Los estudios más recientes llevados a cabo con pequeñas proteínas globulares describen este proceso cinético como un proceso complejo, al menos bifásico, que resulta como consecuencia de un colapso hidrofóbico que podría ocurrir durante el tiempo de mezcla (unos pocos μs) seguido de la formación de algunos contactos específicos entre determinados restos de la cadena polipeptídica (Magg *et al.*, 2006; Lapidus *et al.*, 2007; Sinha y Udgaonkar, 2007). Esto estaría de acuerdo con el carácter cooperativo observado para estos procesos (formación de estructura) y con las escalas de tiempo medidas experimentalmente para el colapso hidrofóbico ($5\mu\text{s}$) (Eaton *et al.*, 1997; Krieger *et al.*, 2003). Además, Roder *et al.* (Roder *et al.*, 2006) han observado cierta correlación de los valores cinéticos de estas fases con la topología estructural (orden de contacto).

En el caso de la quimera SPCp41, como se observa en la Figura 5.5.2, su fase ultra-rápida presenta una amplitud de signo opuesto a las propias del plegamiento, lo que merece ser analizado con detenimiento. Aunque ésta no sea la situación esperada, no es el único ejemplo descrito, ya que también sucede algo similar con la α -Lactoglobulina y la CheY (Roder y Colon, 1997), así como con la nucleasa de estafilococo (Maki *et al.*, 2007), la sintasa de triptófano (Chaffotte *et al.*, 1992), la ribonucleasa A (Welker *et al.*, 2004), la proteínas de tipo barril TIM (Wu, Y. *et al.*, 2005) o la Im7 (Capaldi *et al.*, 2002). En este último caso se pudo analizar dicha fase mediante la medida de cinéticas ultra-rápidas usando la técnica de flujo continuo (cuyo tiempo muerto está dentro del orden del μs) concluyéndose que la presencia de esa fase cinética no puede justificarse mediante la existencia de intermedios, I, no productivos ($I \rightleftharpoons D \rightleftharpoons N$), sino productivos ($D \rightleftharpoons I \rightleftharpoons N$). Esto indicaría que, a pesar de que estructuralmente la especie I parece organizarse a través de la formación de contactos no-nativos, estos intermedios pueden ser cinéticamente productivos.

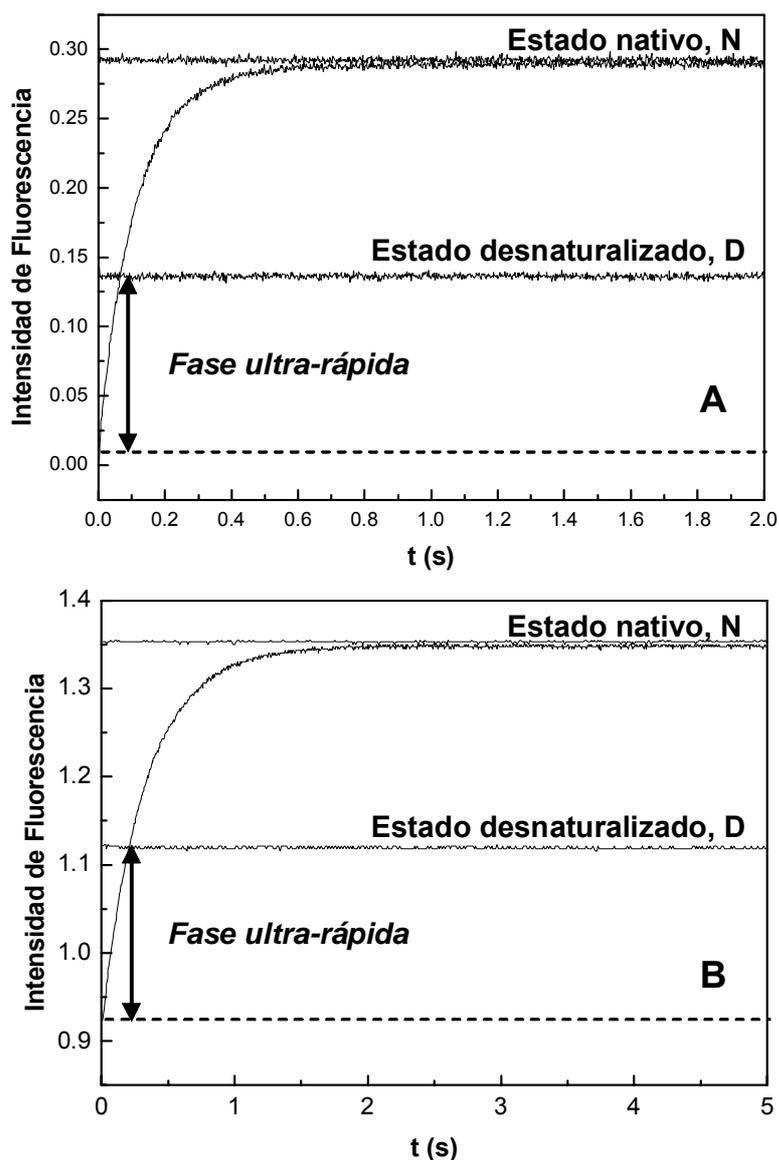


Figura 5.5.3. A. Fase ultrarrápida detectada en el experimento de plegamiento mediante la técnica de flujo detenido del permutante circular S19P20s. B. Fase ultrarrápida detectada en el experimento de plegamiento mediante la técnica de flujo detenido del dominio SH3 natural de α -espectrina. En ambos casos, la flecha indica el cambio de amplitud que ocurre durante el tiempo muerto del instrumento, que en nuestro caso son 3 ms. Los valores de la fluorescencia del estado nativo y desnaturalizado aparecen indicados.

Para intentar conocer un poco mejor los posibles orígenes de la fase ultrarrápida en el plegamiento de la quimera SPCp41, hemos comprobado si ésta tiene lugar en las cinéticas de plegamiento del permutante circular S19P20s (ver Figura 5.5.3.A), lo que indicaría en tal caso que dicha fase no resulta como consecuencia de los nuevos elementos estructurales incluidos en el diseño de la quimera, es decir, la secuencia de conexión entre el permutante y el péptido p41, que forma un giro, y el propio el péptido p41, que se organiza en hélice de poliprolina II. Esto podría ser plausible dado que las hélices de poliprolina II son una característica común en algunos estados desnaturalizados (Creamer, 1998).

A la vista de nuestros resultados, la fase ultrarrápida tampoco parece tener su origen en la topología del propio permutante circular S19P20s, puesto que el dominio SH3 natural de α -espectrina, a partir del cual se creó dicha permutación circular, presenta a su vez un comportamiento similar al observado con SPCp41, como podemos observar en la Figura 5.5.3.B, aunque en este caso la evolución cinética del plegamiento sea monofásica (Viguera *et al.*, 1994b; Martínez *et al.*, 1999). Otra posible explicación podría deberse a que esta fase ultra-rápida surgiera a partir de la rápida organización estructural local de algún segmento de la cadena de la proteína. Existe un análisis estructural por RMN (Viguera *et al.*, 1995) en el que no se observó estructuración alguna para varios fragmentos peptídicos obtenidos a lo largo de todo el dominio Spc-SH3, aunque quizás la selección un tanto arbitraria de los fragmentos podría haber falseado los resultados. Así, el análisis de un fragmento correspondiente a la horquilla β distal, considerada como parte del núcleo de plegamiento de Spc-SH3 (Martínez *et al.*, 1998), sí presentaba alguna evidencia de la existencia de un equilibrio conformacional en los restos del giro β de la horquilla, con una dinámica conformacional muy rápida, del orden de μ s, similar a la comprobada experimentalmente para otras horquillas β . Para explorar la influencia del giro distal sobre la existencia de la fase ultra-rápida, hemos analizado el comportamiento cinético de esta fase en el mutante D48G de Spc-SH3. Esta mutación favorece entrópicamente el formación de la horquilla distal y, por tanto, la nucleación del plegamiento de la proteína (Martínez *et al.*, 1998; Martínez y Serrano, 1999), luego debería afectar a la fase ultra-rápida. Al hacer el análisis de amplitudes para el mutante D48G de Spc-SH3 se obtiene una representación análoga a la obtenida para la quimera SPCp41 (Figura 5.5.1), lo que indica que la fase ultra-rápida no desaparece, sino que sigue estando presente.

Quizás la explicación más plausible podría surgir de resultados muy recientes. Así, Li *et al.* (Li *et al.*, 2007) han detectado la formación de una conformación de tipo helicoidal en la fase ultra-rápida del dominio SH3 del oncogén Src. Esta organización estaría de acuerdo con los resultados de nuestro experimento realizado con SPCp41 (Figura 5.5.1) puesto que esta conformación no-nativa se formaría como consecuencia de un colapso inicial debido al cambio de disolvente (de alta a baja concentración de urea) que da origen a la reacción de plegamiento. En esta situación, la fase ultra-rápida podría representar el desplegamiento de esta estructura no-nativa, para continuar posteriormente con los procesos más lentos descritos anteriormente (plegamiento + isomerización de prolinas) y así conseguir alcanzar el estado nativo.

Aún teniendo en cuenta el carácter no nativo de su estructura, estos intermedios parecen ser productivos para el plegamiento, tal y como hemos comentado para Im7 y algún otro caso de los mencionados anteriormente, quizás porque el colapso estructural que parece ocurrir limita el espacio conformacional accesible, lo que favorecería el plegamiento global, puesto que en ese espacio conformacional restringido sí se encuentra el estado nativo.

5.6. Análisis del estado de transición de plegamiento de la quimera SPCp41

Con objeto de obtener los gráficos de Eyring y determinar los parámetros de activación de los procesos de plegamiento y desplegamiento, hemos analizado la dependencia con la temperatura de las constantes de velocidad de los procesos cinéticos. Para ello hemos registrado las cinéticas de plegamiento y desplegamiento

de la quimera SPCp41 en tampón fosfato sódico 50 mM, pH 7.0 y a 5°C, 15°C, 25°C y 35°C.

En general, a todas las temperaturas de estudio, las constantes de velocidad de los procesos de plegamiento y desplegamiento de la quimera SPCp41 se calcularon a partir del ajuste de las cinéticas utilizando una ecuación biexponencial (de la misma forma que como se ha explicado en la Sección 5.1) para cada concentración de urea. Sin embargo, a 5°C en el proceso de plegamiento desaparece una fase, mientras que en el desplegamiento siguen apareciendo las dos fases cinéticas. Sin embargo, a 35°C tiene lugar la situación opuesta, las cinéticas de plegamiento presentan dos fases, mientras que las de desplegamiento presentan una única fase cinética. A 15°C y a 25°C, tanto para el plegamiento como para el desplegamiento, el ajuste de las curvas cinéticas resultó ser biexponencial. Esto indicaría que desaparece la barrera de activación correspondiente a la fase lenta en esas condiciones de temperatura.

A partir de los valores de las constantes de velocidad obtenidas hemos construido los gráficos de Chevron para la quimera SPCp41 a cada una de las temperaturas de estudio. En el caso de las constantes cinéticas correspondientes a la primera fase (en los casos en los que se observaron dos), éstas describen bien un gráfico de Chevron (Figura 5.6.1), mientras que las constantes de la segunda fase presentan una gran dispersión en este tipo de gráfico. En la Figura 5.6.1 se representan con línea continua los ajustes de los datos experimentales de estos Chevron a la ecuación 6.12.10, dando lugar a las constantes de plegamiento y desplegamiento en ausencia de urea, k_{p1} y k_{d1} , y las pendientes, m_{p1} y m_{d1} , respectivamente, a las diferentes temperaturas. A partir de estos resultados hemos representados el gráfico de Eyring, en el que se representa $\ln(k/T)$ frente a $1/T$, donde k puede ser la constante de plegamiento (k_{p1}) o la de desplegamiento (k_{d1}), tal como se muestra en la Figura 5.6.2.

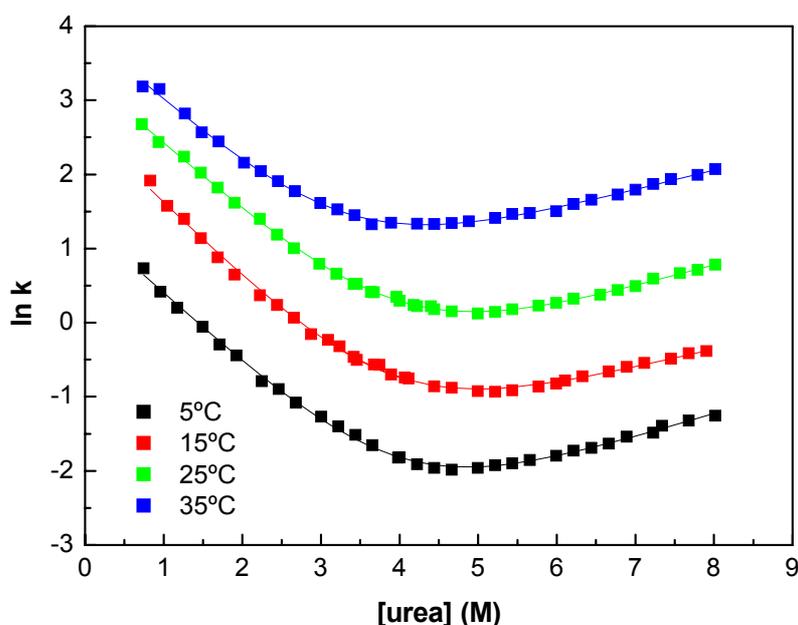


Figura 5.6.1. Gráficos de Chevron de SPCp41 realizado a 5°C, 15°C, 25°C y 35°C en tampón fosfato sódico 50 mM, pH 7. Los símbolos cerrados representan los datos obtenidos experimentalmente. La línea continua, en todos los casos, corresponde al análisis realizado mediante la ecuación 6.12.10.

Según la Teoría del Estado de Transición, cuando se produce un cambio de capacidad calorífica entre el estado nativo o el desplegado y el estado de transición, se produce una desviación de la linealidad en los gráficos de Eyring. En ese caso los gráficos de Eyring se ajustan a la siguiente ecuación:

$$\ln(k/T) = A + B(T_0/T) + C \ln(T_0/T) \quad (5.6.1)$$

siendo:

$$A = [-\Delta C_p^\ddagger + \Delta S^\ddagger(T_0)]/R - \ln(h/k_B) \quad (5.6.2)$$

$$B = [\Delta C_p^\ddagger - \Delta S^\ddagger(T_0)]/R - \Delta G^\ddagger(T_0)/RT_0 \quad (5.6.3)$$

$$C = -\Delta C_p^\ddagger/R \quad (5.6.4)$$

donde k_B y h son las constantes de Boltzmann, y de Planck, ΔS^\ddagger y ΔH^\ddagger son la entropía y entalpía de activación, respectivamente, del proceso cinético analizado, ΔC_p^\ddagger es el cambio de capacidad calorífica entre el estado nativo o el desplegado y el estado de transición, ΔG^\ddagger es la energía de Gibbs de activación del plegamiento o del desplegamiento y T_0 es la temperatura de referencia, en nuestro caso 25°C.

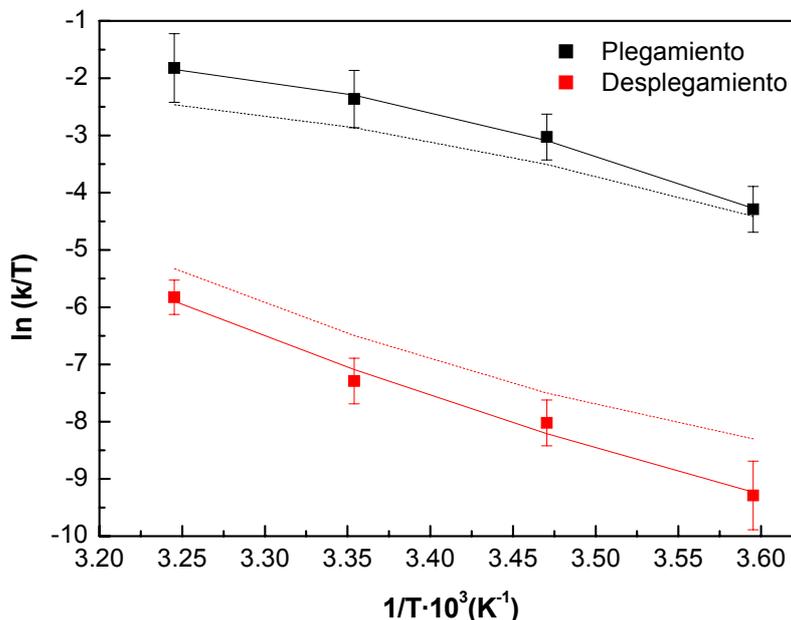


Figura 5.6.2. Gráficos de Eyring de la dependencia con la temperatura de las constantes de plegamiento y desplegamiento en agua de la quimera SPCp41 en fosfato sódico 50 mM, pH 7.0. Los símbolos representan los datos experimentales obtenidos para el plegamiento (en negro) y para el desplegamiento (en rojo). La línea continua representa el ajuste realizado con la ecuación 5.6.1. Con línea discontinua hemos representado los gráficos de Eyring del permutante circular S19P20s para el plegamiento (en negro) y para el desplegamiento (en rojo) obtenidos en trabajos previos (Martinez et al., 1999).

Así, los parámetros termodinámicos correspondientes al estado de transición, relativos al estado nativo y al desplegado, se pueden calcular de la siguiente manera:

$$\Delta C_p^\ddagger = -CR \quad (5.6.5)$$

$$-T_0\Delta S^\ddagger(T_0) = -(A - C - 23.76)RT_0 \quad (5.6.6)$$

$$\Delta G^\ddagger(T_0) = -(A + B - 23.76)RT_0 \quad (5.6.7)$$

$$\Delta H^\ddagger = -(B + C)RT_0 \quad (5.6.8)$$

Tabla 5.6.1. Parámetros de activación para el plegamiento y desplegamiento de la quimera SPCp41 calculados a partir del gráfico de Eyring de acuerdo con la ecuación 5.6.1 a 25°C.

	ΔC_p^\ddagger (kJ·K ⁻¹ ·mol ⁻¹)	ΔS^\ddagger (J·K ⁻¹ ·mol ⁻¹)	ΔH^\ddagger (kJ·mol ⁻¹)	ΔG^\ddagger (kJ·mol ⁻¹)
Plegamiento				
SPCp41 (‡-D)	-2.2 ± 1.5	-65 ± 5	45 ± 5	65 ± 6
*S19P20s (‡-D)	-1.5 ± 0.4	-94 ± 6	38 ± 2	66 ± 4
Desplegamiento				
SPCp41 (‡-N)	1.2 ± 1.5	31 ± 8	86 ± 4	76 ± 5
*S19P20s (‡-N)	1.8	19 ± 15	80 ± 4	75 ± 6

*Los valores de los parámetros de activación para el permutante circular S19P20s han sido calculados previamente (Martinez *et al.*, 1999).

En la Figura 5.6.2 observamos los gráficos de Eyring obtenidos para la quimera SPCp41, junto con los gráficos obtenidos en trabajos previos para el permutante circular S18P20s (Martinez *et al.*, 1999). En ambos casos se aprecia una ligera desviación de la linealidad para el plegamiento, que no es tan evidente para el desplegamiento. Como hemos mencionado, la desviación de la linealidad en un gráfico de Eyring tiene su origen en los cambios en la capacidad calorífica desde el estado desnaturalizado o nativo al estado de transición, y suele ser más apreciable en el plegamiento ya que el cambio de área expuesta al disolvente desde el estado nativo al estado de transición suele ser mayor, que desde el estado desnaturalizado al estado de transición. Si nos fijamos en la ecuación 5.6.1, el término logarítmico que va acompañando a C cuando ΔC_p^\ddagger es distinto de 0, es el responsable de que exista o no curvatura, y cuanto mayor sea el valor de ΔC_p^\ddagger mayor curvatura presenta el gráfico de Eyring. La pequeña curvatura en el gráfico del plegamiento para la quimera SPCp41, al igual que para el permutante circular S19P20s, indica el enterramiento de una superficie hidrofóbica en el estado de transición.

En la Tabla 5.6.1 están recogidos los valores obtenidos para el cambio de capacidad calorífica, entalpía, entropía y energía de Gibbs de activación de los procesos de plegamiento y de desplegamiento de las dos proteínas. Los parámetros de activación del plegamiento, dentro del elevado error de determinación de estos parámetros en ambos casos, son parecidos, sugiriendo que el núcleo de plegamiento de SPCp41 y de S19P20s pueden ser bastante similares. Ocurre igual en el caso del desplegamiento, lo que podría indicar que el segmento adicional que incluye el deca péptido p41 en la quimera no parece estar contribuyendo al proceso de

despliegamiento de forma cooperativa. Este hecho está de acuerdo con los resultados obtenidos en el Capítulo 4, que pusieron de manifiesto la existencia de un equilibrio conformacional acusado que afecta las interacciones en el sitio de unión en el estado nativo.

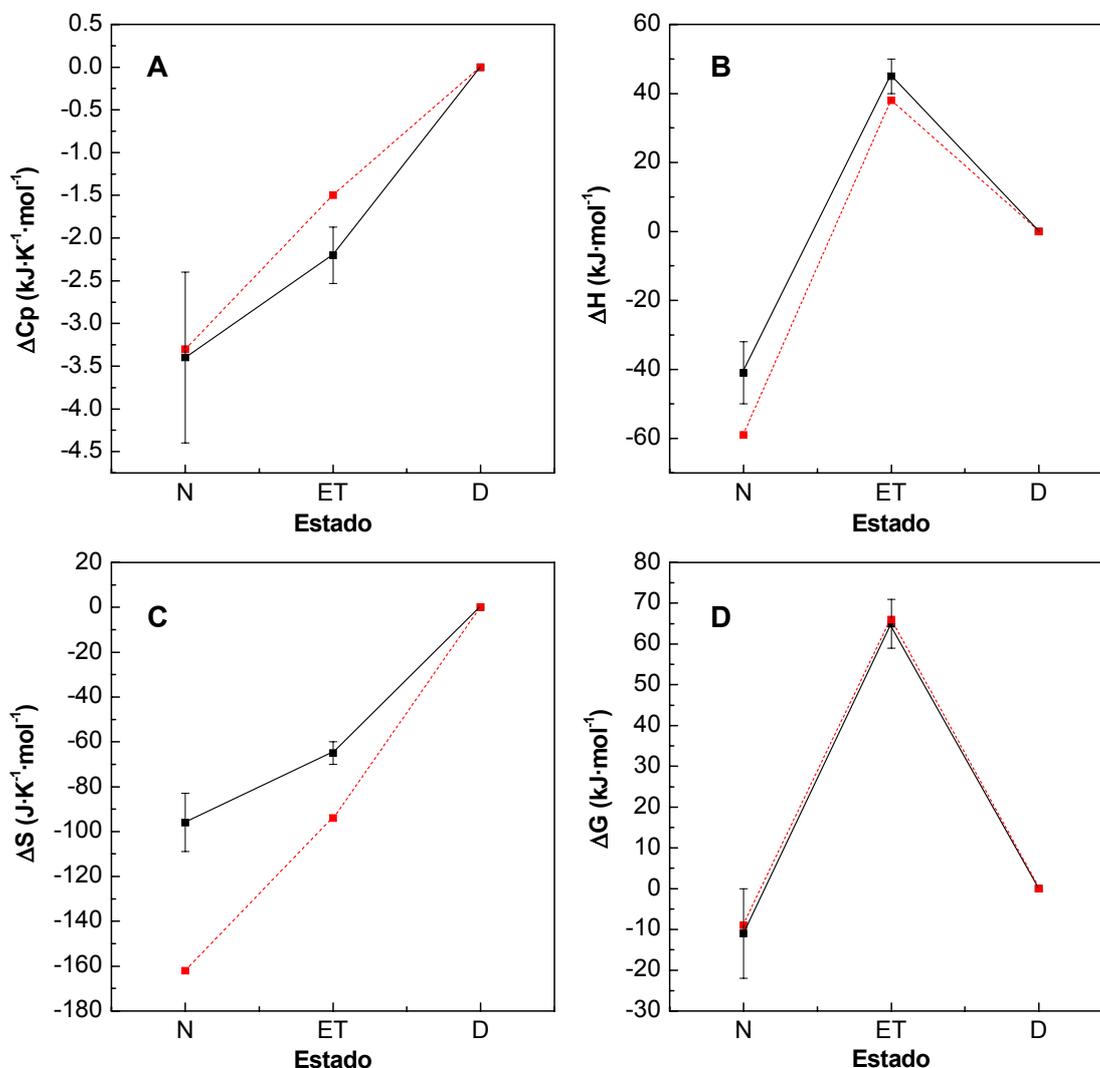


Figura 5.6.3. Perfiles de reacción de la quimera SPCp41 bajo condiciones nativas en fosfato sódico 50 mM, pH 7.0. La reacción va desde el estado nativo (N), a través del estado de transición (ET), al estado desnaturalizado (D), que lo hemos tomado como punto de referencia. **A.** Cambio de capacidad calorífica entre los tres estados; **B.** Cambio de entalpía entre los tres estados; **C.** Cambio de entropía entre los tres estados; **D.** Cambio de energía de Gibbs entre los tres estados. Hemos comparado los datos obtenidos para la quimera SPCp41 (en negro) con los obtenidos en trabajos previos para el permutante circular S19P20s (en rojo) (Martínez et al., 1999).

Basándonos en toda esta información termodinámica hemos construido los perfiles de reacción de la quimera SPCp41 y de S19P20s, que se representan en la Figura 5.6.3. Todos ellos son similares, lo que habla a favor de una similitud energética para ambos estados de transición de plegamiento, con la excepción del componente entrópico, quizás debido a la mayor longitud de la cadena en SPCp41 que en S19P20s. Las entalpías de activación obtenidas tanto para la quimera como para el permutante indican que el estado de transición se alcanza desde el estado nativo a

través de la rotura de numerosas interacciones que no se compensan con la ganancia de entalpía por solvatación.

A partir de los parámetros obtenidos del análisis de los gráficos de Chevron de los mutantes Pro-Ala del segmento p41, en comparación con la quimera SPCp41 original, se puede aplicar el método de la ingeniería de proteínas propuesto por Fersht *et al.* (Fersht, 1995) (Sección 6.12.5 de Materiales y Métodos), para obtener los valores del parámetro $\phi_{\ddagger-D}$ (ecuación 6.12.11), que representan la pérdida de energía en el estado de transición inducida por una mutación puntual en la secuencia de la proteína, en relación a la que se produce en el estado nativo. Este parámetro proporciona información de si en dicho estado de transición la región que rodea la cadena lateral del residuo mutado se encuentra plegada, parcialmente organizada o desplegada. Así el análisis de los valores del parámetro $\phi_{\ddagger-D}$ para las mutaciones Pro-Ala del segmento p41 de las quimeras nos proporcionaría información adicional sobre el grado de participación del ligando p41 en la estructura del estado de transición de plegamiento.

Hemos determinado los valores de $\phi_{\ddagger-D}$ para las mutaciones puntuales realizadas en la región correspondiente al péptido p41 de la quimera SPCp41: P69A, P73A, P74A, P75A, P76A y P77A, con los que se llevó a cabo el estudio termodinámico-estructural descrito en el Capítulo 4 de esta Memoria.

Así, se realizó el análisis de los gráficos de Chevron de las fases rápidas (que como ya se ha explicado en la Sección 5.4.1 corresponden al plegamiento y desplegado de la quimera y sus mutantes Pro-Ala) a pH 7.0, tal como se describe en la Sección 5.4.1 y en la Sección 6.12.4 de Materiales y Métodos, y se han calculado los valores de $\phi_{\ddagger-D}$ como se describe en la Sección 6.12.5 de Materiales y Métodos, siendo éstos cercanos a cero para todos los mutantes. La determinación de los valores de los parámetros $\phi_{\ddagger-D}$ presenta errores de elevada magnitud. La razón de esto se encuentra en que los cambios en la energía de plegamiento producidos por las mutaciones, $\Delta\Delta G_{N-D}$, son muy pequeños, por lo que al estar en el denominador de la ecuación 6.12.11, producen grandes errores en $\phi_{\ddagger-D}$. Trabajos previos discrepan entre diferentes límites en los valores de $\Delta\Delta G_{N-D}$, a partir de los cuales se puede estimar de forma fiable el valor de $\phi_{\ddagger-D}$. Así, Sánchez *et al.* (Sanchez y Kiefhaber, 2003) consideran que cuando el valor de $\Delta\Delta G_{N-D}$ es $< 7 \text{ kJ}\cdot\text{mol}^{-1}$, la fiabilidad en el cálculo de los valores de $\phi_{\ddagger-D}$ es baja. Sin embargo, Fersht *et al.* (Fersht y Sato, 2004) determinan que se pueden obtener valores de $\phi_{\ddagger-D}$ fiables, si se realizan las mutaciones adecuadas, cuando el valor de $\Delta\Delta G_{N-D}$ está comprendido entre 2.5 y 7 $\text{kJ}\cdot\text{mol}^{-1}$, debido a que la estructura del estado de transición de una proteína puede alterarse cuando se realizan mutaciones que presentan un valor de $\Delta\Delta G_{N-D} \gg 8 \text{ kJ}\cdot\text{mol}^{-1}$. Una forma de abordar el estudio e interpretación de los valores de $\phi_{\ddagger-D}$ podría consistir en realizar diferentes mutaciones en la misma posición (Fersht y Sato, 2004) o seguir un camino alternativo consistente en cambiar las condiciones del medio para una mutación en una posición concreta, a través de cambios en el pH (Cobos, *et al.*, ; Martinez y Serrano, 1999), con el fin de buscar unas condiciones en las que la estabilidad de la proteína varíe más significativamente.

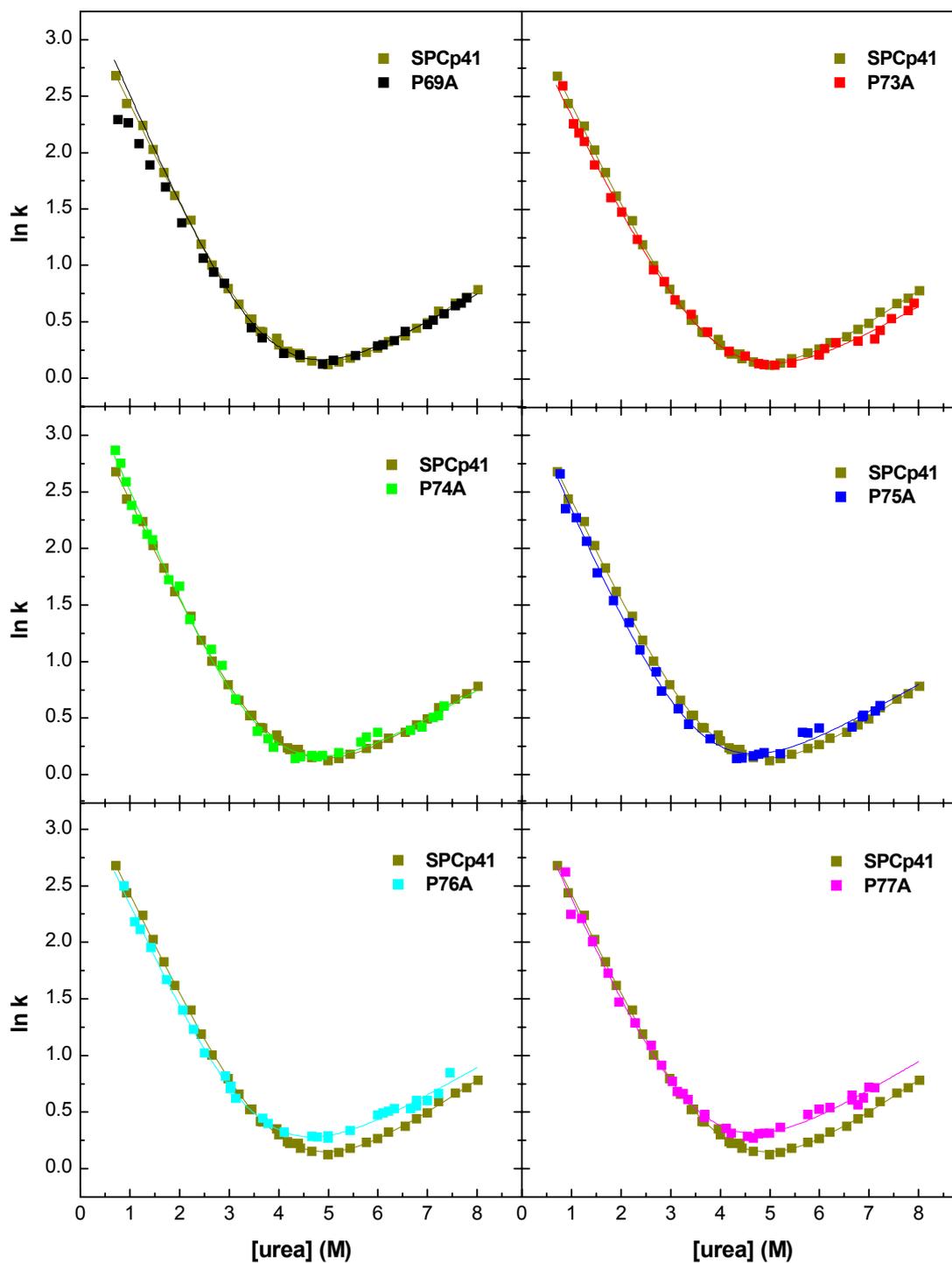


Figura 5.6.4. Gráficos de Chevron de la quimera SPCp41 (referencia) y sus mutantes Pro-Ala obtenidos de los datos experimentales realizados en tampón fosfato sódico 50 mM, pH7.0 y a 25°C. Los símbolos representan los datos obtenidos experimentalmente para la fase rápida. La línea continua, en todos los casos, corresponde al análisis realizado mediante la ecuación 6.12.10.

Cuando los valores de $\phi_{\ddagger-D}$ son iguales o muy próximos a cero se debe a que en el gráfico de Chevron las constantes de velocidad de plegamiento de la referencia y el mutante coinciden, indicando que la mutación no afecta a la cinética de plegamiento y, por tanto, que la interacción afectada por la mutación no estará involucrada en la formación del estado de transición. Es decir, que la región de la proteína que contiene la mutación se encuentra desplegada en el estado de transición. Según la Figura 5.6.4, en la que se comparan los gráficos de Chevron de cada mutante con la quimera SPCp41 (considerada como la especie de referencia para el cálculo de los valores de $\phi_{\ddagger-D}$), se observa que las ramas de plegamiento coinciden de forma general en todos los casos, observándose únicamente diferencias, no demasiado significativas, en la rama de desplegamiento.

En cualquier caso, y a falta de analizar en más profundidad el comportamiento cinético de SPCp41, los valores de $\phi_{\ddagger-D}$ cercanos a cero estarían perfectamente de acuerdo con el análisis de los parámetros de activación, puesto que indicarían la escasa o nula contribución energética del segmento p41 al estado de transición de plegamiento del dominio SH3, manteniéndose por consiguiente el núcleo de plegamiento en el dominio SH3, al que posteriormente se uniría el péptido p41 a través de un proceso de unión intramolecular, similar al descrito para el péptido libre, lo que argumenta claramente a favor de la efectividad de esta quimera de diseño para abordar estudios de unión SH3-péptido.

5.7. Análisis de los errores en los parámetros cinéticos y en los valores ϕ

A partir de la teoría de errores hemos propuesto un método general del cálculo de los errores asociado a los parámetros cinéticos, tal como se describe en las Secciones siguientes y que se ha aplicado a los datos experimentales obtenidos en esta Memoria. Este procedimiento lo hemos publicado muy recientemente en la revista *Biophysical Journal* (Cobos, *et al.*).

5.7.1. Desarrollo de una base de datos experimentales

Hemos considerado cinco conjuntos de datos experimentales cinéticos (de plegamiento y desplegamiento) del mutante D48G de Spc-SH3 obtenidos en fosfato sódico 50 mM a pH 7.0, y 25°C, que resultan de la determinación independiente de unas 1000 medidas de constantes cinéticas, con el fin de evaluar los errores experimentales y la propagación de los mismos en los parámetros cinéticos derivados de los gráficos de Chevron. Estos experimentos se realizaron con dos aparatos de flujo detenido distintos, por diferentes operadores y en distintos años, lo que permite evitar y evaluar los diferentes errores experimentales originados. Para comparar, hemos realizado un análisis paralelo de los gráficos de Chevron a pH 3.5 y 2.5, obtenidos de un único conjunto de datos experimentales. Los detalles acerca de los procedimientos experimentales para realizar estos experimentos están explicados en diferentes trabajos publicados anteriormente (Martinez *et al.*, 1998; Martinez y Serrano, 1999; Cobos *et al.*, 2003), así como en la Sección 6.12 de Materiales y Métodos.

5.7.2. Errores en las constantes cinéticas

En primer lugar se pretende obtener el error estándar de cada constante cinética obtenida del ajuste no lineal a la ecuación 5.1.1, de las dependencias experimentales de la fluorescencia frente a tiempo obtenidas a cada concentración de agente desnaturante. Un análisis de la distribución del gráfico de los residuos obtenidos de este ajuste constituye un criterio para diferenciar si la ecuación utilizada para el ajuste es la adecuada, como se explica en la Sección 5.1. Este proceso es fundamental ya que puede aparecer más de una fase cinética que no se pueda detectar en el análisis y que haga que se obtengan valores de las constantes cinéticas erróneos. Una estrategia para mejorar la relación señal/ruido experimental consiste en hacer la media de un conjunto de diferentes curvas cinéticas obtenidas en las mismas condiciones experimentales. Aunque esta media reduce el error, la exactitud en la determinación de los datos disminuye, porque algunas veces se introducen fases “artefactuales”.

Basándonos en la teoría de errores (Bevington y Robinson, 2003), podemos determinar que el error asociado a la constante cinética es $\sigma_x = \pm 3\sigma_i$ (donde σ_i es el error estándar del ajuste realizado con la ecuación 5.1.1), considerando que al menos el 90% de los valores de k en estas condiciones experimentales se pueden incluir en dicho intervalo. En la Figura 5.7.1 hemos representado los valores de las constantes cinéticas en función de la concentración de agente desnaturante para el plegamiento del del mutante D48G de Spc-SH3 a pH 7.0, incluyendo en la misma gráfica las barras de error. En la Figura 5.7.1.B ($\sigma_x = \pm 3\sigma_i$) se observa que aproximadamente el 90% de las barras de error se superponen muy bien con la línea del ajuste, lo que indica la buena calidad de los datos experimentales. A la vista de la Figura 5.7.1.C se observa que considerando un error de $\sigma_x = \pm 6\sigma_i$ la superposición de las barras de error con la línea del ajuste no mejora. Más aún, considerando un error de $\sigma_x = \pm 2\sigma_i$ (Figura 5.7.1.A), la superposición es incluso peor.

Por tanto, una estimación adecuada de los errores estándar derivados de los ajustes por mínimos cuadrados, podría dar información sobre la precisión sobre la determinación de la constante de velocidad, aunque en algunos casos otros errores pueden contribuir a disminuir la exactitud de la determinación, muchos de los cuales son errores instrumentales debidos a la termostatación, jeringas, pipetas, refractómetro, etc. y/o producen una media errónea de las curvas cinéticas. Los valores obtenidos para $\sigma_k = \pm 3\sigma_i$ se encuentran en el intervalo comprendido entre el 1% y el 10% en valores relativos (un valor medio de 4.7%, Figura 5.7.2.D). Estos resultados están de acuerdo con los resultados obtenidos previamente por De los Ríos *et al.*, si se comparan experimentos realizados en diferentes laboratorios usando diferente instrumentación, pero no cuando se llevan a cabo con diferencias temporales significativas (De Los Rios *et al.*, 2006). De los Ríos *et al.* evalúan que los errores de los ajustes ($\pm\sigma_i$) y concluyen que las barras de error más razonables son aproximadamente de un 10% (lo que equivale a $\pm 6\sigma_i$ en nuestros cálculos). Sin embargo, del análisis de nuestro conjunto de datos experimentales podemos concluir que en general es suficiente con considerar que los errores en las constantes son $\sigma_k = \pm 3\sigma_i$ (aproximadamente un 5%). Por lo tanto, del análisis de nuestro conjunto de datos experimentales podemos concluir en que, en general, el error es $\sigma_x = \pm 3\sigma_i$.

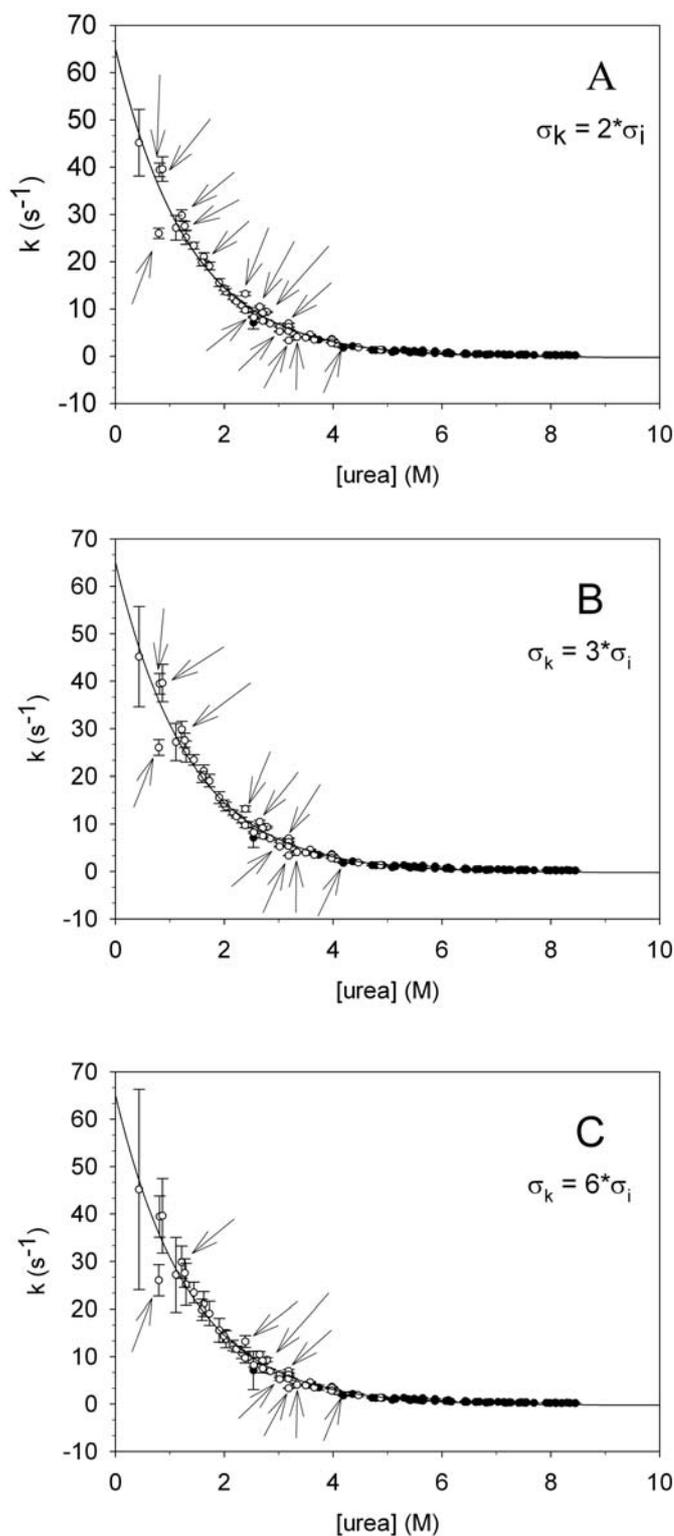


Figura 5.7.1. Constantes cinéticas obtenidas a diferentes concentraciones de urea para el mutante D48G de Spc-SH3 en fosfato sódico 50 mM (pH 7.0, 25°C). Los experimentos se han hecho en dos instrumentos distintos, por diferentes operadores y en distintos años. La línea continua representa los ajustes por mínimos cuadrados a una función exponencial. Los símbolos abiertos y cerrados representan los datos experimentales obtenidos para los experimentos de plegamiento y desplegamiento respectivamente. Las flechas indican las barras de error que no se superponen con la línea del ajuste. Las barras de error se han calculado de las estimaciones de los errores del ajuste que resultan con la ecuación 5.1.1. **A.** Dos veces los errores del ajuste ($2\sigma_i$); **B.** Tres veces los errores del ajuste ($3\sigma_i$); **C.** Seis veces los errores del ajuste ($6\sigma_i$).

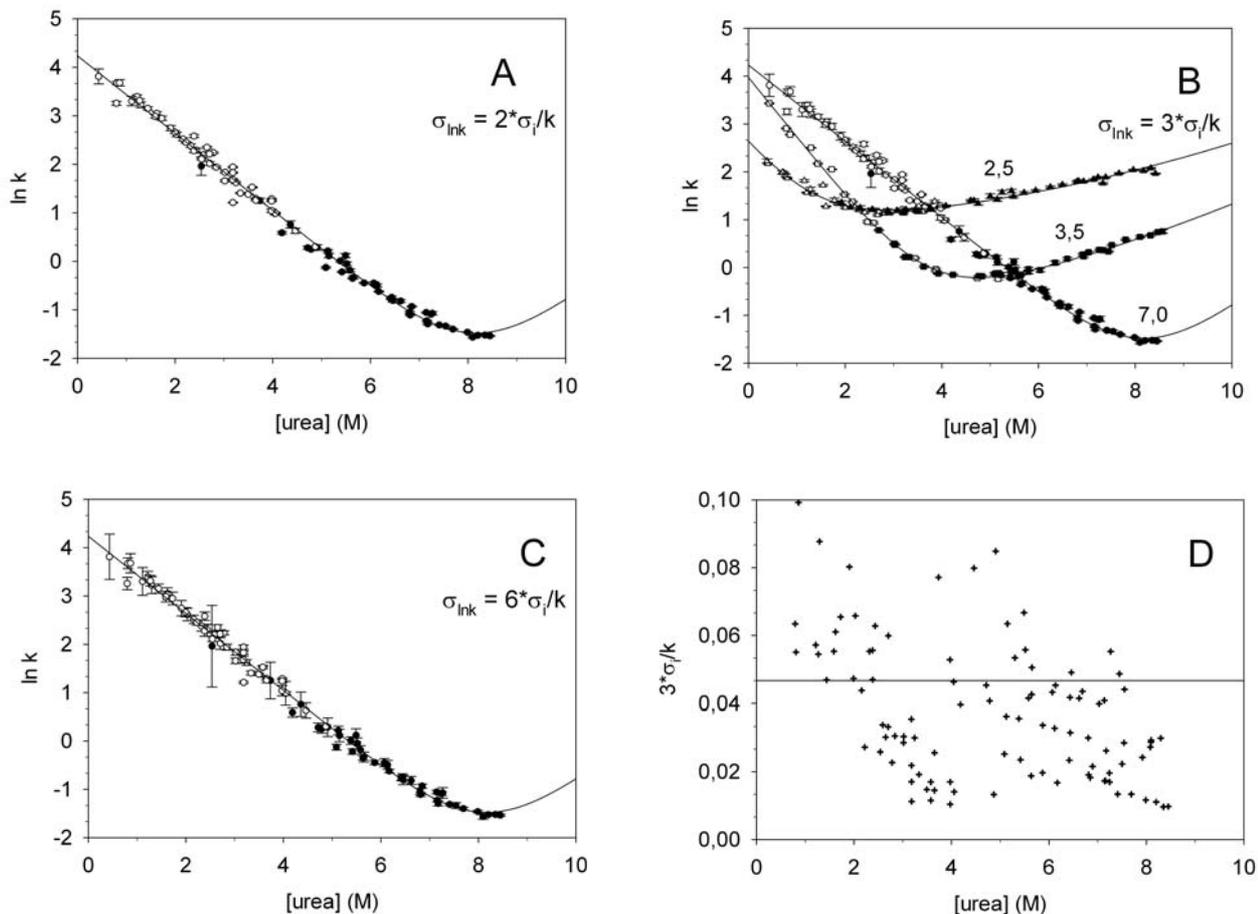


Figura 5.7.2. Gráficos de Chevron del mutante D48G de Spc-SH3 en fosfato sódico 50 mM (pH 7.0, 25°C). Los experimentos se han hecho en dos instrumentos distintos, por diferentes operadores y en distintos años. Las líneas continuas en los paneles A, B y C representan el ajuste por mínimos cuadrados a la ecuación del Chevron. Los símbolos abiertos y cerrados representan los datos experimentales obtenidos para los experimentos de plegamiento y desplegamiento respectivamente. Las flechas indican las barras de error que no se superponen con la línea del ajuste. Las barras de error se han calculado de las estimaciones de los errores del ajuste que resultan con la ecuación 5.1.1. **A.** Dos veces los errores del ajuste ($2\sigma_i$); **B.** Tres veces los errores del ajuste ($3\sigma_i$) a tres valores de pH: pH 7.0 (fosfato sódico 50 mM), pH 3.5 y pH 2.5 (glicina 50 mM); **C.** Seis veces los errores del ajuste ($6\sigma_i$). **D.** Errores relativos en k estimados de la misma forma que en el panel B como función de sus respectivas concentraciones de urea. En este caso, la línea continua representa el valor medio que es 4.7% (0.047).

5.7.3. Errores en los parámetros cinéticos obtenidos del análisis de los gráficos de Chevron

Un gráfico de Chevron, como se describe en la Sección 6.12.4 de Materiales y Métodos, representa la dependencia del logaritmo de la constante cinética de velocidad con la concentración de agente desnaturalizante a una temperatura determinada. En la Figura 5.7.2 (panel A, B y/o C) se muestran los gráficos de Chevron obtenidos para el mutante D48G de Spc-SH3. Las barras de error para las constantes cinéticas se han calculado a partir de la Figura 5.7.1 [$\sigma_k = \pm 2\sigma_i$ (panel A), $\sigma_k = \pm 3\sigma_i$ (panel B), $\sigma_k = \pm 6\sigma_i$ (panel C)], donde σ_i es el error estándar que se obtiene al realizar el ajuste por mínimos cuadrados mediante el algoritmo de Marquardt usando

la ecuación 5.1.1. Así, para un logaritmo natural el error coincide con la magnitud del error relativo: $\sigma_{\ln k} = \sigma_k/k$.

En los gráficos de Chevron que se muestran en la Figura 5.7.2 los ajustes de los datos por mínimos cuadrados a la ecuación que describe el Chevron (ver (Martinez *et al.*, 1998; Martinez y Serrano, 1999) para más detalles) se representan con una línea continua. De este análisis se han obtenido los valores de $k_{\pm N}$, $k_{\pm D}$, $m_{\pm N}$ y $m_{\pm D}$, que están recogidos en la Tabla 5.7.1. En los ajustes a pH 7.0, el valor $m_{\pm N}$ se fijó a un valor de -0.42 M^{-1} (valor medio que se obtiene para los mutantes estudiados en el análisis de los valores de ϕ a pH 7.0, como veremos después), debido a la falta de información en la rama de desplegamiento como consecuencia de la elevada estabilidad de la proteína a este pH. Como los datos experimentales se han obtenido a partir de conjuntos independientes de experimentos, los errores obtenidos deberían relacionarse con los errores “reales”, lo que significa que hay un 68% de probabilidad de incluir el valor verdadero. Los errores relativos, que se obtienen como $3\sigma_x/x$ (aproximadamente con un 90% de confianza “real” a partir del análisis realizado en la Figura 5.7.1), son de un 13% para $k_{\pm D}$, de un 22% para $k_{\pm N}$ y de un 4% para $m_{\pm D}$. Las diferencias entre los errores relativos obtenidos para las constantes cinéticas (casi el doble) pueden deberse a la larga extrapolación a concentración 0 M de agente desnaturante que se hace en la rama de desplegamiento.

Tabla 5.8.1. Parámetros cinéticos con las estimaciones de los errores para el mutante D48G de Spc-SH3 bajo diferentes condiciones experimentales a 25°C.

pH	Método	$m_{\pm N}$ (M^{-1})	$m_{\pm D}$ (M^{-1})	$k_{\pm N}$ (s^{-1})	$k_{\pm D}$ (s^{-1})
Fosfato 50 mM pH 7.0	Ajuste global	0.42 [†]	-0.80 ± 0.03 (4%)	0.00160 ± 0.00036 (22%)	69 ± 9 (13%)
	Ajuste individual	0.42 [†]	-0.79 ± 0.03 (4%)	0.00165 ± 0.00012 (8%)	71 ± 8 (12%)
	Valor medio de los mutantes	0.42 ± 0.02 (4%)	-0.90 ± 0.02 (4%)		
Glicina 50 mM pH 3.5	Ajuste individual	0.50 ± 0.06 (12%)	-1.20 ± 0.09 (8%)	0.08 ± 0.02 (25%)	50 ± 7 (14%)
	Valor medio de los mutantes	0.52 ± 0.06 (12%)	-1.24 ± 0.09 (8%)		
Glicina 50 mM pH 2.5	Ajuste individual	0.37 ± 0.03 (8%)	-0.92 ± 0.18 (20%)	0.94 ± 0.18 (19%)	11.6 ± 1.5 (13%)
	Valor medio de los mutantes	0.39 ± 0.06 (15%)	-0.96 ± 0.09 (9%)		

* Los errores se han calculado como $3\sigma_x$, donde σ_x es el ajuste o la media del error de x. Los valores entre paréntesis representan los errores relativos calculados como $(3\sigma_x/x) \cdot 100$.

[†]Valores fijados en el ajuste por mínimos cuadrados.

Se puede hacer una aproximación alternativa al análisis de errores con los ajustes individuales de los distintos conjuntos de datos a pH 7.0 y con la media de los resultados (Tabla 5.7.1). Si la comparamos con los ajustes globales no se observa dispersión en los datos. Los errores, σ_{μ} , se han calculado usando la siguiente ecuación (Bevington y Robinson, 2003):

$$\sigma_{\mu} = \frac{DE}{\sqrt{N}} \quad (5.7.1)$$

donde DE es la desviación estándar y N es el número de datos, y son aproximadamente de la misma magnitud, excepto para $k_{\pm N}$ cuyo error mejora. Esta mejora es consecuencia de la reproducibilidad (baja dispersión) de los datos de $k_{\pm N}$ que se obtienen de los distintos ajustes. No obstante, puede ser poco fiable, ya que el valor de $m_{\pm N}$ se ha fijado a -0.42 M^{-1} , lo que reduce notablemente la dispersión en los valores de $k_{\pm N}$.

Por otro lado, los errores en el $\ln k$ que se obtienen a diferentes concentraciones de agente desnaturalizante (Figura 5.7.2.D) están distribuidos aleatoriamente alrededor del valor medio (4.7% en nuestro caso), a causa de la falta de correlación con la concentración de agente desnaturalizante. Esto se puede explicar también por el bajo intervalo de variabilidad de los valores de $\ln k$, lo que definitivamente enmascara cualquier correlación entre la magnitud de estos valores y sus correspondientes errores (que se puede traducir en una forma de V en los puntos de la Figura 5.7.2.D), debido a que predominan otras incertidumbres experimentales. Sin embargo, de acuerdo con (Bevington y Robinson, 2003), las barras de error en las constantes cinéticas a diferentes concentraciones de urea se pueden considerar iguales a la media ($\sigma_i \approx \sigma$), y a sus valores extrapolados a concentración 0 M de agente desnaturalizante, que se pueden estimar correctamente de los errores del ajuste, en lugar de usar las fórmulas de propagación, lo que simplifica nuestros cálculos de errores. La disponibilidad de un conjunto de datos experimentales contribuye positivamente a esta equivalencia. Así, las constantes cinéticas en agua también se pueden estimar correctamente a partir de una estimación adecuada de los errores de los ajustes, como se ha hecho previamente (Zarrine-Afsar y Davidson, 2004).

Comparando los resultados obtenidos a pH 7.0 con los obtenidos del análisis de los experimentos a pH 3.5 y 2.5, podemos concluir que los errores en las constantes cinéticas son similares en magnitud, mientras que los errores asociados a los valores de m aumentan drásticamente (Tabla 5.7.1). En ningún caso se ha hecho ninguna restricción en los valores de los parámetros, lo que hace que estos análisis se conviertan en una aproximación más realista que una evaluación cinética común. Hay una discrepancia entre las magnitudes de los errores relativos de $k_{\pm D}$ (12-14%) y $k_{\pm N}$ (19-25%) que era esperada. El error más alto para la constante cinética de desplegamiento resulta de la larga extrapolación a concentración 0 M de desnaturalizante que se hace en este caso. Además, el valor del 25% podría ser considerado como el límite más bajo del error experimental real en los casos de la extremada alta estabilidad, que corresponde a las constantes cinéticas que no se pueden determinar con toda confianza. Así, usando los datos de pH 7.0 como ejemplo, el error en la constante cinética de desplegamiento puede ser más grande si se deja como parámetro flotante del ajuste la pendiente de desplegamiento.

Otra observación interesante, en ausencia de cualquier influencia significativa, es la exactitud en los errores. Si los errores se obtienen de los parámetros cinéticos

derivados de las representaciones Chevron a pH 7.0 (incluyendo los datos de los cinco conjuntos de datos), estos errores se pueden comparar con los estimados a pH 3.5 y 2.5 (solamente de un conjunto de datos) (Tabla 5.7.1), y se puede ver, que a pH 7.0 los errores acontecen para los efectos más relevantes y comunes en la exactitud, junto con las fluctuaciones que vienen de la precisión, mientras que a pH 3.5 y 2.5 acontece solo este último. Es sorprendente que los errores se puedan comparar en magnitud. Así, la posible contribución a los errores sistemáticos (exactitud) parece contrarrestarse por el mayor número de datos de constantes cinéticas que se tienen a pH 7.0, lo que reduce la magnitud del error.

Los errores relativos de los valores de m , se pueden evaluar usando los datos obtenidos para los distintos mutantes diseñados para el análisis de ϕ , que en nuestro caso es solamente $m_{\ddagger-N}$. Es de esperar que el carácter no-destructivo de las mutaciones diseñadas para el análisis de ϕ (Viguera *et al.*, 1996; Martínez *et al.*, 1998; Martínez y Serrano, 1999; Martínez *et al.*, 1999; Vega *et al.*, 2000; Cobos *et al.*, 2003) pueda evitar cualquier efecto significativo sobre los valores de m , principalmente relacionados con la exposición macroscópica de los grupos hidrofóbicos (Myers *et al.*, 1995). De hecho, una comparación de los gráficos de Chevron de los distintos mutantes no revela ninguna diferencia significativa entre sus pendientes (ver Tabla 2 en (Martínez y Serrano, 1999)), que se refleja en el bajo valor estándar de la media (Tabla 5.7.1). Hemos determinado la media de los valores de $m_{\ddagger-N}$ y $m_{\ddagger-D}$. Como podemos ver en la Tabla 5.7.1, los errores obtenidos de los ajustes globales y de los ajustes individuales coinciden.

5.7.4. Análisis de los errores en las energías de Gibbs y en los valores de ϕ

A partir de los valores de las $k_{\ddagger-N}$ y $k_{\ddagger-D}$ que se obtienen del Chevron, se pueden obtener fácilmente los valores de $\Delta\Delta G$ y de ϕ , usando las siguientes ecuaciones:

$$\Delta\Delta G_{\ddagger-D} = -RT \cdot \ln\left(\frac{k_{\ddagger-D,ref}}{k_{\ddagger-D,mut}}\right) \quad \Delta\Delta G_{\ddagger-N} = -RT \cdot \ln\left(\frac{k_{\ddagger-N,ref}}{k_{\ddagger-N,mut}}\right) \quad (5.7.2)$$

y

$$\phi = \frac{\Delta\Delta G_{\ddagger-D}}{\Delta\Delta G_{N-D}} \quad (5.7.3)$$

El valor de $\Delta\Delta G_{N-D}$ se puede calcular de una curva de equilibrio de desnaturalización, aunque también se pueden obtener de los datos cinéticos como $\Delta\Delta G_{N-D} = \Delta\Delta G_{\ddagger-D} - \Delta\Delta G_{\ddagger-N}$.

Utilizando la siguiente ecuación:

$$\sigma_x \approx \sqrt{\sigma_v^2 \left(\frac{\partial x}{\partial v}\right)^2 + \sigma_w^2 \left(\frac{\partial x}{\partial w}\right)^2 + \dots + 2\sigma_{vw}^2 \left(\frac{\partial x}{\partial v}\right) \left(\frac{\partial x}{\partial w}\right) + \dots} \quad (5.7.4)$$

y las ecuaciones 5.7.2 se pueden obtener las expresiones para el cálculo del error de $\Delta\Delta G$, $\sigma_{\Delta\Delta G}$:

$$\sigma_{\Delta\Delta G_{\ddagger-D}} \approx RT \left(\sigma_{\ln k_{\ddagger-D,ref}} + \sigma_{\ln k_{\ddagger-D,mut}} + \sqrt{2 \cdot \sigma_{\ln k_{\ddagger-D,ref}} \cdot \sigma_{\ln k_{\ddagger-D,mut}}} \right) \quad (5.7.5)$$

$$\sigma_{\Delta\Delta G_{\ddagger-N}} \approx RT \left(\sigma_{\ln k_{\ddagger-N,ref}} + \sigma_{\ln k_{\ddagger-N,mut}} + \sqrt{2 \cdot \sigma_{\ln k_{\ddagger-N,ref}} \cdot \sigma_{\ln k_{\ddagger-N,mut}}} \right) \quad (5.7.6)$$

donde el tercer término es el término de cruzamiento (*cross-term*) que acontece por la interdependencia de ambos errores. Esto, junto con la ecuación 5.7.3 nos da la fórmula para calcular los errores en los valores de ϕ :

$$\sigma_{\phi} \approx \sqrt{\frac{\Delta\Delta G_{N-D}^2 \cdot \sigma_{\Delta\Delta G_{\ddagger-D}}^2 - \Delta\Delta G_{\ddagger-D}^2 \cdot \sigma_{\Delta\Delta G_{N-D}}^2 + 2 \cdot \Delta\Delta G_{N-D} \cdot \Delta\Delta G_{\ddagger-D} \cdot \sigma_{\Delta\Delta G_{N-D}} \cdot \sigma_{\Delta\Delta G_{\ddagger-D}}}{\Delta\Delta G_{N-D}^4}} \quad (5.7.7)$$

donde $\sigma_{\Delta\Delta G_{N-D}} = \sigma_{\Delta\Delta G_{\ddagger-D}} + \sigma_{\Delta\Delta G_{\ddagger-N}}$, o alternativamente, es el error que corresponde a los experimentos de desplegamiento en equilibrio, que deben de tener la misma magnitud o mayor. Los dos primeros términos nos dan los errores en las energías de Gibbs de activación y de equilibrio respectivamente, mientras que el tercero origina el término de cruzamiento (*cross-term*) de la ecuación 5.7.4. En todos los casos hemos considerado que $\sigma_{vw} = \sigma_v \cdot \sigma_w$.

5.7.5. Estrategia para el trabajo experimental y el análisis de errores

En base a los argumentos que hemos descrito, un análisis correcto de los errores ($\geq 90\%$ de confianza en la práctica) se puede hacer fácilmente a partir de $\sigma_{\ln k} = \pm 3\sigma_i/k$, donde σ_i representa el error estándar obtenido para $k_{\ddagger-D}$ y $k_{\ddagger-N}$, que se obtienen a partir de los ajustes de los gráficos de Chevron. Las ecuaciones 5.7.5, 5.7.6 y 5.7.7 nos sirven para estimar los errores en las funciones de $\Delta\Delta G$ y en los valores de ϕ respectivamente.

CAPÍTULO 6: MATERIALES Y MÉTODOS

CAPÍTULO 6: MATERIALES Y MÉTODOS

6.1. Clonación de las quimeras basadas en la topología del permutante circular S19P20s por mutagénesis dirigida

Actualmente es posible sintetizar, manipular y disponer de cantidades ilimitadas de cualquier gen que codifique una proteína de interés. Además podemos producir alteraciones en el material genético incluyendo cambios en la secuencia (mutaciones), eliminación de parte de la misma (supresiones) o introducción de fragmentos nuevos en la misma (inserciones). Todo esto ha permitido el nacimiento de lo que se conoce como *Ingeniería de Proteínas*. El método de la *Reacción en Cadena de la Polimerasa* (RCP) (Mullis y Faloona, 1987; Ausubel *et al.*, 1992; Creighton, 1993) nos permite producir las alteraciones deseadas en la secuencia de ADN que se use como molde, diseñando los oligonucleótidos adecuados. Así, se generan copias del nuevo ADN conteniendo la secuencia del gen que codifica la nueva proteína, al que se denomina ADN *recombinante* por proceder de la “combinación” de moléculas de ADN de distinto origen. Éste método se ha escogido para la producción de las proteínas quiméricas que se estudian en esta Memoria. Nuestro fin fue enlazar por su extremo N-terminal la secuencia del decapeptido p41 al extremo C-terminal de la cadena del permutante circular S19P20s del dominio de Spc-SH3 mediante una secuencia de conexión de 2 ó 3 aminoácidos. Hemos clonado cuatro quimeras con secuencias de conexión de diferentes longitudes y composición.

Con la RCP el ADN se amplifica según un proceso cíclico, en el que cada uno de sus ciclos consta de tres pasos básicos (Figura 6.1.1): en el primer paso se despliega completamente la doble cadena del ADN molde que contiene la secuencia que se desea modificar, calentando la mezcla de reacción entre 1 y 3 minutos a 94°C. A continuación se reduce la temperatura para que tenga lugar la hibridación específica de los oligonucleótidos con sus respectivas cadenas complementarias del ADN molde. La elección de la temperatura de hibridación es un factor crítico, ya que la especificidad de la reacción depende de ella. Cuando la temperatura es muy elevada, no se produce hibridación y, si es muy baja, tiene lugar de forma inespecífica. En general la temperatura suele estar entre 55°C y 65°C. El tiempo de duración de esta etapa es de 1 minuto. Por último, se copian las secuencias molde a partir de los dos oligonucleótidos mediante la actuación de la Taq ADN-polimerasa, cuya temperatura óptima de actuación son 72°C, que coincide con la temperatura a la que se produce la extensión de los oligonucleótidos. El tiempo de duración de esta etapa es de 1 minuto.

Tras realizarse estos tres pasos, se completa el primer ciclo de la reacción, con lo que se obtienen dos nuevos fragmentos de ADN a partir del molde. Así, en los ciclos sucesivos se produce una amplificación del ADN de forma exponencial hasta alcanzar 2^{n-1} copias del nuevo ADN, donde n es el número total de ciclos que normalmente está entre 25-35. Una vez que se ha completado el último ciclo programado se mantienen las mezclas de reacción entre 5 y 15 minutos a 72°C, con objeto de asegurar una completa extensión de todas las cadenas de ADN. En nuestro caso el esquema que se siguió para realizar las RCPs que se recogen en esta Memoria viene descrito en la Figura 6.1.2. Las reacciones se han preparado siguiendo el siguiente protocolo:

- a) 68 μ L de agua estéril.
- b) 10 μ L de tampón KCl 100 mM, Tris-HCl 200 mM (pH 8.8 a 25°C), $(\text{NH}_4)_2\text{SO}_4$ 100 mM, MgSO_4 20 mM y 1% de Triton X-100 (10x).
- c) 10 μ L de oligonucleótido 5' de concentración 10 μ M.

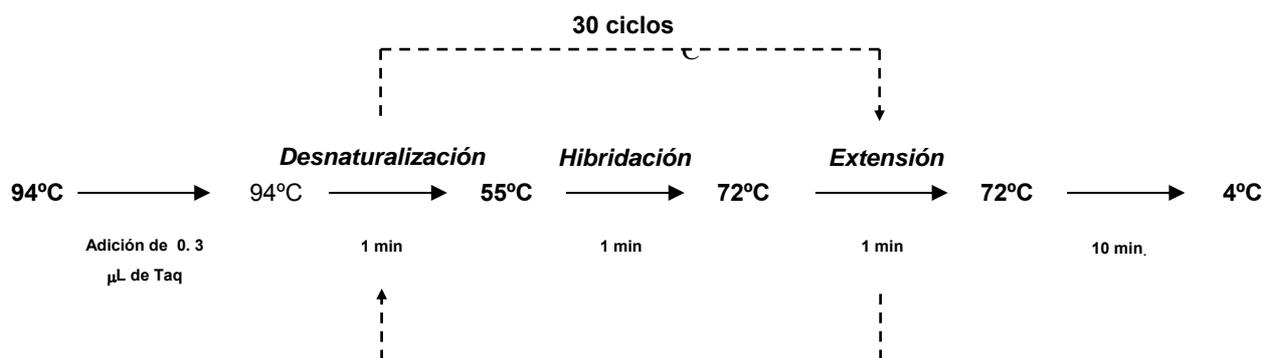


Figura 6.1.2. Esquema general de los ciclos programados para la realización de las RCPs llevadas a cabo en esta Memoria.

6.1.1. Diseño e introducción de alteraciones en fragmentos de ADN

El diseño de los oligonucleótidos definidos como 5' y 3' en el protocolo de RCP, sigue comúnmente una serie de pautas que repercuten en el rendimiento de la RCP: i) la longitud de éstos que debe estar comprendida entre 20 y 30 nucleótidos, aunque pueden diseñarse conteniendo hasta 80-100 nucleótidos; ii) el contenido en bases debe ser tal que entre el 40% y el 60% del total sean citosinas (C) y guaninas (G); iii) deben acotarse en sus extremos con guaninas y citosinas siempre que sea posible, ya que estas bases establecen al aparearse más puentes de hidrógeno que los formados entre las bases A y T, mejorando así la hibridación de los extremos de las cadenas a la hebra correspondiente de ADN.

Para cada cebador diseñado podemos calcular valor de la temperatura de fusión, T_f , parámetro puramente empírico y orientativo que nos permite evaluar la temperatura a la que se desnaturaliza la doble hebra formada por el ADN molde y el oligonucleótido correspondiente. La T_f se calcula con la siguiente relación (Thein, 1986):

$$T_f = [2 \cdot (N_A + N_T)] + [4 \cdot (N_G + N_C)] \text{ (}^\circ\text{C)} \quad (6.1.1)$$

siendo N_A , N_T , N_C y N_G el número total de adeninas, timinas, citosinas y guaninas respectivamente contenidas en parte de la secuencia del oligonucleótido que hibrida con el correspondiente ADN molde durante la RCP. La T_f usual de los cebadores deber ser al menos igual a la temperatura de hibridación, que normalmente está entre 55°C y 65°C.

En todos los casos, el oligonucleótido 5' comienza con la secuencia diana de una enzima de restricción, a la que debe seguir la secuencia de las bases ATG, correspondiente al aminoácido metionina (M), necesaria para que pueda iniciarse la traducción del ADN. En nuestro caso hemos usado el enzima de restricción *NcoI* (CCATGG) que delimita el comienzo de la secuencia en la dirección 5'-3' de la misma. Su secuencia diana incluye el codón ATG del aminoácido metionina que constituye además el punto de inicio de la secuencia del gen. El oligonucleótido 3' se hibrida sobre la cadena complementaria del ADN molde. Seguidamente se incluyen dos codones de terminación (ATCATC) para impedir que en la traducción del ADN puedan codificarse por error las secuencias de nuevos restos aminoacídicos que se añadirían a la secuencia original. A continuación se incluye la secuencia diana correspondiente a

la enzima de restricción *HindIII* (TTCGAA). Por otra parte, cuando se desea realizar una mutación, los oligonucleótidos contienen el fragmento de la secuencia del gen con la mutación que se desea realizar. Para diseñarlos es necesario realizar en la secuencia de ADN de la proteína las sustituciones de las bases apropiadas que codifiquen los nuevos aminoácidos en las posiciones seleccionadas. Como se observa en la Figura 6.1.3, algunos aminoácidos pueden estar codificados por distintas combinaciones de tres bases. Esto permite escoger a conveniencia los cambios en las mismas al hacer una mutación, procurando alterar mínimamente la secuencia original del gen de la proteína.

		Segunda letra					
		U	C	A	G		
Primera letra (extremo 5')	U	UUU] phe UUC] UUA] leu UUG]	UCU] ser UCC] UCA] UCG]	UAU] tyr UAC] UAA detención UAG detención	UGU] cys UGC] UGA detención UGG detención	U C A G	Tercera letra (extremo 3')
	C	CUU] leu CUC] CUA] CUG]	CCU] pro CCC] CCA] CCG]	CAU] his CAC] CAA] CAG]	CGU] arg CGC] CGA] CGG]	U C A G	
	A	AUU] ile AUC] AUA] AUG met	ACU] thr ACC] ACA] ACG]	AAU] asn AAC] AAA] lys AAG]	AGU] ser AGC] AGA] arg AGG]	U C A G	
	G	GUU] val GUC] GUA] GUG]	GCU] ala GCC] GCA] GCG]	GAU] asp GAC] GAA] glu GAG]	GGU] gly GGC] GGA] GGG]	U C A G	

Figura 6.1.3. Representación del código genético. Los grupos de tres nucleótidos (codones) (escritos en la dirección 5'→3') de una molécula de ADN se traducen en aminoácidos de acuerdo con las reglas indicadas.

El proceso de diseño de los oligonucleótidos necesarios para producir las mutaciones en la proteína se ha realizado con ayuda del software *BioEdit Sequence Alignment Editor* (<http://www.mbio.ncsu.edu/BioEdit/bioedit.html>), que permite la modificación de las secuencias de ADN que se le introduzcan, traducirlas a secuencia de aminoácidos y seleccionar los fragmentos dentro de la misma, que servirán para el diseño de los oligonucleótidos. Estos fragmentos de ADN se obtienen por síntesis química y se han comprado a BONSAI TECHNOLOGIES.

Quimera basada en la topología del permutante circular S19P20s

A

```

5'-ATG GGA CCT CGG GAG GTG ACT ATG AAG AAA GGA GAT ATT CTA ACC
TAC CCT GGA GCC CTC CAC TGA TAC TTC TTT CCT CTA TAA GAT TGG
M G P R E V T M K K G D I L T

CTG CTC AAC AGC ACC AAC AAG GAC TGG TGG AAG GTT GAA GTT AAC
GAC GAG TTG TCG TGG TTG TTC CTG ACC ACC TTC CAA CTT CAA TTG
L L N S T N K D W W K K V E V

GAT CGT CAG GGC TTT GTA CCA GCT GCC TAT GTG AAA AAA CTA GAT
CTA GCA GTC CCG AAA CAT GGT CGA CGG ATA CAC TTT TTT GAT CTA
N D R Q G F V P A A Y K K L D

TCC GGT ACT GGT AAA GAG CTT GTG CTA GCA CTC TAT GAT TAC CAA
AGG CCA TGA CCA TTT CTC GAA CAC GAT CGT GAG ATA CTA ATG GTT
S G T G K E L V L A L Y D Y Q

GAG AGT X GCT CCG AGT TAT AGT CCG CCA CCT CCA CCG-3'
CTC TCA CGA GGC TCA ATA TCA GGC GGT GGA GGT GGC
E S A P S Y S P P P P P P
    
```

▪ **Oligonucleótido 5':**

B

```

NcoI
5'-CGC GCG CC ATG GGA CCT CGG GAG GTG ACT ATG AAG AAA GG-3'
M G P R E V T M K K
    
```

▪ **Oligonucleótidos 3':**

Secuencia de conexión **ggc**

```

HindIII Stop Stop
3'-GCG CGC AAG CTT CTA CTA CCG TGG AGG TGG CGG ACT ATA ACT CGG AGC
P P P P S Y S P A
ggc gcc gcc ACT CTC TTG GTA ATC-5'
G G G S E Q Y D
    
```

Secuencia de conexión **gdn**

```

HindIII Stop Stop
3'-GCG CGC AAG CTT CTA CTA CCG TGG AGG TGG CGG ACT ATA ACT CGG AGC
P P P P P S Y S P A
ggt gtc acc ACT CTC TTG GTA ATC -5'
N D G S E Q Y D
    
```

Secuencia de conexión **gng**

```

HindIII Stop Stop
3'-GCG CGC AAG CTT CTA CTA CCG TGG AGG TGG CGG ACT ATA ACT CGG AGC
P P P P P S Y S P A
gcc gtt atc ACT CTC TTG GTA ATC-5'
G D N S E Q Y D
    
```

Secuencia de conexión **gnt**

```

HindIII Stop Stop
3'- GCG CGC AAG CTT CTA CTA CCG CGG CGG CGG CGG GCT ATA GCT CGG CGC
P P P P P S Y S P A
gnt atc GCT TTC CTG ATA ATC-5'
N D S E Q Y D
    
```

Figura 6.1.4. A. Secuencia de ADN del gen que codifica a la quimera basada en la topología del permutante circular y su correspondiente traducción según el mapa genético de *E. Coli*. La **X** representa las distintas secuencias de conexión. En color rojo y azul se resaltan las zonas de hibridación de los oligonucleótidos 5' y 3' respectivamente. **B.** Secuencia de los oligonucleótidos empleados para realizar las RCPs encaminadas a producir las quimeras estudiadas en esta memoria. Se resaltan en gris, las secuencias diana correspondientes a los dos enzimas de restricción empleados (*NcoI* y *HindIII*) para cortar específicamente el producto de amplificación y en negro se resaltan los dos codones de terminación de la traducción. Las secuencias de conexión se indican resaltadas en verde.

Para clonar las proteínas quiméricas hemos diseñado cada pareja de oligonucleótidos de acuerdo con la secuencia de bases que se deseaba obtener finalmente (Figura 6.1.4.A). Cada uno de los oligonucleótidos se hibrida con un extremo de la cadena del ADN molde. Los oligonucleótidos 5' se hibridan con el extremo que corresponde al extremo N-terminal de la proteína. En este caso diseñamos uno solamente (Figura 6.1.4.B), porque la secuencia de bases en el extremo N-terminal no iba a variar, ya que la secuencia del péptido precedido de la secuencia de conexión se ha enlazado al extremo C-terminal de la cadena del dominio SH3. Por tanto, hemos diseñado cuatro oligonucleótidos 3' diferentes (Figura 6.1.4.A) que son los que se hibridan con el extremo C-terminal y en los que varía la secuencia de conexión. Como ADN molde hemos usado el gen que codifica al permutante circular S19P20s del dominio de Spc-SH3.

6.1.2. Digestión de los productos de la RCP

El producto final purificado que se obtiene de la RCP se digiere a continuación. Este proceso se lleva a cabo gracias a la actuación de dos endonucleasas de restricción: *NcoI* y *HindIII*. Estas enzimas producen la rotura en un punto específico de un fragmento de ADN que constituye el sitio de reconocimiento o de restricción del enzima, y que se caracteriza por estar constituido por una secuencia concreta de bases, cuya cadena complementaria opuesta es idéntica (secuencia palindrómica). En la Figura 6.1.4.B podemos observar el sitio de corte para estas enzimas. El protocolo usado para la digestión es el siguiente:

- a) 2.5 μL de tampón de corte Tris HCl 10 mM, MgCl_2 5 mM, NaCl 100 mM, 2-mercaptoetanol 1 mM, pH 8.0 (37 °C) (10x).
- b) 1 μL de disolución de *NcoI* (10 U/ μL).
- c) 1 μL de disolución de *HindIII* (15 U/ μL).
- d) 20.5 μL del producto de RCP puro.

Para que se produzca la digestión en condiciones óptimas es necesario incubar la reacción a una temperatura de 37°C durante 2 horas. Una vez completada la reacción, la enzima que queda sin reaccionar se inactiva incubando la mezcla de reacción a 65°C durante 15 minutos. Los productos del corte, se vuelven a purificar según el protocolo comercial de purificación de QIAGEN (*PCR Purification Kit, Qiaquick*).

6.1.3. Ligación del ADN digerido al vector de expresión

La ligación permite unir el ADN digerido (inserto) a un vector de expresión para formar un nuevo plásmido. En la práctica se pueden producir otros tipos de asociación entre fragmentos de restricción con extremos compatibles que dan lugar a la aparición de productos secundarios no deseados. Para evitar esto, antes de realizar la digestión se prepara el vector que se va a emplear, mediante un proceso de digestión previa (con los mismos enzimas de restricción utilizados en la digestión del ADN), purificación y desfosforilación de los extremos del mismo mediante incubación durante 1 hora con fosfatasa alcalina para evitar su recircularización. En nuestro caso, los vectores de expresión utilizados son pBAT4 (Peranen *et al.*, 1996) y pETM-11. Las ligaciones se han realizado según un protocolo comercial de ligación de ROCHE (*Rapid Ligation Kit*).

6.1.4. Introducción del plásmido en el organismo de expresión

El plásmido que contiene el gen que codifica la proteína nos sirve para la introducción y expresión de dicho gen recombinante en bacterias *Escherichia coli*. En este proceso, conocido como transformación, el ADN penetra en la bacteria mediante el plásmido y es capaz de replicarse de forma autónoma dentro de la célula. Ésta, además, adquiere resistencia a un antibiótico determinado gracias a un gen específico contenido en el plásmido, de forma que podemos hacer una selección de las bacterias que contengan el plásmido cultivándolas en un medio que contenga dicho antibiótico. En nuestro caso, de los dos vectores utilizados el vector pBAT4 contiene el gen de resistencia al antibiótico ampicilina y el vector pETM-11 contiene el gen de resistencia al antibiótico kanamicina.

Después de su selección en placa de LB-agar, las células se trasladan y cultivan en medio LB. La estrategia seguida para introducir el plásmido y expresar la proteína en las células implica una transformación previa en células de la cepa XL1-Blue que poseen elevada capacidad de transformación lo que nos permite amplificar el plásmido que procede de la ligación. Una vez amplificado se transforman con él las células competentes de la cepa BL21 (DE3) para expresar la proteína ya que éstas tienen una gran capacidad de expresión de proteína. El paso previo de transformación y amplificación en células XL1-Blue es necesario por que las células de la cepa BL21 (DE3) no poseen una capacidad transformante suficiente para captar las pequeñas cantidades de plásmido procedentes de una reacción de ligación. El protocolo de transformación utilizado es el protocolo comercial de NOVAGEN para las células BL21 (DE3) y el de STRATAGENE para las células XL1-Blue.

6.1.5. Ensayos de expresión

Se seleccionan varias colonias (4-5) de las placas Petri donde se han incubado las células transformadas la noche anterior y se prepara para cada una un minicultivo de 3 mL de medio LB en presencia de ampicilina (0.5 $\mu\text{L}/\text{mL}$), que se deja crecer en un incubador a 37°C con agitación fuerte. Para cada colonia se hacen ensayos variando el tiempo de inducción con isopropil- β -D-tiogalactopiranosido (IPTG) (0.5 $\mu\text{L}/\text{mL}$), que es un agente inductor de la expresión del plásmido, de manera que se puede así establecer también qué grado de crecimiento del cultivo es el óptimo para inducir la expresión, si bien lo más usual es inducirla cuando las células se encuentran en fase exponencial de crecimiento, lo que se puede comprobar midiendo la absorbancia del medio de cultivo a 600 nm y comprobando que ésta tiene un valor situado entre 0.6 y 0.8 aproximadamente. Los minicultivos se dejan crecer toda la noche en el incubador a 37°C y a la mañana siguiente se recogen muestras que se analizan mediante electroforesis en gel de poliacrilamida al 16% en presencia de SDS (SDS-PAGE) según el método de (Laemmli, 1970) para estimar el nivel de expresión de la proteína recombinante. De todas las colonias que expresan nuestra proteína se eligen las que presentan las bandas correspondientes de mayor intensidad en el gel.

6.1.6. Purificación del plásmido

El paso siguiente implica la purificación del plásmido de las células que expresan la proteína partiendo de las colonias seleccionadas en las placas. Para ello se traspasan las colonias elegidas a un minicultivo de 15 mL de LB con ampicilina y se incuban a 37°C toda la noche. A continuación se recogen las células por centrifugación. El proceso que sigue para aislar el plásmido comprende el uso de un

conjunto de reactivos comerciales de ROCHE (*High Pure Plasmid Isolation Kit*), obteniéndose un volumen total de 50 μL de plásmido puro en agua estéril, que se conserva congelado a $-20\text{ }^{\circ}\text{C}$.

6.1.7. Secuenciación

Los plásmidos obtenidos tras todo este proceso de clonación se secuencian para comprobar que la secuencia de bases del gen que codifica nuestra proteína es la correcta. En nuestro caso se ha utilizado el Servicio de Secuenciación del Instituto de Parasitología y Biomedicina “López Neyra” del Consejo Superior de Investigaciones Científicas, en Granada. Otra prueba para comprobar la identidad de la proteína consiste en, una vez expresada y purificada según el protocolo correspondiente, analizarla mediante espectrometría de masas para verificar si su peso molecular es el esperado: para ello hemos utilizado el servicio de Espectrometría de Masas del Centro de Instrumentación Científica (C.I.C.), de la Universidad de Granada.

6.2. Clonación de mutantes de la quimera SPCp41

Todos los mutantes de la quimera SPCp41 se han clonado utilizando un protocolo comercial de mutagénesis dirigida de la casa comercial STRATAGEN (*Quickchange Site-Directed Mutagenesis Kit*). Este método presenta muchas ventajas respecto a otras técnicas, ya que en sólo cuatro pasos genera ADN mutado con más de un 80% de eficiencia. Con este método se pueden hacer mutaciones puntuales y múltiples en un mismo plásmido. El procedimiento básico utiliza el vector de ADN molde de cadena doble que contiene el gen de la proteína original y dos oligonucleótidos sintéticos que contienen la mutación deseada, complementarios y de cadena sencilla (como ya se ha explicado en la Sección 6.1). En este caso para llevar a cabo la RCP se utiliza la PfuTurbo® ADN polimerasa que tiene capacidad correctora. El plásmido mutado que se genera se trata con la enzima endonucleasa *DpnI*, que se usa para digerir el ADN molde y así obtener el ADN sintetizado que contiene la(s) mutación(es) deseadas. El vector que incorpora la mutación es transformado seguidamente en células competentes comerciales XL1-Blue.

La poca cantidad de ADN molde requerida, la alta fidelidad de la PfuTurbo® ADN polimerasa y el bajo número de ciclos necesarios en el programa del termociclador dan como resultado una alta eficiencia de mutación y una minimización de las mutaciones aleatorias durante la reacción.

En la Figura 6.1.4.A se representa la secuencia del gen que codifica a la quimera SPCp41, que es la construcción sobre la que se han diseñado todos los mutantes estudiados en esta Memoria.

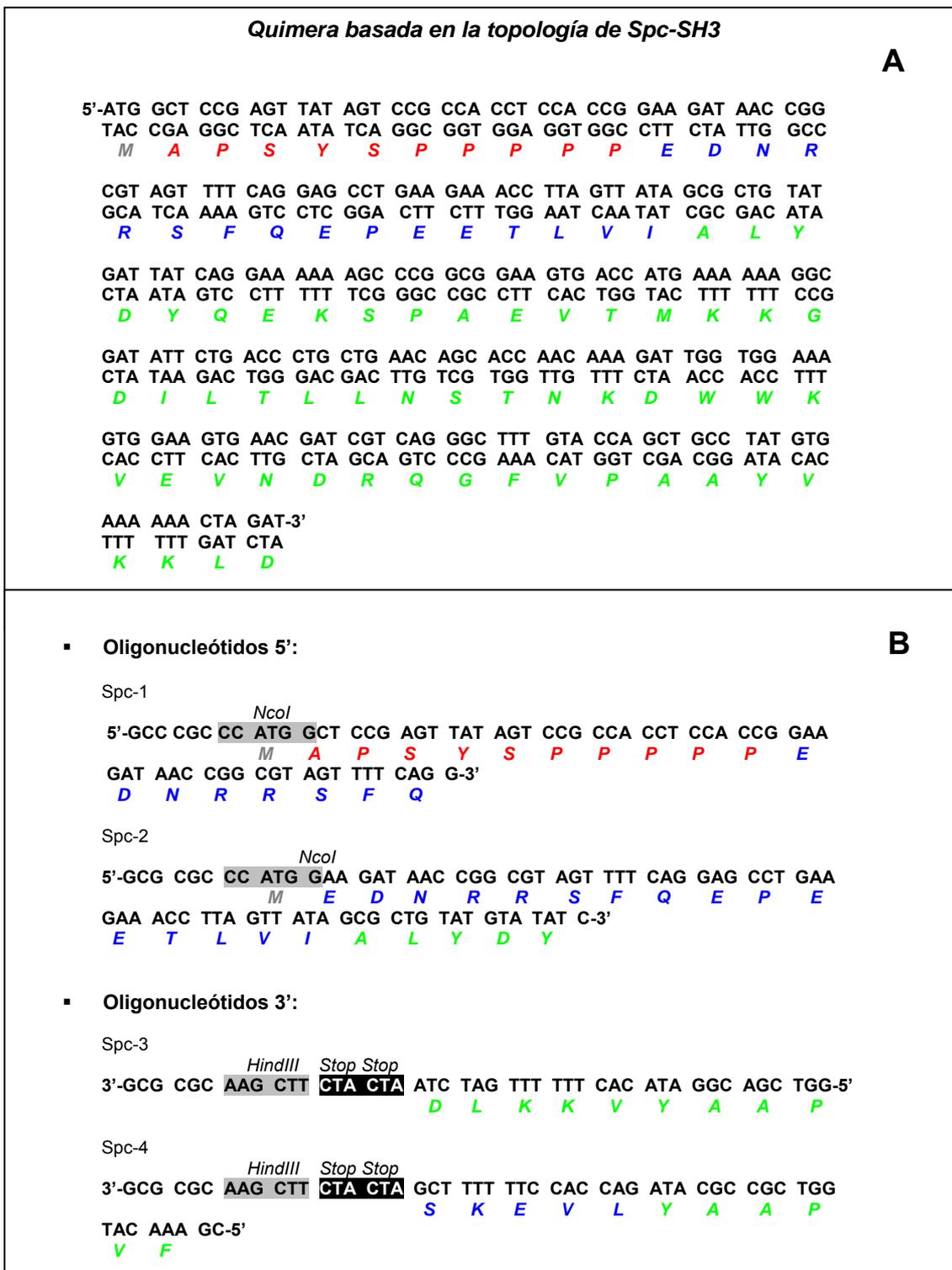


Figura 6.3.1. A. Secuencia de ADN del gen que codifica a la quimera basada en la topología de Spc-SH3 y su correspondiente traducción según el mapa genético de *E. Coli*. En rojo se ha representado la secuencia que corresponde al péptido p41, en azul la secuencia de conexión del dominio Itk-SH3 y en verde la secuencia del dominio natural de Spc-SH3. **B.** Secuencia de los oligonucleótidos empleados para realizar las RCPs encaminadas a producir las quimeras. Se resaltan en gris, las secuencias diana correspondientes a los dos enzimas de restricción empleados (*NcoI* y *HindIII*) para cortar específicamente el producto de amplificación y en negro se resaltan los dos codones de terminación de la traducción.

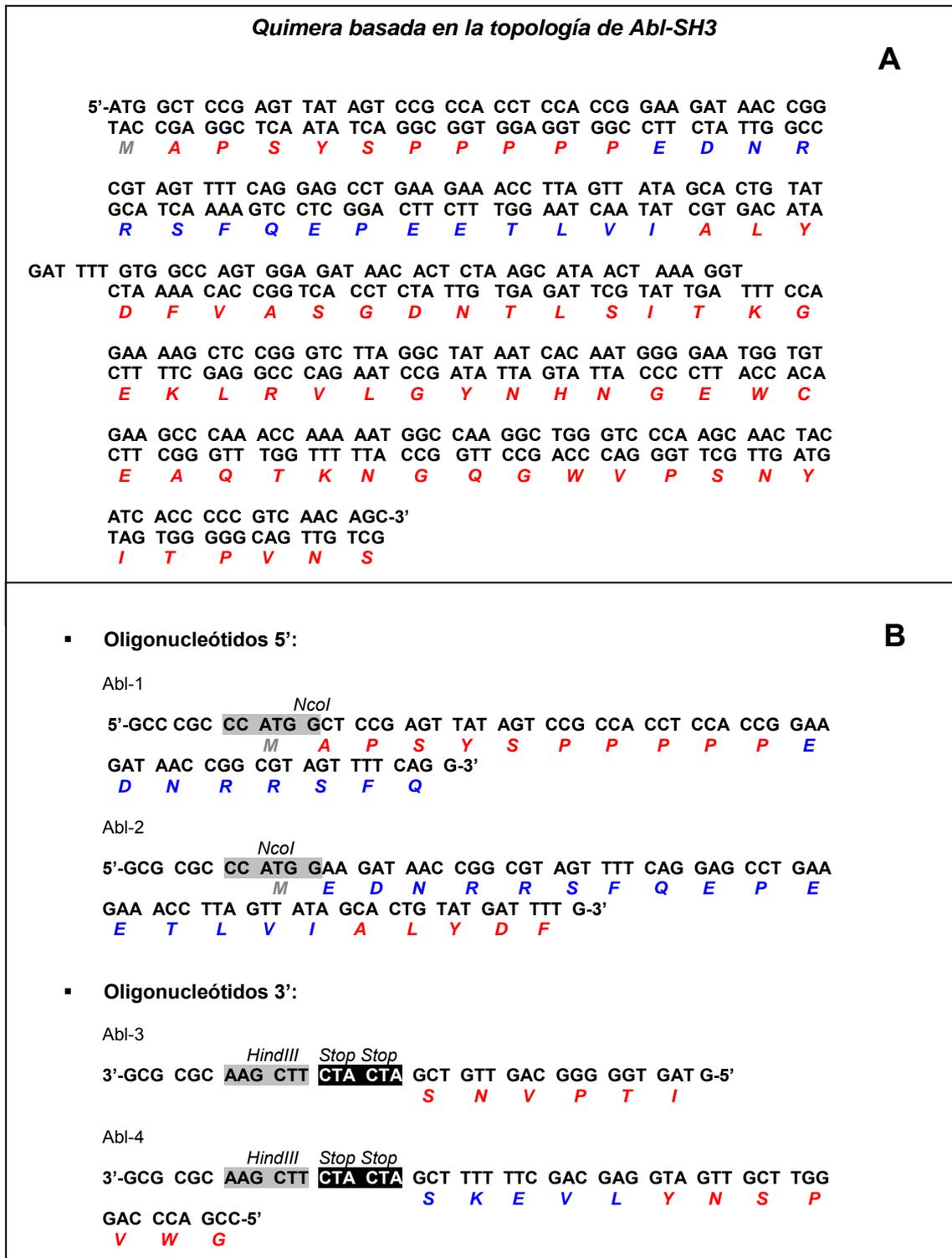


Figura 6.3.2. A. Secuencia de ADN del gen que codifica a la quimera basada en la topología de Abl-SH3 y su correspondiente traducción según el mapa genético de *E. Coli*. En rojo se ha representado la secuencia que corresponde al péptido p41 y la secuencia del dominio Abl-SH3 y en azul la secuencia de conexión del dominio *lck*-SH3. **B.** Secuencia de los oligonucleótidos empleados para realizar las RCPs encaminadas a producir la quimera. Se resaltan en gris, las secuencias diana correspondientes a los dos enzimas de restricción empleados (*NcoI* y *HindIII*) para cortar específicamente el producto de amplificación y en negro se resaltan los dos codones de terminación de la traducción.

6.3. Clonación de las quimeras basadas en la topología de Itk-SH3

Para clonar las quimeras de Spc-SH3 y Abl-SH3 con topología basada en Itk-SH3, se ha diseñado cada pareja de oligonucleótidos de acuerdo con la secuencia de bases que se deseaba obtener finalmente (ver Figuras 6.3.1 y 6.3.2) y se ha seguido el protocolo descrito en la Sección 6.1. La clonación ha requerido el uso de dos RCP, que denominaremos, RCP#1 y RCP#2, y de una tercera, RCP#3, para modificar el extremo C-terminal de acuerdo con la secuencia del dominio Itk-SH3. Las dos primeras reacciones de RCP se realizan de forma análoga en cuanto al procedimiento, que es el mismo que se ha descrito en la Sección 6.1.1, aunque difieren en los oligonucleótidos que intervienen en la reacción, así:

- RCP#1: oligonucleótido Spc-2/Abl-2 + oligonucleótido Spc-3/Abl-3
- RCP#2: oligonucleótido Spc-1/Abl-1 + oligonucleótido Spc-3/Abl-3

En la RCP#1 se adicionó la secuencia de conexión al dominio correspondiente, y en la RCP#2 se añadió la secuencia del péptido p41 al extremo N-terminal de las construcciones obtenidas en la RCP#1, en la que se ha usado como ADN molde el gen que codifica al dominio natural de Spc-SH3, mientras que en la RCP#2 el ADN molde ha sido el producto purificado obtenido en la RCP#1.

La RCP#3 se prepara de igual manera que las otras dos RCPs, usando como ADN molde el producto purificado de la RCP#2, y usando los siguientes oligonucleótidos:

- RCP#3: oligonucleótido Spc-1/Abl-1 + oligonucleótido Spc-4/Abl-4

Una vez clonadas las quimeras, se purifican de acuerdo con el procedimiento descrito en la Sección 6.4.1.

6.4. Expresión y purificación de proteínas

Para la purificación de las distintas variantes de la quimera SPCp41 se han utilizado dos métodos distintos, según el vector de expresión en el que insertemos el gen correspondiente. En el caso de pBAT4, la purificación se hace por cromatografía de exclusión molecular, mientras que en el caso de pETM-11 la proteína contiene una cola de histidina en el extremo N-terminal para su purificación por cromatografía de afinidad en columna de Níquel.

Se han utilizado dos vectores distintos debido a un cierto grado de heterogeneidad encontrado en algunas muestras de las proteínas purificadas por cromatografía de exclusión molecular. Esta heterogeneidad producía duplicidad de picos en los espectros de RMN, lo que dificultaba extraordinariamente su análisis. El espectro de masas de éstas mostraba dos picos: uno mayoritario con la masa molecular correcta y otro menor con unos 770 Da menos. Para elucidar el problema se realizó la secuenciación del extremo N-terminal de la proteína y el análisis cuantitativo de aminoácidos en el Servicio de Secuenciación del Centro de Investigaciones Biológicas (CIB) de Madrid. Los resultados pusieron de manifiesto la existencia de dos especies, una mayoritaria que se correspondía con la proteína completa y otra minoritaria que aparecía truncada por el extremo N-terminal desde la Met8. Los 7 primeros aminoácidos tienen un peso molecular de aproximadamente 770 Da, que es la diferencia de masa encontrada entre los dos picos del espectro de masas.

Para solucionar este problema nos planteamos varias posibilidades: i) repurificar y secuenciar de nuevo el plásmido de la quimera por si tuviera alguna heterogeneidad, ii) utilizar un cocktail de inhibidores en la purificación de la proteína por si hubiera alguna proteasa en el medio que la estuviera fraccionando, y iii) cambiar la temperatura de crecimiento de las células a 25°C, para que la expresión fuera más específica. Ninguna de estas posibilidades nos permitió eliminar la especie cortada de nuestra proteína.

Así, la solución definitiva y que nos permitió eliminar completamente esta segunda especie, fue cambiar el inserto de nuestra proteína al vector pETM-11 que incluye una cola de histidina en el extremo N-terminal. De esta forma, la cola de histidina se incorpora a la secuencia de la forma completa de la proteína y no a la troncada, lo que nos permite separarlas por cromatografía de afinidad en columna de Níquel. La mayoría del trabajo de esta Memoria se ha realizado con proteínas purificadas por cromatografía de afinidad. No obstante, como se describe en la Sección 3.5 del Capítulo 3, el extremo N-terminal de la cadena se encuentra desordenado en la estructura de las quimeras y tras realizar controles exhaustivos tanto en los estudios de estabilidad térmica como en las cinéticas de plegamiento-desplegamiento hemos obtenido resultados idénticos con ambos tipos de proteínas, por lo que no se ha hecho distinción entre las mismas.

6.4.1. Cromatografía de exclusión molecular

Se transforman células competentes BL21 (DE3) de *E.Coli.* con el plásmido, mediante el procedimiento de transformación comercial. La expresión se induce con 1mM de IPTG cuando la absorbancia a 600 nm está entre 0.6 y 0.8 y se incuba durante toda la noche a 37°C. Las células se recogen a la mañana siguiente por centrifugación a 4°C (HETTICH *Roto Super 40*, rotor A6.9) y se resuspenden en tampón citrato sódico 5 mM a pH 3.5; después se rompen en una prensa French a 1000 psi y se eliminan los restos celulares por ultracentrifugación (BECKMAN *Optima LE-80*, rotor 45Ti) a 30000 rpm y a 4 °C durante 30 minutos. La proteína se obtiene del sobrenadante por precipitación con sulfato amónico al 75% de saturación y posteriormente se solubiliza en el mínimo volumen posible en tampón fosfato sódico 50 mM, cloruro sódico 100 mM, pH 7.0, que contiene urea a una concentración mayor de 7 M.

Por último la proteína se purifica por cromatografía de exclusión molecular en una columna *HiLoad™ Superdex™ 75 26/60* (AMERSHAM BIOSCIENCES) acoplada a un equipo *AKTA™FPLC System* de PHARMACIA BIOTECH, equilibrando previamente la columna en el mismo tampón. Finalmente las fracciones de proteína pura se dializan de forma prolongada frente a tampón glicina 20 mM, pH 3.0, para eliminar la urea y así renaturalizar la proteína. Al hacer la diálisis a pH ácido conseguimos también eliminar los fragmentos de ADN que hayan podido eluirse de la columna conjuntamente con la proteína, lo cual requiere hacer al menos 4 cambios de 5 L del tampón de diálisis. Las diferentes etapas de la purificación, así como la pureza final de la proteína, se controlan mediante SDS-PAGE al 16%. La pureza e identidad de las proteínas se confirmaron mediante espectrometría de masas en el servicio del Centro de Instrumentación Científica (C.I.C.) de la Universidad de Granada.

6.4.2. Cromatografía de afinidad

El plásmido pETM-11 (*Protein Expression and Purification Core Facility, EMBL Heidelberg, Alemania*) contiene el gen que codifica a las proteínas quiméricas unido a una cola de histidinas (*6xHis-tag*) en el extremo N-terminal mediante una secuencia de conexión que incluye un sitio de reconocimiento para la proteasa TEV lo que nos permite eliminar la cola de histidinas de las proteínas tras su purificación. La expresión se realizó en *E. Coli* BL21 (DE3). El sitio de reconocimiento para la proteasa TEV es Glu-Asn-Leu-Tyr-Phe-Gln-Gly-Ala y el corte se produce entre Gln y Gly, por lo que a la secuencia de nuestra proteína tendrá una Gly y una Ala adicionales en el extremo N-terminal. La expresión se induce con 1mM IPTG cuando la absorbancia a 600 nm está entre 0.6 y 0.8, y se deja crecer durante toda la noche a 37°C. Las células se recogen por centrifugación y se resuspenden en tampón fosfato sódico 50 mM, cloruro sódico 0.3 M, pH 8.0 (Tampón de Equilibrado de Columna: TEC); después se rompen mediante prensa French a 1000 psi y posteriormente se ultracentrifugan a 30000 rpm, 30 minutos a 4°C. El sobrenadante se carga en una columna con 5 mL de resina Ni-NTA (QUIAGEN) previamente equilibrada con TEC y se llevan a cabo sucesivos lavados con TEC, TEC + imidazol 20 mM y TEC + imidazol 50 mM.

Por último, la proteína se eluye con TEC + imidazol 500 mM. Las fracciones que contienen la proteína se dializan frente al tampón de corte de la proteasa TEV (Tris-HCl 50 mM pH 8.0, EDTA 50 mM) para, también, eliminar el imidazol ya que inhibe su actividad. A continuación la proteína se incuba con la proteasa TEV, que fue producida en nuestro laboratorio, en presencia de 1 mM ditiotreitól a temperatura ambiente durante toda la noche. Los productos de la proteólisis se someten a otra etapa de cromatografía en Ni-NTA, de modo que la cola de histidinas y la proteína no hidrolizada se retienen en la columna, recogiendo la proteína pura en las primeras fracciones, las cuales se concentran hasta 2-3 mg·mL⁻¹ y se almacenan a -20°C en el mismo tampón. Las diferentes etapas de la purificación así como la pureza final de la proteína se controlan mediante SDS-PAGE al 16%. Además, la pureza de la proteína y su correcto peso molecular se comprobaron mediante espectrometría de masas.

6.4.3. Purificación de proteína marcada con ¹⁵N y/o ¹³C

Para el marcaje de proteína con ¹⁵N y/o ¹³C se parte de una colonia de bacterias de una placa Petri que se siembra en 10 mL de medio LB con el antibiótico correspondiente y se incuba toda la noche a 37°C con agitación. A la mañana siguiente se inoculan 3 mL de este cultivo a 1 L de medio M9. El medio M9 se prepara a partir de las siguientes disoluciones, que deben ser autoclavadas o esterilizadas filtrándolas con un tamaño de poro de 0.22 μm de manera independiente, para tener un volumen final de 1 L aproximadamente:

*Sales M9	200 mL·L ⁻¹
Agua	780 mL·L ⁻¹
MgSO ₄ ·7H ₂ O (1 M)	1 mL·L ⁻¹
CaCl ₂ ·H ₂ O (0.05 M)	2 mL·L ⁻¹
FeCl ₃ ·6H ₂ O (10%)	4 mL·L ⁻¹
(¹³ C ₆)D-Glucosa (20%)	10 mL·L ⁻¹
Tiamina (0.5%)	2 mL·L ⁻¹
Antibiótico (100 mg/mL)	0.5 mL·L ⁻¹
¹⁵ NH ₄ Cl	1 g·L ⁻¹

*Sales M9 5x (1 L) (pH 7.4)	30 g·L ⁻¹ Na ₂ HPO ₄ 15 g·L ⁻¹ KH ₂ PO ₄ 2.5 g·L ⁻¹ NaCl
--------------------------------	---

La expresión se induce con 1mM IPTG con una absorbancia a 600 nm entre 0.6 y 0.8, y se deja crecer durante toda la noche a 37° con agitación. A partir de aquí, el protocolo de purificación es el mismo que el descrito en la Sección anterior.

6.4.4. Expresión y purificación de la ciclofilina humana *hCyp18*

Se transforman células XL1-Blue de *E.Coli.* con el plásmido pQE70/*hCyp18* (Fanghanel y Fischer, 2003) (el plásmido fue proporcionado por el Dr. Guntert Fischer, Instituto Max Planck, Halle, Alemania, a través del Dr. Thomas Kiefhaber, Universidad de Basel, Suiza), mediante el procedimiento de transformación comercial. La expresión se induce con 1mM de IPTG cuando la absorbancia a 600 nm está entre 0.6 y 0.8 y se incuba durante toda la noche a 37°C. Las células se recogen a la mañana siguiente por centrifugación a 4°C (HETTICH *Roto Super 40*, rotor A6.9) y se resuspenden en tampón tricina 10 mM, pH 8.0 (tampón A); después se rompen en una prensa French a 1000 psi y se eliminan los restos celulares por ultracentrifugación (BECKMAN *Optima LE-80*, rotor 45Ti) a 30000 rpm y a 4 °C durante 30 minutos. La ciclofilina se obtiene del sobrenadante por cromatografía de intercambio aniónico en una columna *Mono Q[®] HR 5/5* de PHARMACIA BIOTECH, equilibrada previamente en el tampón A. La *hCyp18* se eluye con un gradiente de KCl desde 0 a 3 M en tampón A.

A continuación las fracciones que contienen *hCyp18* se pasan por una columna *Blue Sepharosa* de PHARMACIA BIOTECH y se eluye con un gradiente de KCl desde 0 a 3 M en tampón A. Las fracciones que contienen la ciclofilina se dializan en HEPES 10 mM, pH 7.8, KCl 150 mM, MgCl₂ 1.5 mM y 2 mM DTT (tampón B). La *hCyp18* dializada se purifica por cromatografía de exclusión molecular en una columna *HiLoad[™] Superdex[™] 75 26/60* (AMERSHAM BIOSCIENCIAS) acoplada a un equipo *ÄKTA[™]FPLC System* de PHARMACIA BIOTECH, equilibrando previamente la columna en el tampón B. Las diferentes etapas de la purificación, así como la pureza final de la *hCyp18*, se controlan mediante SDS-PAGE al 16%. La actividad de la *hCyp18* se confirmó mediante un ensayo de actividad (Liu *et al.*, 1990).

6.5. Preparación de muestras

6.5.1. Preparación de tampones y de disoluciones de proteína

Las muestras empleadas para la realización de los experimentos recogidos en esta Memoria se dializan en una membrana de diálisis de SPECTRA-POR con un punto de corte de 3500 Daltons frente al tampón correspondiente, según el pH que se desee conseguir. Las disoluciones tampón se preparan a partir de las formas ácida y básica del propio tampón para lo que podemos utilizar una página web como un acceso útil (<http://www.liv.ac.uk/buffers/buffercalc.html>) que nos da las cantidades necesarias de ácido y base para que al mezclarlas obtengamos el pH deseado. El pH se comprueba y corrige en un pHmetro (CRISON). Las muestras dializadas se centrifugan en una microcentrífuga o se filtran con un filtro de nitrocelulosa de tamaño de poro de 0.45 µm (SARSTEDT FiltroPur®). La concentración de las muestras se determina espectrofotométricamente usando un coeficiente de extinción molar de 16502 M⁻¹·cm⁻¹ a 280 nm, determinado previamente por el método de Gill y von Hippel (Gill y Vonhippel, 1989).

Para los experimentos de RMN se emplearon muestras de la quimera SPCp41 y del mutante P74A marcadas con ^{15}N y/o ^{13}C y sin marcar, todas ellas previamente liofilizadas. Se disuelve la cantidad necesaria de proteína, determinada por pesada, para obtener una concentración aproximada de 1 mM en 550 μL de tampón glicina-D5 20 mM, 93% H_2O :7% D_2O (en volumen) o en 100% D_2O . El pH se ajusta a 3.5 con DCI o NaOD utilizando un microelectrodo de pH *Ingold* de 3 mm de diámetro, sin aplicar ninguna corrección por efectos isotópicos.

6.5.2. Preparación de disoluciones concentradas de agente desnaturizante

Las disoluciones de urea se descomponen lentamente originando iones amonio y cianato (Hagel *et al.*, 1971). Estos últimos pueden interaccionar con los grupos amino de la proteína (Stark, 1965) de manera que es aconsejable usar una disolución fresca de urea, preparada el mismo día que se van a hacer los experimentos o en su defecto, el día anterior. Es posible también, congelar la disolución y utilizarla al cabo del tiempo. A continuación se detallan los pasos seguidos en la preparación de una disolución madre de urea, aproximadamente 10 M, en tampón fosfato sódico 50 mM, pH 7.0:

En primer lugar se pesa la cantidad necesaria de urea para disolverla después en un volumen de agua Milli-Q menor que el volumen final deseado, debido al aumento de volumen que tiene lugar durante la solubilización. Dado que estamos muy cerca de la saturación (la solubilidad de la urea a 25°C es 10.49 M) y con objeto de acelerar el proceso de solubilización, calentamos ligeramente sin llegar a superar los 30°C, ya que se favorecería la descomposición. A continuación purificamos la disolución de urea, de los iones fruto de su descomposición, utilizando una resina de intercambio iónico, AG501-X8(D) (BIORAD). Una vez purificada la urea, preparamos el tampón según se ha descrito en la Sección anterior. Después se mide el pH para confirmar que no hay cambios significativos ya que la presencia de desnaturizantes influye en las medidas de pH. Para finalizar se filtra la disolución resultante con un filtro de 0.45 μm .

La concentración exacta de las disoluciones de urea se determina a partir de la medida del índice de refracción, n (Pace *et al.*, 1989), mediante un refractómetro de mano ATAGO modelo R5000. La diferencia entre el índice de refracción de la disolución de urea y el tampón acuoso en el que se disuelve, $\Delta n = n_{\text{urea}} - n_{\text{tampón}}$, nos dará un Δn , que está relacionado con la concentración de urea a través de la ecuación (Warren y Gordon, 1966):

$$[\text{urea}] = 117.66 \cdot (\Delta n) + 29.753 \cdot (\Delta n)^2 + 185.56 \cdot (\Delta n)^3 \quad (6.5.1)$$

6.6. Calorimetría diferencial de barrido (CDB)

La calorimetría diferencial de barrido (CDB) 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 y biomembranas. Se han realizado gran cantidad de estudios de CDB de la desnaturización térmica de proteínas (Privalov y Khechinashvili, 1974; Mabrey *et al.*, 1978; Bach y Chapman,

1980; Privalov y Potekhin, 1986; Privalov *et al.*, 1989; Freire *et al.*, 1990; Martinez *et al.*, 1994; Martinez *et al.*, 1995; Sanchez-Ruiz, 1995; Filimonov *et al.*, 1999).

Un calorímetro diferencial de barrido registra la capacidad calorífica aparente de una disolución de macromolécula en función de la temperatura, obteniéndose así un *termograma*. Su posterior análisis utilizando modelos basados en la Termodinámica del equilibrio puede proporcionar la caracterización termodinámica completa del proceso de desplegamiento térmico de una proteína. Esto permite obtener la capacidad calorífica parcial absoluta del sistema en función de la temperatura, así como los parámetros termodinámicos asociados, tales como el cambio de entalpía, ΔH_D , de entropía, ΔS_D , de energía de Gibbs, ΔG_D , y de capacidad calorífica, ΔC_p , correspondientes a la transición inducida por la 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 de equilibrio más simple que existe para el análisis del desplegamiento térmico de proteínas es el modelo de dos estados, el cual supone que las proteínas sólo pueden encontrarse en estado nativo o en estado desplegado. Otros modelos pueden suponer la existencia de intermedios de plegamiento significativamente poblados y/o estados con diferentes grados de asociación.

6.6.1. El calorímetro diferencial de barrido

Para la realización de todos los experimentos de CDB de esta Memoria se ha empleado el microcalorímetro VP-DSC, de MICROCAL INC. (Plotnikov *et al.*, 1997), 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 lo más parecidas posible que contienen la disolución de la muestra objeto del estudio y una disolución de referencia.

6.6.2. Realización del experimento calorimétrico

El experimento calorimétrico se inicia llenando ambas células con el tampón en el cual va disuelta la proteína y se registra un barrido de temperatura, que corresponde a lo que se denomina *línea base instrumental*. Se pueden registrar varios barridos sucesivos 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 y se realiza el primer barrido con la proteína. Posteriormente puede realizarse un segundo barrido sin extraer la proteína del calorímetro para comprobar la reversibilidad del proceso de desplegamiento.

6.6.2.1. Cálculo de la capacidad calorífica molar parcial de la proteína

Los ficheros generados por el programa de adquisición de datos del calorímetro contienen el valor de una señal de voltaje (en mV) proporcional a las diferencias de capacidad calorífica entre las células en función de la temperatura. Los ficheros se analizan con el programa *MicroCal Origin for DSC v 4.1* en el que se lleva a cabo de forma automática la conversión de la señal de unidades de mV a unidades típicas de capacidad calorífica (en $\text{cal}\cdot\text{°C}^{-1}$ o $\text{J}\cdot\text{°C}^{-1}$), así como la corrección del tiempo de respuesta del calorímetro (corrección dinámica del termograma), basándose en la última calibración realizada y en el tiempo de respuesta preseleccionado. Una vez aquí se resta la línea base instrumental del termograma obtenido con la muestra de proteína y se normalizan los datos según la concentración de proteína y el volumen

efectivo de la célula, expresando la diferencia de capacidad calorífica aparente entre la disolución de macromolécula y el disolvente, $\Delta C_{p,ap}(T)$, en términos de capacidad calorífica molar en $\text{kJ}\cdot\text{K}^{-1}\cdot\text{mol}^{-1}$. Finalmente, la capacidad calorífica molar parcial de la proteína se obtiene según (Privalov y Khechinashvili, 1974):

$$C_p = \left[C_{p,s}(T) \cdot \frac{V_p(T)}{V_s(T)} M_p + \Delta C_{p,ap}(T) \right] \quad (6.6.1)$$

donde $C_{p,s}(T)$ es la capacidad calorífica parcial del disolvente, $V_p(T)$ y $V_s(T)$ son los correspondientes volúmenes específicos parciales de la proteína en disolución y del disolvente respectivamente y M_p es la masa molecular de la proteína. El volumen específico parcial que se usa generalmente como promedio para las proteínas globulares es $V_p=0.73 \text{ mL}\cdot\text{g}^{-1}$ (Makhatadze y Privalov, 1990).

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

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

6.6.2.2.1. 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 equilibrio:



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

$$K_D = \frac{[D]}{[N]} \quad (6.6.2)$$

La función de partición del sistema se define como:

$$q = 1 + K_D \quad (6.6.3)$$

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

$$f_N = \frac{1}{q} \quad (6.6.4)$$

$$f_D = \frac{K_D}{q} \quad (6.6.5)$$

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

$$\langle H \rangle = f_N \cdot H_N + f_D \cdot H_D = H_N + f_D \cdot \Delta H_D = H_N + \frac{K_D}{1+K_D} \cdot \Delta H_D \quad (6.6.6)$$

donde ΔH_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:

$$C_p = \left[\frac{\partial \langle H \rangle}{\partial T} \right]_P = C_{p,N} + \frac{\Delta H_D}{(1+K_D)^2} \left[\frac{\partial K_D}{\partial T} \right]_P + \frac{K_D}{(1+K_D)} \left[\frac{\partial \Delta H_D}{\partial T} \right]_P \quad (6.6.7)$$

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

$$\frac{\partial \Delta H_D}{\partial T} = \Delta C_p \quad (6.6.8)$$

$$\frac{\partial \Delta S_D}{\partial T} = \frac{\Delta C_p}{T} \quad (6.6.9)$$

$$K_D = \exp\left(-\frac{\Delta G_D}{T}\right) \quad (6.6.10)$$

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

Así, se obtienen las siguientes ecuaciones para $C_{p,D}$, $C_{p,N}$ y ΔC_p :

$$C_{p,D} = a + b \cdot T + c \cdot T^2 \quad (6.6.11)$$

$$C_{p,N} = d + e \cdot T \quad (6.6.12)$$

$$\Delta C_p = (a - d) + (b - e) \cdot T + c \cdot T^2 \quad (6.6.13)$$

Para obtener las dependencias de ΔH_D y ΔS_D con la temperatura integramos las ecuaciones 6.5.10 y 6.5.11:

$$\Delta H_D(T) = \Delta H_m + \int_{T_m}^T \Delta C_p \cdot dT \quad (6.6.14)$$

$$\Delta S_D(T) = \Delta S_m + \int_{T_m}^T \frac{\Delta C_p}{T} \cdot dT \quad (6.6.15)$$

donde ΔH_m y ΔS_m son los incrementos de entalpía y entropía de desplegamiento a la temperatura de la transición (T_m). Esa temperatura de la transición es aquella a la que f_D se hace igual a 0.5, ΔG_D se anula y por lo tanto K_D se hace igual a 1. En ese punto, el cambio de entropía a la temperatura T_m , ΔS_m , se puede obtener según la ecuación siguiente:

$$\Delta S_m = \frac{\Delta H_m}{T_m} \quad (6.6.16)$$

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

$$\Delta H_D = \Delta H_m + (a-d) \cdot (T - T_m) + \frac{(b-e)}{2} \cdot (T^2 - T_m^2) + \frac{1}{3} \cdot c \cdot (T^3 - T_m^3) \quad (6.6.17)$$

$$\Delta S_D = \frac{\Delta H_m}{T_m} + (a-d) \cdot \ln\left(\frac{T}{T_m}\right) + (b-e) \cdot (T - T_m) + \frac{1}{2} \cdot c \cdot (T^2 - T_m^2) \quad (6.6.18)$$

$$\Delta G_D = \Delta H_D - T \cdot \Delta S_D \quad (6.6.19)$$

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

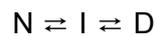
$$C_p = C_{p,N} + \frac{(\Delta H_D)^2}{RT^2} \cdot \frac{K_D}{(1+K)^2} + f_D \cdot \Delta C_p \quad (6.6.20)$$

Los ajustes por mínimos cuadrados no lineales de las curvas de C_p se han realizado usando las ecuaciones del modelo de dos estados, utilizando el programa Origin 6.1 (*Microcal Software Inc.*).

La ecuación empleada para los ajustes fue la 6.6.20 pero antes es preciso definir todas las magnitudes de las que depende C_p en dicha ecuación, en función de la temperatura. De esta manera la C_p queda expresada en función de la temperatura y una serie de parámetros de ajuste adecuados. De los ajustes obtenemos los valores de T_m , ΔH_m y ΔC_p .

6.6.2.2.2. Modelo de equilibrio de tres estados

En ocasiones existen transiciones calorimétricas reversibles que no pueden ser explicadas con un modelo tan sencillo y necesitamos usar modelos de equilibrio multiestados. En nuestro caso, como se describe en la Sección 4.6 del Capítulo 4, hemos utilizado un modelo de equilibrio de tres estados en el que consideramos la existencia de un estado intermedio de plegamiento significativamente poblado (I), además del estado nativo (N) y el desplegado (D). Así, el esquema del proceso de desplegamiento es el siguiente:



Las constantes de equilibrio K_1 y K_2 para cada una de las transiciones se definen como:

$$K_1 = \frac{[I]}{[N]} \quad (6.6.21)$$

$$K_2 = \frac{[D]}{[I]} \quad (6.6.22)$$

La función de partición del sistema en función de estas constantes de equilibrio es:

$$q = 1 + K_1 + K_1 K_2 \quad (6.6.23)$$

Usando las ecuaciones anteriores las poblaciones para cada estado se pueden calcular como:

$$f_N = \frac{1}{q} \quad (6.6.24)$$

$$f_I = \frac{K_1}{q} \quad (6.6.25)$$

$$f_D = \frac{K_1 K_2}{q} \quad (6.6.26)$$

La entalpía molar parcial para un sistema de tres estados se expresa como:

$$\langle H \rangle = f_N \cdot H_N + f_I \cdot H_I + f_D \cdot H_D = H_N + \langle \Delta H \rangle \quad (6.6.27)$$

donde H_N , H_I y H_D son las entalpías correspondientes a cada estado y $\langle \Delta H \rangle$ es el cambio de entalpía del sistema. Tomando como referencia el estado nativo, nos queda:

$$\langle \Delta H \rangle = \langle H \rangle - H_N = f_I \cdot \Delta H_1 + f_D (\Delta H_1 + \Delta H_2) \quad (6.6.28)$$

Aquí, ΔH_1 y ΔH_2 representan los cambios de entalpía correspondientes a la primera y a la segunda transición respectivamente. Derivando la ecuación 6.6.28 respecto de la temperatura, la capacidad calorífica del sistema será:

$$\langle C_p \rangle = \langle \Delta C_p^{\text{int}} \rangle + \langle \Delta C_p^{\text{exc}} \rangle \quad (6.6.29)$$

donde $\langle \Delta C_p^{\text{int}} \rangle$ corresponde a la capacidad calorífica química o interna de la proteína en disolución, mientras que $\langle \Delta C_p^{\text{exc}} \rangle$ refleja el exceso de capacidad calorífica producido por la reacción de desplegamiento:

$$\langle \Delta C_p^{\text{int}} \rangle = C_{p,N} + f_I \cdot \Delta C_{p,1} + f_D \cdot (\Delta C_{p,1} + \Delta C_{p,2}) \quad (6.6.30)$$

$$\langle \Delta C_p^{\text{exc}} \rangle = \frac{\langle \Delta H^2 \rangle - \langle \Delta H \rangle^2}{RT^2} \quad (6.6.31)$$

El cambio de capacidad calorífica para cada una de las transiciones se define como:

$$\Delta C_{p,1} = C_{p,I} - C_{p,N} \quad (6.6.32)$$

$$\Delta C_{p,2} = C_{p,D} - C_{p,I} \quad (6.6.33)$$

donde la dependencia de la capacidad calorífica con la temperatura del estado nativo se puede aproximar a una función lineal, mientras que para los estados I y D tenemos un polinomio de segundo grado:

$$C_{p,N} = a + bT \quad (6.6.34)$$

$$C_{p,I} = c + dT + eT^2 \quad (6.6.35)$$

$$C_{p,D} = f + gT + hT^2 \quad (6.6.36)$$

Finalmente, las constantes de equilibrio K_1 y K_2 , a cada temperatura se pueden calcular como:

$$K_1 = \exp\left(-\frac{\Delta G_1}{RT}\right) = \exp\left(-\frac{\Delta H_1}{RT} + \frac{\Delta S_1}{R}\right) \quad (6.6.37)$$

$$K_2 = \exp\left(-\frac{\Delta G_2}{RT}\right) = \exp\left(-\frac{\Delta H_2}{RT} + \frac{\Delta S_2}{R}\right) \quad (6.6.38)$$

Las variaciones de entalpía y entropía con la temperatura vienen dadas por:

$$\Delta H_1 = \Delta H_{m,1} + \int_{T_{m,1}}^T \Delta C_{p,1} \cdot dT \quad (6.6.39)$$

$$\Delta H_2 = \Delta H_{m,2} + \int_{T_{m,2}}^T \Delta C_{p,2} \cdot dT \quad (6.6.40)$$

$$\Delta S_1 = \Delta S_{m,1} + \int_{T_{m,1}}^T \frac{\Delta C_{p,1}}{T} \cdot dT \quad (6.6.41)$$

$$\Delta S_2 = \Delta S_{m,2} + \int_{T_{m,2}}^T \frac{\Delta C_{p,2}}{T} \cdot dT \quad (6.6.42)$$

En estas ecuaciones $\Delta H_{m,1}$, $\Delta H_{m,2}$, $\Delta S_{m,1}$ y $\Delta S_{m,2}$ representan las magnitudes correspondientes a las temperaturas $T_{m,1}$ y $T_{m,2}$, respectivamente.

Para los análisis recogidos en esta Memoria en los que hemos aplicado este modelo de equilibrio, hemos tomado como $T_{m,2}$ a la temperatura de la transición correspondiente al desplegamiento del permutante circular S19P20s a cada pH, que es la temperatura en la que: $\Delta S_{m,2} = \Delta H_{m,2} / T_{m,2}$. En el caso de $T_{m,1}$ hemos elegido como temperatura de referencia de la transición 25°C. En este caso, $\Delta S_{m,1}$ viene dada por:

$$\Delta S_{m,1} = \frac{\Delta H_{m,1}}{T_{m,1}} + R \ln K_{m,1} \quad (6.6.43)$$

donde $K_{m,1}$ es la constante de equilibrio para la primera transición a la temperatura $T_{m,1}$.

Este conjunto de ecuaciones nos permite hacer un ajuste global de las curvas de capacidad calorífica de cada quimera medida a distintos pHs (ver Figura 4.6.1 en el Capítulo 4). Para simplificar el procedimiento de ajuste y reducir el número de parámetros ajustables, hemos hecho una serie de suposiciones y aproximaciones:

- 1) En el estado intermedio, los 12 residuos del extremo C-terminal en los que se encuentra la secuencia del péptido, están completamente desordenados y expuestos al disolvente, por lo que la extensión de la cadena no influye en la termodinámica del desplegamiento del permutante S19P20s. Así, para la transición entre los estado I y D hemos considerado las mismas magnitudes termodinámicas que para el desplegamiento de S19P20s, lo que nos ha permitido fijar todos los parámetros de esta transición para cada pH (Martinez *et al.*, 1999).
- 2) Los coeficientes de $C_{p,D}$ han sido calculados a partir de la secuencia aminoacídica (Makhatadze y Privalov, 1990). Los coeficientes para $C_{p,I}$ los hemos calculado sumando las funciones de capacidad calorífica del estado nativo para el permutante (Martinez *et al.*, 1999) y la capacidad calorífica de la extensión de la cadena. Ambas funciones han sido fijadas en el análisis global.
- 3) Hemos considerado independientes del pH las diferencias de entalpía y de capacidad calorífica entre los estados N e I, lo que nos ha permitido definir ambos parámetros de ajuste comunes para todos los valores de pH.

Con estas aproximaciones, los parámetros de ajuste global del conjunto de curvas de C_p para cada proteína quimérica a diferentes valores de pH, se reducen a la constante de equilibrio, $K_{m,1}$, para cada pH, al cambio de entalpía, $\Delta H_{m,1}$, y al cambio de capacidad calorífica, $\Delta C_{p,1}$, a 25°C de la primera transición, además de los coeficientes que definen la dependencia lineal de $C_{p,N}$ con la temperatura.

6.7. Dicroísmo circular (DC)

La espectroscopia de dicroísmo circular (DC) es una de las técnicas más empleadas en el estudio de las propiedades conformacionales de proteínas y ácidos nucleicos en disolución (Schmid, F.X., 1989; Johnson, W. C. J., 1990). El fenómeno del dicroísmo circular se basa en la diferencia de absorción de la luz circularmente polarizada hacia la derecha e izquierda por parte de las moléculas asimétricas que son óptimamente activas, también llamadas quirales. Cada uno de los tipos de luz circularmente polarizada (a derecha e izquierda) sigue la ley de Lambert-Beer. Según esto, un espectro de DC para una muestra de un compuesto dado puede definirse inicialmente mediante la diferencia entre las absorbancias que presenta para las dos formas de luz polarizada, en función de la longitud de onda, λ :

$$\Delta Abs(\lambda) = Abs_I(\lambda) - Abs_D(\lambda) = [\varepsilon_I(\lambda) - \varepsilon_D(\lambda)]c \cdot l \quad (6.7.1)$$

$$\Delta Abs(\lambda) = \Delta \varepsilon \cdot c \cdot l \quad (6.7.2)$$

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

Las bandas de un espectro de DC pueden ser negativas o positivas, dependiendo de qué tipo de luz se absorba más y sólo se dan a longitudes de onda correspondientes a las bandas de absorción. Actualmente, los instrumentos que se utilizan para realizar estudios mediante DC reciben el nombre de espectropolarímetros y miden de forma directa la diferencia de absorbancia $\Delta Abs(\lambda)$. Cuando la luz polarizada plana incide sobre una muestra que exhibe DC, las dos componentes circularmente polarizadas que la componen se absorben en diferente magnitud. El DC es proporcional a la excentricidad de la elipse que caracteriza la luz polarizada resultante. La elipticidad, θ , es una medida angular de esta excentricidad y constituye la forma original para detectar la medida de DC. La mayoría de los instrumentos comerciales están calibrados en unidades de elipticidad (θ). Generalmente los espectros de DC se expresan en términos de elipticidad molar $[\theta]$ en $\text{grados} \cdot \text{dmol}^{-1} \cdot \text{cm}^2$. Su relación con la diferencia entre los coeficientes de extinción molar viene dada por la relación:

$$[\theta] = 3298 \cdot \Delta \varepsilon \quad (6.7.3)$$

que incluye un factor de 100 debido a la conversión de la concentración en mol/L a $\text{dmol} \cdot \text{cm}^{-3}$.

En todos los experimentos de DC recogidos en esta Memoria las medidas de elipticidad se han normalizado generalmente a unidades de elipticidad molar media por residuo ($[\theta]$) en $(\text{grado} \cdot \text{dmol}^{-1} \cdot \text{cm}^2)$, para lo cual se utiliza la siguiente ecuación:

$$[\theta] = \frac{\theta}{\frac{C}{Pm} \cdot l \cdot 10 \cdot n} \quad (6.7.4)$$

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

Mediante DC podemos obtener mucha información acerca de las propiedades conformacionales de las moléculas biológicas, más concretamente:

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

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

Los espectros de DC en esta región son muy sensibles a los cambios de estructura secundaria de las proteínas, ya que la contribución de la estructura terciaria a la señal de DC en estas longitudes de onda es despreciable. A estas longitudes de onda el cromóforo es el enlace peptídico. Aunque el enlace peptídico no es intrínsecamente una estructura quiral, la interacción entre grupos peptídicos adyacentes dispuestos en determinadas orientaciones hace que las proteínas presenten actividad óptica en la región del ultravioleta lejano. Cada elemento principal de estructura secundaria de proteínas: hélices α , láminas β , giros β y estructuras aleatorias presenta un espectro de DC característico (Compton y Johnson, 1986). En la Figura 6.7.1 se representan los espectros típicos de cada uno de ellos y se observa que estos espectros difieren entre sí lo suficiente como para determinar el contenido de estas estructuras en la proteína. Las hélices α presentan un espectro de DC muy característico en la región del ultravioleta lejano con un doble mínimo a 222 y a 208-210 nm y un máximo más intenso a 191-193 nm. Las proteínas con estructura β suelen presentar un único pico negativo y un único pico positivo cuyas intensidades son menores que las de las hélices. El pico positivo es de mayor intensidad que el negativo y aparece entre 190 y 200 nm. El mínimo se encuentra normalmente entre 210 y 225 nm, aunque se han descrito algunos ejemplos (Manavalan y Johnson, 1983; Wu, J. *et al.*, 1992) de proteínas que, aunque se ha determinado por RMN que son todo tipo β , sus espectros de DC tienen una banda negativa intensa alrededor de 200 nm, similar a los péptidos con estructura aleatoria. La explicación que se ha dado en estos casos es que estas proteínas tienen hojas β cortas y bastante irregulares.

Los péptidos y proteínas con estructura no regular (aleatoria), los bucles de las conformaciones nativas y, en general, las conformaciones desplegadas presentan habitualmente una banda negativa fuerte cerca de 200 nm y algunas bandas débiles entre 220 y 230 nm que pueden ser tanto positivas como negativas.

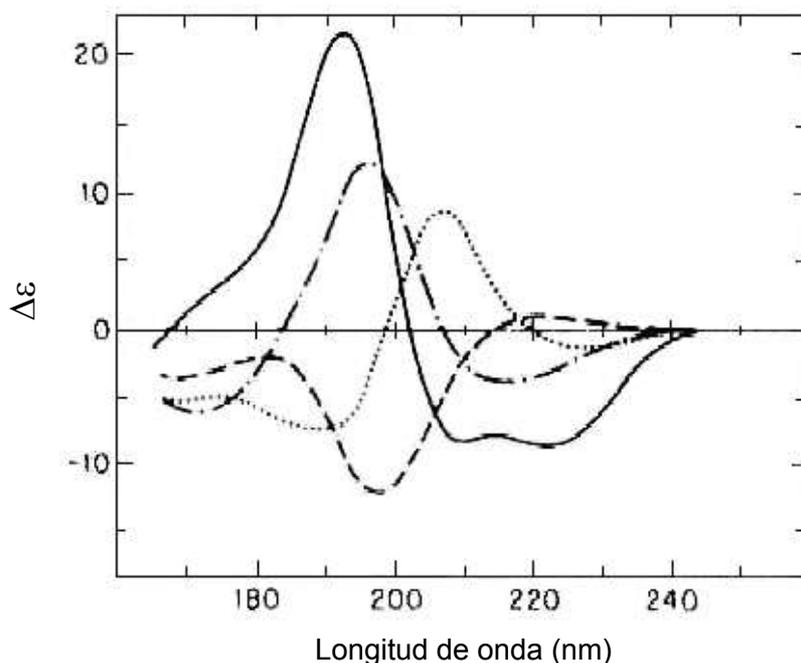


Figura 6.7.1. Espectros representativos de los cuatro tipos principales de estructuras secundarias desde 178 hasta 260 nm: Hélice alfa (—), hoja beta (—●—●—●), giro beta (●●●●●) y otras estructuras aleatorias (-----), tomado de la referencia (Brahms y Brahms, 1980).

Hay que recordar, sin embargo, que en el ultravioleta lejano también absorben los aminoácidos aromáticos y que, aunque sus cadenas laterales no son quirales, se comportan como tales cuando así lo propicia su entorno proteico. Por esta razón el espectro de algunas proteínas aparece dominado por contribuciones peculiares debidas a estos residuos. Con esta salvedad, la actividad óptica de los enlaces peptídicos es de gran utilidad para conocer el contenido en estructura secundaria de las proteínas. Así, cada proteína tiene un espectro de DC en el ultravioleta lejano que es, aproximadamente, una combinación lineal de los espectros característicos de las estructuras secundarias y no regulares que contiene (Woody, 1994).

6.7.2. Dicroísmo circular en la zona del UV-cercano (250-350 nm)

A estas longitudes de onda los cromóforos son los aminoácidos aromáticos (Trp, Tyr y Phe) 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 de DC es pequeña, debido a su bajo coeficiente de extinción; además, es poco sensible a las alteraciones de su ambiente. Presenta una banda de absorción en la región 250-270 nm. La tirosina tiene otra más intensa a 276 nm y un hombro a 283 nm (banda que puede desplazarse hacia el rojo debido a la formación de un puente de hidrógeno por el grupo fenólico). El triptófano presenta la banda de absorción aún más intensa centrada a 282 nm (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 UV-lejano, las absorciones de las cadenas laterales de estos residuos aromáticos no se mezclan con las del grupo peptídico (que no absorbe en esta región).

El espectro de DC de una proteína en el UV-cercano no puede analizarse cuantitativamente, ya que depende del entorno asimétrico concreto de cada cromóforo y, por tanto, constituye una pequeña huella dactilar de la conformación nativa cuya pérdida nos indicará que han ocurrido cambios conformacionales apreciables.

6.7.3. Obtención de espectros de dicroísmo circular

Los experimentos se han llevado a cabo en un espectropolarímetro JASCO modelo J-715 equipado con un control de temperatura del compartimento de células. En varias tesis doctorales del grupo de investigación Biofísica y Biotecnología Molecular de la Universidad de Granada (Sadqi, 2000; Cobos, 2002) se han explicado ampliamente todos los fundamentos y características del instrumento, por lo que nos limitamos a recomendar estas referencias para un conocimiento más detallado. Una vez preparada la disolución de proteína, como se ha indicado en la Sección 6.5, se introduce en una cubeta de cuarzo de paso de luz adecuado. Éste se escoge en función de la concentración de la muestra y del intervalo de longitud de onda seleccionado para las medidas para evitar la saturación de la señal.

Para la caracterización preliminar de la quimera SPCp41 hemos registrado los espectros de DC en la región del ultravioleta lejano (entre 190 y 250 nm) y en la del ultravioleta cercano (entre 250 y 320 nm) usando cubetas de 1 y 5 mm de paso de luz, respectivamente. Todos los espectros se han registrado a 25°C. La concentración de la muestra fue de 0.2-0.25 mg/mL para el ultravioleta lejano y de 2-2.5 mg/mL para el cercano. Se registró asimismo un espectro con el tampón de diálisis que se restó del espectro de cada proteína, para eliminar las contribuciones del tampón e instrumentales. Para cada uno de los espectros finales se promediaron cuatro registros consecutivos (ver Figuras 2.3.4 y 2.3.5 del Capítulo 2).

También se han realizado los espectros de DC en el UV-lejano y cercano de las formas mutantes de la quimera, con objeto de observar posibles cambios en la estructura correspondiente al estado nativo de la quimera SPCp41 producidos por las mutaciones. Los experimentos en el ultravioleta lejano se registraron desde 250 hasta 190 nm con una cubeta de 1 mm de paso de luz. Para una mayor exactitud de la medida de la señal de elipticidad a 228 nm, estos espectros se volvieron a registrar desde 250 a 215 nm utilizando una concentración de proteína de $\approx 0.25\text{-}0.50\text{ mg}\cdot\text{mL}^{-1}$. Para el ultravioleta cercano los espectros se registraron desde 320 hasta 250 nm usando una cubeta de 5 mm de paso de luz y concentraciones de proteína de $\approx 1.5\text{-}2.5\text{ mg}\cdot\text{mL}^{-1}$. Se registro así mismo una línea base con tampón de diálisis que se restó del espectro de cada proteína. Para cada uno de los espectros finales se promediaron veinte registros consecutivos (ver Figuras 4.2.3 y 4.2.5).

6.8. Espectroscopia de fluorescencia

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. Algunas proteínas que poseen otros grupos cromóforos como cofactores presentan también absorción en la región visible del espectro. En el caso de las proteínas los tres residuos 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; Demchenko, 1981; Schmid, F.X., 1989; Permyakov, 1993). Estos aminoácidos son relativamente poco abundantes en proteínas, lo cual facilita la

interpretación de los datos espectrales desde un punto de vista estructural. Una característica importante de la fluorescencia de proteínas es la alta sensibilidad del triptófano a los cambios en su entorno próximo, de tal manera que se pueden observar cambios en los espectros de emisión de triptófano como respuesta a transiciones conformacionales de la proteína, a fenómenos de asociación de subunidades, a procesos de unión de ligandos y a la propia desnaturalización de la proteína. Todos estos procesos pueden afectar el entorno del anillo indólico del triptófano y, como consecuencia, provocar cambios en su fluorescencia intrínseca, de ahí su utilidad.

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. Además, por el hecho de absorber a longitudes de onda mayores, a menudo la energía absorbida por la fenilalanina y tirosina se transfiere a los residuos 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/o de su entorno local.

La emisión de fluorescencia de proteínas se excita por lo general a un máximo de absorción cercano a 280 nm o incluso a longitudes de onda mayores, por lo que la fenilalanina no se excita normalmente en un experimento de emisión convencional (Chen, R. F., 1967). Puede tener lugar una transferencia de energía desde la tirosina hasta el triptófano mediante procesos no radiativos, con lo que la emisión de las tirosinas también se puede ver afectada (Cantor y Schimmel, 1980; Penzer, 1980; Eftink, 1991), pero para que esto ocurra es preciso que ambos residuos se encuentren en la misma proteína y orientados de forma favorable en el espacio. Si se excita a longitudes de onda superiores a 295 nm, la absorción se debe principalmente al triptófano, que puede ser excitado de forma selectiva a longitudes de onda entre 295 y 305 nm.

La fluorescencia puede emplearse para el estudio de procesos de unión de ligandos a las proteínas. En algunas ocasiones dichos ligandos poseen fluorescencia propia, que cambia cuando se unen a una proteína (Daniel y Weber, 1966; Anderson y Weber, 1969; Condie y Quay, 1983). La unión de ligandos a proteínas también puede provocar cambios en la estructura tridimensional de las mismas. Si esos cambios estructurales suponen un efecto en el entorno de un fluoróforo intrínseco o extrínseco a la proteína, esto supondrá una variación en el espectro de fluorescencia de la misma. Tales cambios de fluorescencia a una longitud de onda determinada, pueden emplearse para obtener la constante de disociación (K_d) de la proteína para ese ligando, donde K_d es una medida de la afinidad de la proteína por el ligando.

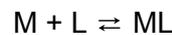
6.8.1. Análisis de la unión proteína-ligando mediante espectroscopia de fluorescencia

Las medidas de fluorescencia se han hecho en un instrumento denominado espectrofluorímetro de la casa comercial PERKIN-ELMER modelo LS-55B.

Hemos realizado una serie de experimentos de titulación por adición de volúmenes variables de una disolución madre de péptido p41 sobre una disolución de

permutante circular S19P20s, midiéndose la intensidad de fluorescencia tras cada adición, con objeto de compararlo con el espectro de fluorescencia de la quimera SPCp41. Así, se partió de una disolución inicial de proteína, dializada frente al tampón correspondiente, de una concentración en torno a 25 μM . Por otro lado se prepara una disolución madre de péptido p41 en el mismo tampón a una concentración de aproximadamente 3 mM. Se utilizó una cubeta de cuarzo de 3 mm de pequeño volumen para la muestra y en ella se introdujeron 100 μL de la disolución madre de proteína, registrándose su espectro de fluorescencia. Sobre esta disolución se adicionaron volúmenes crecientes de la disolución madre de p41 y se registraron los espectros de emisión de fluorescencia después de cada adición, a las diferentes concentraciones de p41 (desde 0 hasta 2 mM aproximadamente) (ver Figura 2.3.1 del Capítulo 2).

El equilibrio de unión que se establece entre la proteína y el ligando se describe:



siendo la constante de unión, K_d y la fracción de saturación de proteína, θ :

$$K_d = \frac{[M][L]}{[ML]} \quad (6.8.1)$$

$$\theta = \frac{[ML]}{[M] + [ML]} \quad (6.8.2)$$

Es posible escribir la fracción de saturación en función de la constante de disociación, siendo M_T la concentración de proteína total en la muestra y L_T la concentración de ligando total en la muestra:

$$\theta = \frac{(K_d + M_T + L_T) \pm \sqrt{(K_d + M_T + L_T)^2 - 4 \cdot M_T \cdot L_T}}{2 \cdot M_T} \quad (6.8.3)$$

La intensidad de fluorescencia normalizada por la concentración de proteína total en la muestra, M_T , en cada medida del experimento de titulación nos queda como:

$$F_{i,n} = \frac{F_i}{M_T} = F_M + (F_{ML} - F_M) \cdot \theta_i = F_M + \Delta F \cdot \theta_i \quad (6.8.4)$$

También se registraron espectros de fluorescencia de cada uno de los mutantes de la quimera SPCp41 para caracterizar los cambios estructurales que producían las mutaciones correspondientes. Todos estos espectros fueron normalizados por la concentración de proteína (ver Figura 4.2.1 del Capítulo 4).

Todos los espectros de emisión de fluorescencia de las diferentes muestras consideradas se registraron a 25°C entre 305 y 400 nm utilizando una longitud de onda de excitación de 298 nm, y las rendijas de excitación y emisión se dispusieron a 2.5 y 5 nm respectivamente, para evitar problemas de saturación de la señal.

6.9. Estructura tridimensional de proteínas mediante Resonancia Magnética Nuclear (RMN)

La Resonancia Magnética Nuclear (RMN) se ha consolidado en los últimos años como la técnica más poderosa para la determinación de la estructura tridimensional de biomoléculas en disolución. Mediante la utilización de la RMN homo- y hetero-nuclear, es posible la interpretación de los complejos espectros de RMN de proteínas y la obtención de los parámetros necesarios para la determinación de sus estructuras. Los estudios estructurales de proteínas por RMN constan de cuatro partes claramente diferenciadas. La primera es el registro de los espectros, la segunda consiste en la asignación de los mismos, en tercer lugar se extrae la información estructural de los espectros asignados y, finalmente, se procede a calcular las estructuras que son compatibles con esta información experimental. En la Figura 6.9.1 se muestra un esquema del proceso que se sigue en la determinación estructural de proteínas en disolución por RMN, y es el proceso que nosotros hemos seguido para obtener la estructura tridimensional de la quimera SPCp41, tal y como se describe en esta Sección. Los parámetros de RMN aportan información acerca de la estructura de las proteínas porque se encuentran relacionados con parámetros geométricos de las moléculas, y pueden traducirse en restricciones para el cálculo de la estructura tridimensional. La información que se obtiene de cada uno de estos parámetros se detalla en la Figura 6.9.1. En nuestro caso, los parámetros utilizados para el cálculo son el desplazamiento químico (δ) y el Efecto Overhauser Nuclear (NOE).

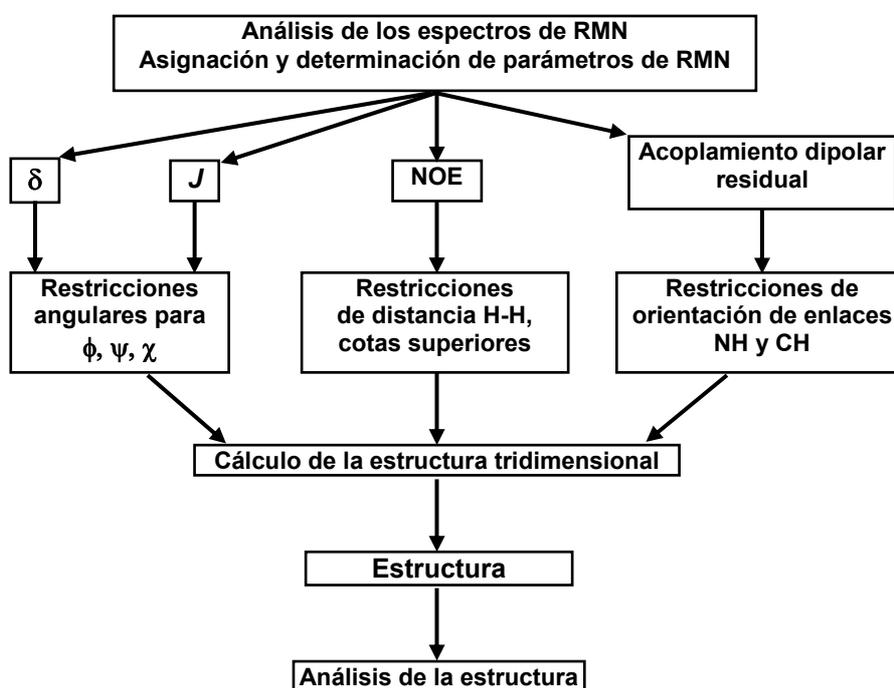


Figura 6.9.1. Sinopsis del procedimiento para la determinación de estructura de proteínas por RMN.

Desplazamiento químico: la frecuencia de resonancia de un núcleo, ν , depende de la intensidad del campo magnético B_0 , es decir, del espectrómetro empleado para la adquisición de los espectros de RMN, de acuerdo con las siguientes expresiones:

$$\nu = \gamma B_{\text{efectivo}} / 2\pi \quad (6.9.1)$$

$$B_{\text{efectivo}} = B_0 (1 - \sigma) \quad (6.9.2)$$

donde B_{efectivo} es el campo magnético real que experimenta el núcleo, γ es la constante giromagnética del núcleo, y σ es la constante de apantallamiento que hace referencia al fenómeno que se origina cuando la circulación de los electrones provoca un campo magnético inducido que se opone al principal (B_0). El desplazamiento químico δ , surge ante la necesidad de hacer comparables las posiciones de las señales de RMN en espectros obtenidos en distintos espectrómetros. Se expresa en ppm (partes por millón) y se define como $\delta_{\text{núcleo}} = (\nu_{\text{núcleo}} - \nu_{\text{referencia}}) 10^6 / \nu_0$, siendo $\nu_{\text{núcleo}}$ la frecuencia del núcleo observado, $\nu_{\text{referencia}}$ la frecuencia de resonancia de un núcleo elegido como referencia y ν_0 la frecuencia del núcleo considerado (^1H , ^{13}C o ^{15}N) en el espectrómetro, todas ellas expresadas en hercios (Hz). Las referencias más comunes en disolución acuosa son el TSP (3-trimetilsililpropionato de sodio) y el DSS (2,2-dimetil-2-silpentano-5 sulfonato de sodio).

La dispersión de δ en los espectros de RMN de una proteína es un notable indicador del estado nativo o desnaturalizado de las proteínas debido a que el δ de cada núcleo depende de su entorno químico, siendo muy sensible a la conformación de la proteína. En proteínas desnaturalizadas cada núcleo tiene su señal de resonancia en una zona característica del espectro, mientras que las proteínas en estado nativo se caracterizan por una dispersión de las señales mucho mayor. Los desplazamientos químicos de los núcleos de cada residuo encontrados experimentalmente pueden utilizarse para la predicción de la estructura secundaria de la proteína a través del índice de desplazamiento químico (*Chemical-Shift Index*; CSI) (Wishart y Sykes, 1994). Este método consiste en comparar los desplazamientos químicos de los $^1\text{H}_\alpha$, $^{13}\text{C}_\alpha$, $^{13}\text{C}_\beta$ y ^{13}CO con los tabulados para esos núcleos en conformación desordenada y asignar un índice de +1, 0 o -1 según esas diferencias sean positivas o negativas y superen en magnitud ciertos umbrales previamente establecidos. Esta información permite predecir con poco margen de error la estructura secundaria de la proteína a lo largo de la cadena. Además, a partir de esta información obtenemos restricciones para los ángulos dihedros *phi* y *psi* de la cadena principal de la proteína que pueden utilizarse posteriormente para el cálculo estructural.

Efecto Overhauser Nuclear (NOE): es el parámetro de RMN que proporciona la información estructural más valiosa e imprescindible para la determinación de la estructura tridimensional de una proteína. La transferencia de magnetización por efecto NOE entre núcleos próximos depende de un factor de movilidad (el tiempo de correlación, τ_m) y de la inversa de la sexta potencia de la distancia internuclear ($\propto 1/r^6$). Esta dependencia hace que el efecto NOE entre dos protones sea observable si éstos se encuentran a una distancia inferior a 4.5-5.0 Å, puesto que por encima de esta distancia la intensidad del NOE se hace tan pequeña que no se detecta experimentalmente. De todas formas estas distancias son del orden de las distancias de enlace, por lo que son suficientes para nuestro propósito. La observación del efecto NOE entre protones alejados en la secuencia de la proteína indica que éstos se encuentran espacialmente próximos, lo que restringe en gran medida el espacio conformacional de la proteína. En el cálculo de estructuras tridimensionales de

proteínas, la intensidad de las conectividades NOE (conocidos habitualmente como NOEs) se traduce en cotas superiores (restricciones) para las distancias entre pares de protones.

Para poder hacer uso de toda esta información estructural es preciso realizar la asignación de todas las señales de resonancia observadas en los espectros de RMN. Los métodos a seguir dependen del tamaño de la proteína, siendo necesario el empleo de muestras etiquetadas con ^{15}N o doblemente etiquetadas ($^{15}\text{N}/^{13}\text{C}$) si se trata de proteínas grandes. En las secciones siguientes se dan detalles de los métodos existentes y se describe el protocolo seguido en esta Memoria.

6.9.1. Registro de los espectros de RMN

Los espectros utilizados para la asignación de las señales de RMN de la proteína SPCp41 se registraron a 30°C en un espectrómetro de RMN Bruker Avance 800 US-2 del Instituto de Química Física Rocasolano del Consejo Superior de Investigaciones Científicas (C.S.I.C) de Madrid, que opera a 800.13 MHz para el protón. Está equipado con una sonda triple de 5 mm inversa tipo TXI (^1H , ^{13}C , ^{15}N) con gradientes de campo en los ejes X, Y, Z. Los experimentos de triple resonancia para la asignación del mutante P74A de la quimera SPCp41 se registraron en el espectrómetro Varian VNS 600 del Centro de Instrumentación Científica (C.I.C.) de la Universidad de Granada, que opera a 600.25 MHz para el protón, equipado con una sonda PFG-XYZ triple-resonancia $^1\text{H}\{^{13}\text{C}, ^{15}\text{N}\}$. La temperatura de la sonda de RMN se calibró utilizando una muestra de metanol (Duerst y Merbach, 1965).

Tabla 6.9.1. Experimentos de RMN adquiridos para la quimera SPCp41.

Experimento	Muestra	Correlaciones	Información obtenida
2D			
^{15}N -HSQC	^{15}N	H-N	Asignación esqueleto
TOCSY	^1H	H-H	Asignación cadenas laterales
NOESY	^1H	H-H	Restricciones distancia
3D			
CBCA(CO)NH	^{15}N ^{13}C	H, N, C_α , C_β	Asignación esqueleto
CACBNH	^{15}N ^{13}C	H, N, C_α , C_β	Asignación esqueleto
HNCA	^{15}N ^{13}C	H, N, C_α	Asignación esqueleto
HN(CO)CA	^{15}N ^{13}C	H, N, C_α	Asignación esqueleto
HNCO	^{15}N ^{13}C	H, N, CO	Asignación esqueleto
HCCH-TOCSY	^{15}N ^{13}C	H, C, H	Asignación cadenas laterales
^{15}N -TOCSY-HSQC	^{15}N	H, N, H	Asignación cadenas laterales
^{15}N -NOESY-HSQC	^{15}N	H, N, H	Restricciones distancia

Tabla 6.9.2. *Parámetros de adquisición y procesado de los distintos experimentos.*

Experimento	Anchura espectral (ppm)	Número de puntos complejos	Número de barridos	^a Número total de puntos
¹⁵ N-HSQC	¹ H → 6.51 ¹⁵ N → 30	¹ H x ¹⁵ N 1024 x 64	8	1024 x 512
TOCSY	¹ H → 13 ¹ H → 13	¹ H x ¹ H 1024 x 256	40	2048 x 1024
NOESY	¹ H → 13 ¹ H → 13	¹ H x ¹ H 1024 x 512	64	2048 x 1024
CBCA(CO)NH	¹ H → 6.51 ¹⁵ N → 30 ¹³ C → 75	¹ H x ¹⁵ N x ¹³ C 1024 x 27 x 75	8	1024 x 128 x 512
CBCANH	¹ H → 6.51 ¹⁵ N → 30 ¹³ C → 75	¹ H x ¹⁵ N x ¹³ C 1024 x 27 x 75	8	1024 x 128 x 512
HNCA	¹ H → 6.51 ¹⁵ N → 30 ¹³ C → 75	¹ H x ¹⁵ N x ¹³ C 1024 x 27 x 75	8	1024 x 128 x 512
HN(CO)CA	¹ H → 6.51 ¹⁵ N → 30 ¹³ C → 75	¹ H x ¹⁵ N x ¹³ C 1024 x 27 x 75	8	1024 x 128 x 512
HNCO	¹ H → 4.88 ¹⁵ N → 30 ¹³ C → 14	¹ H x ¹⁵ N x ¹³ C 1024 x 27 x 75	8	768 x 128 x 256
HCCH-TOCSY	¹ H → 6.51 ¹³ C → 75 ¹ H → 7.5	¹ H x ¹ H x ¹³ C 1024 x 64 x 32	8	1024 x 256 x 512
¹⁵ N-TOCSY-HSQC	¹ H → 6.51 ¹⁵ N → 30 ¹ H → 13	¹ H x ¹⁵ N x ¹ H 1024 x 32 x 128	8	1024 x 256 x 512
¹⁵ N-NOESY-HSQC	¹ H → 6.51 ¹⁵ N → 30 ¹ H → 13	¹ H x ¹⁵ N x ¹ H 1024 x 32 x 128	8	1024 x 256 x 512

^aDespués del procesado.

La intensa señal de resonancia del agua, que aparece entre 4.6 y 4.8 ppm, constituye un problema en los espectros de RMN de disoluciones acuosas de proteínas pues resulta ser una importante interferencia en los espectros que podría enmascarar señales de la proteína. Así, la supresión de la señal del agua se realizó por presaturación selectiva de ésta o incluyendo una secuencia *WATERGATE* (Piotto *et al.*, 1992) justo antes de la secuencia de pulsos propiamente dicha de cada espectro. De esta forma la intensidad de la señal del agua se reduce considerablemente haciendo visibles el resto de las señales.

En la Tabla 6.9.1 se indican los espectros registrados con cada una de las muestras de SPCp41, así como la información obtenida de cada uno de ellos. Utilizando la muestra de la quimera doblemente etiquetada con ^{15}N y ^{13}C , obtenida como se ha descrito en el Sección 6.4.3 de este Capítulo, se registraron espectros tridimensionales CBCA(CO)NH (Geen y Freeman, 1991), CBCANH (Patt, 1992), HNCA (Clore y Gronenborn, 1991), HN(CO)CA (Emsley, 1994), HNCO (Kay, L. E. *et al.*, 1994) y HCCH-TOCSY (Kay, L. E. *et al.*, 1993). Utilizando muestras de proteína etiquetada con ^{15}N se registraron espectros ^{15}N -HSQC (Bodenhausen y Ruben, 1980), ^{15}N -TOCSY-HSQC y ^{15}N -NOESY-HSQC (Palmer *et al.*, 1992). Con muestras de la quimera SPCp41 sin etiquetar se registraron experimentos bidimensionales TOCSY (60 ms de tiempo de mezcla) (Bax y Davis, 1985) y NOESY (100 ms de tiempo de mezcla) (Kumar *et al.*, 1980). En la Tabla 6.9.2 se indican el número de puntos registrados para cada dimensión de los espectros y el número de barridos realizados.

6.9.2. Procesado y análisis de los espectros de RMN

Los espectros bidimensionales y tridimensionales se procesaron con el programa XWINNMR (Bruker Biospin, Karlsruhe, Alemania) o NMRPipe (Delaglio *et al.*, 1995), mientras que para la visualización y el análisis de los espectros se utilizaron los programas SPARKY (Goddard y Kneller) o NMRView (Johnson, B. A. y Blevins, 1994).

Para el procesado en algunas dimensiones se duplicó el número de puntos mediante llenado con ceros y, dependiendo del experimento, los espectros se filtraron utilizando funciones seno o seno cuadrado desplazadas. Para duplicar el número de puntos para las dimensiones de ^{13}C y ^{15}N se utilizó predicción lineal y llenado de ceros hasta alcanzar el número total de puntos indicado en la Tabla 6.9.2. Una vez realizada la transformada de Fourier de los espectros, se aplicó una corrección de línea base en todas las dimensiones.

6.9.3. Asignación de los espectros de RMN

La estrategia de asignación de espectros de proteínas consta de dos etapas: 1) identificación de las señales de RMN correspondientes a *sistemas de espín* independientes y clasificación por tipos de aminoácidos, 2) asignación secuencial que consiste en asignar los sistemas de espín previamente identificados a residuos específicos de la proteína. El concepto de *sistema de espín* es muy útil para que el análisis de los complejos espectros de ^1H -RMN de proteínas; las diferentes líneas espectrales observadas se pueden clasificar en conjuntos que se corresponden con sistemas completos de espín, que no son otra cosa sino conjuntos de espines acoplados escalarmenete. De esta manera, las señales de resonancia y la red de conectividades a través de acoplamientos escalares forman patrones característicos para cada tipo de aminoácido, que se conservan cuando se encuentran formando parte de proteínas. De este modo, los veinte aminoácidos presentes en las proteínas se agrupan en diversas categorías según el *sistema de espín* que formen sus diferentes protones, desde la glicina, el más sencillo, hasta los aromáticos, más complejos, y se clasifican según una nomenclatura estandarizada (Pople, 1959). Sólo la metionina y los aminoácidos que presenta una cadena lateral aromática (histidina, fenilalanina, tirosina y triptófano) se encuentran constituidos por más de un sistema de espín.

Se dispone de dos estrategias de asignación secuencial de proteínas, una basada en experimentos de correlación homonuclear ^1H - ^1H (Wüthrich, 1986), y la segunda en correlaciones heteronucleares de proteínas enriquecidas en ^{13}C y ^{15}N (Ikura *et al.*, 1990).

6.9.3.1. Asignación de proteínas mediante métodos homonucleares

En la asignación de proteínas pequeñas se identifican en primer lugar los sistemas de espín que corresponden a los distintos aminoácidos a partir de las señales observadas en el espectro 2D TOCSY. El experimento **TOCSY** (**T**otal **C**orrelation **S**pectroscop**Y**) (Bax y Davis, 1985) se basa en la transferencia de magnetización entre núcleos acoplados escalarmente. En este experimento se pueden observar señales de correlación entre todos los protones pertenecientes a un mismo sistema de espín, lo que le hace el más adecuado para la asignación e identificación de sistemas de espín en moléculas complejas que contienen múltiples sistemas de espín. En el espectro TOCSY todas las señales son de absorción y en fase y si el tiempo de mezcla del experimento es suficientemente largo, es posible correlacionar un determinado protón con todos los pertenecientes a su sistema de espín.

A continuación, se establecen correlaciones secuenciales a partir de los NOEs $d_{\text{NN}(i,i+1)}$, $d_{\text{aN}(i,i+1)}$, y $d_{\text{BN}(i,i+1)}$ entre protones de residuos contiguos en la secuencia de la proteína que pueden observarse en espectros 2D NOESY. El experimento **NOESY** (**N**uclear **O**verhauser **E**ffect **S**pectroscop**Y**) (Kumar *et al.*, 1980) se basa en la transferencia de la magnetización entre espines que están acoplados dipolarmente y, por tanto, se encuentran próximos en el espacio. El resultado de este experimento es un espectro donde las frecuencias de los distintos núcleos cercanos en el espacio están correlacionadas. En general, la observación de un NOE en el espectro indica que los núcleos implicados están separados una distancia menor de 5 Å. Los NOEs observados en proteínas correlacionan núcleos pertenecientes al mismo sistema de espín (intrarresiduo), núcleos de aminoácidos consecutivos en la secuencia de la proteína (secuenciales) y núcleos alejados en la secuencia pero próximos en el espacio (NOEs de medio y largo alcance). La intensidad, el número y el tipo de NOEs secuenciales que se observan en cada segmento de cadena peptídica dependen de la estructura secundaria, ya que las distancias entre los protones de residuos contiguos dependen de ella, pero siempre existe al menos una distancia secuencial menor de 3 Å que da lugar a un NOE intenso. La geometría de las hélices α da lugar a NOEs del tipo $d_{\text{Ni,Ni+1}}$, $d_{\text{Bi,Ni+1}}$, $d_{\text{ai,Ni+3}}$ y $d_{\text{ai,Ni+4}}$, mientras que los NOEs $d_{\text{ai,Ni+1}}$, $d_{\text{Ni,Nj+1}}$ y $d_{\text{ai,aj}}$ son característicos de láminas β , donde $d_{\text{Xi,Yi+1}}$ representa la distancia entre el átomo x del aminoácido i y el átomo y del aminoácido $i+1$, siendo i y j aminoácidos de dos hebras contiguas de la lámina. Los diferentes patrones de NOE secuenciales que caracterizan los distintos elementos de estructura secundaria se pueden usar como información adicional para verificar o ampliar el número de asignaciones.

6.9.3.2. Asignación de proteínas mediante métodos heteronucleares

La asignación de las resonancias protónicas, imprescindibles para cualquier estudio estructural por RMN, se ve facilitada en gran medida cuando estas señales se asocian con las de los heteronúcleos, ^{13}C y ^{15}N , a los que se encuentran directa o indirectamente unidos. Adicionalmente, los parámetros de los heteronúcleos contienen información estructural independiente y complementaria a la obtenida de los protones. En particular sus parámetros de relajación, como veremos en la siguiente Sección,

proporcionan un método apropiado para la caracterización de la dinámica interna de las moléculas estudiadas.

Para proteínas grandes se necesita el etiquetado uniforme con ^{15}N y ^{13}C . La conexión secuencial entre los sistemas de espín se basa en un conjunto de espectros 3D, que se basan en los acoplamientos escalares a través de uno o dos enlaces en los que participan los heteronúcleos. En las múltiples estrategias de asignación por métodos heteronucleares (distintos experimentos 3D proporcionan conexiones secuenciales a través de distintos enlaces; ver Tabla 6.9.3), primero se lleva a cabo la asignación de los núcleos del esqueleto peptídico, ^1HN , H^{15}N , ^{13}CO , ^1H , $^{13}\text{C}_\alpha$ y $^1\text{H}_\alpha$, mediante el análisis conjunto del ^{15}N -HSQC y de los espectros de triple resonancia CBCA(CO)NH, CBCANH, HNCA, HN(CO)CA y HNCO, como veremos a continuación, y después se completa la asignación de las cadenas laterales de los distintos aminoácidos. Los experimentos 3D se analizan examinando proyecciones 2D.

Tabla 6.9.3. Conectividades intrarresiduales y secuenciales observadas en algunos espectros 3D útiles para la asignación secuencial a través de enlaces.

Experimento	Conectividad intrarresidual	Conectividad interresidual
CBCANH	$^{13}\text{C}_{\beta(i)}-^{13}\text{C}_{\alpha(i)}-^{15}\text{N}_{(i)}-^1\text{H}_{\text{N}(i)}$	$^{13}\text{C}_{\beta(i)}-^{13}\text{C}_{\alpha(i)}-^{15}\text{N}_{(i+1)}-^1\text{H}_{\text{N}(i+1)}$
CBCA(CO)NH	—	$^{13}\text{C}_{\beta(i)}-^{13}\text{C}_{\alpha(i)}-^{15}\text{N}_{(i+1)}-^1\text{H}_{\text{N}(i+1)}$
HNCA	$^1\text{H}_{\text{N}(i)}-^{15}\text{N}_{(i)}-^{13}\text{C}_{\alpha(i)}$	$^1\text{H}_{\text{N}(i)}-^{15}\text{N}_{(i)}-^{13}\text{C}_{\alpha(i-1)}$
HN(CO)CA	—	$^1\text{H}_{\text{N}(i)}-^{15}\text{N}_{(i)}-^{13}\text{C}_{\alpha(i-1)}$
HNCO	—	$^1\text{H}_{\alpha(i)}-^{15}\text{N}_{(i)}-^{13}\text{CO}_{(i-1)}$

El experimento de correlación ^{15}N -HSQC (Heteronuclear Single Quantum Correlation) (Bodenhausen y Ruben, 1980) es un espectro bidimensional en el que se transfiere la magnetización entre los ^{15}N de los grupos NH_i y los ^1H unidos directamente a ellos, que se encuentran acoplados escalarmente. De esta forma para cada grupo amida del esqueleto peptídico obtenemos un pico de cruce a la frecuencia del ^1H y del ^{15}N al que está unido. También obtenemos picos de cruce para los protones de cadenas laterales unidos a ^{15}N como Asn, Gln, Arg y Trp. El análisis de estos espectros en conjunción con los homonucleares de protón permite la asignación de los grupos ^1H - ^{15}N de los residuos de una cadena polipeptídica y son la base para la espectroscopia tridimensional heteronuclear.

Los nombres de los experimentos de triple resonancia son muy descriptivos, ya que consisten en los nombres de los núcleos a través de los cuales se transfiere la magnetización durante el experimento. Se nombran según el orden de uso de cada núcleo en la ruta de transferencia de magnetización que se produce en el experimento, poniendo entre paréntesis los núcleos que solamente se usan para la transferencia y cuyas frecuencias no son detectadas. Si observamos la Figura 6.9.2, en ella están representados los experimentos de triple resonancia empleados para la asignación de SPc41, mostrando las conectividades que se detectan en cada uno de ellos en color rojo y los átomos que se usan para la transferencia de magnetización pero que no se detectan en color verde. En el experimento **CBCA(CO)NH** (Geen y Freeman, 1991), para cada par de frecuencias ^1H - ^{15}N de cada residuo (i) debe haber dos señales correspondientes al $^{13}\text{C}_\alpha$ y al $^{13}\text{C}_\beta$ del residuo anterior (i-1). En cambio, en el experimento **CBCANH** (Patt, 1992), para cada par de frecuencias ^1H - ^{15}N de cada residuo (i) deben aparecer a la vez las señales correspondientes a los $^{13}\text{C}_\alpha$ y al $^{13}\text{C}_\beta$ del mismo residuo (i) y los del residuo anterior (i-1). Las señales de correlación que se

obtienen con los carbonos $^{13}\text{C}_\beta$ son de signo negativo y con $^{13}\text{C}_\alpha$ son de signo positivo, lo que facilita la asignación. En general, la intensidad de los picos de correlación intrarresiduo es mayor que la de los secuenciales. En el experimento **HNCA** (Clare y Gronenborn, 1991), para cada par de frecuencias ^1H - ^{15}N de cada residuo (i) debe haber dos señales correspondientes a los $^{13}\text{C}_\alpha$ del residuo (i) y del anterior (i-1). En el espectro **HN(CO)CA** (Emsley, 1994) para cada par de frecuencias ^1H - ^{15}N de cada residuo (i) debe aparecer sólo la señal del $^{13}\text{C}_\alpha$ del residuo anterior (i-1). Finalmente, el experimento **HNCO** (Kay, L. E. *et al.*, 1994) correlaciona las señales de ^1H y ^{15}N de cada grupo amida de cada residuo (i) con los las de los ^{13}CO del residuo anterior (i-1).

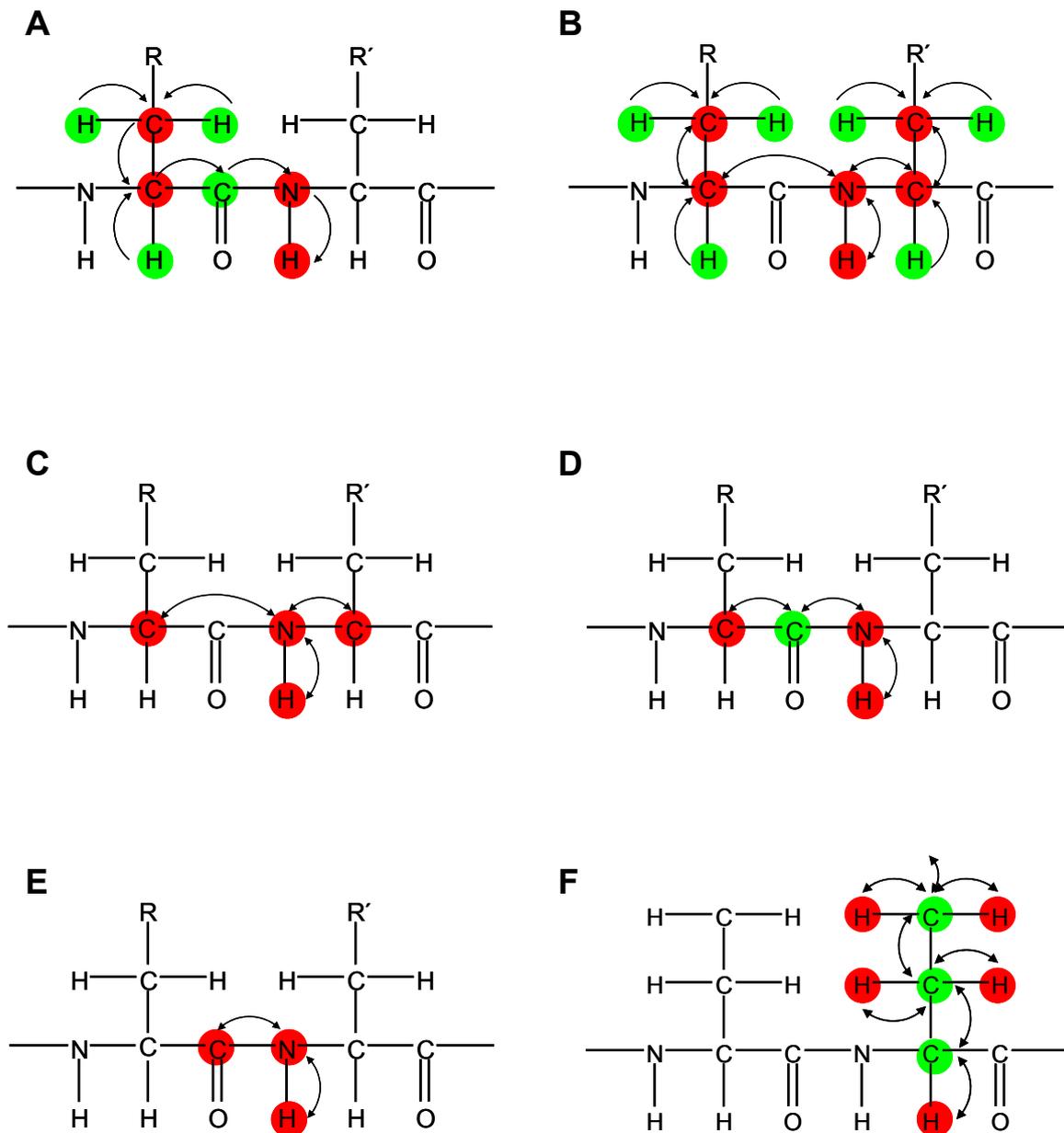


Figura 6.9.2. Experimentos de triple resonancia usados para la asignación de la quimera SPCp41. Los distintos dipéptidos mostrados en la figura muestran las conectividades detectadas (rojo) y los átomos que no se detectan, pero que se usan para la transferencia de la magnetización (verde). **A.** CBCA(CO)NH; **B.** CBCANH; **C.** HNCA; **D.** HN(CO)CA; **E.** HNCO; **F.** HCCH-TOCSY.

El análisis de estos espectros se ha llevado a cabo de forma manual. La información de los desplazamientos químicos de $^{13}\text{C}_\alpha$, y especialmente en combinación con los $^{13}\text{C}_\beta$, permiten asignar tentativamente las señales de algunos tipos de aminoácidos, o aminoácidos concretos como Ala, Gly, Ser y Thr.

Para la asignación de las cadenas laterales alifáticas empleamos los espectros HCCH-TOCSY, ^{15}N -TOCSY-HSQC y ^{15}N -NOESY-HSQC. En el experimento **HCCH-TOCSY** (Kay, L. E. *et al.*, 1993) la magnetización se transfiere por toda la cadena lateral a través de acoplamiento directos entre ^1H - ^{13}C o ^{13}C - ^{13}C unidos por un solo enlace y obtenemos un plano ^{13}C - ^1H donde al desplazamiento químico de cada ^{13}C se encuentran picos de cruce con los protones del mismo sistema de espín. En el experimento **^{15}N -TOCSY-HSQC** (Palmer *et al.*, 1992) la magnetización se transfiere a través de protones vecinos (separados por tres enlaces) y se aprovecha la existencia de un heteronúcleo para incrementar la resolución del experimento aumentando en una dimensión (la del núcleo de ^{15}N) el mismo. El resultado de este experimento es un cubo con dos dimensiones de ^1H y una de ^{15}N . Para cada plano de ^{15}N - ^1H se observan los protones del mismo aminoácido. En el experimento **^{15}N -NOESY-HSQC** (Palmer *et al.*, 1992), sin embargo, la magnetización se transfiere desde el protón amida mediante correlaciones de tipo NOE a los protones cercanos y, al igual que en el ^{15}N -TOCSY-HSQC, se aprovecha la existencia de un heteronúcleo para incrementar la resolución del experimento aumentando en una dimensión (la del núcleo de ^{15}N) el mismo. El resultado de este experimento es un cubo con dos dimensiones de ^1H y una de ^{15}N . Para cada plano de ^{15}N - ^1H se observan los NOEs intrarresiduo e interresiduo facilitando la asignación secuencial.

La asignación de las señales de RMN de la quimera SPC41 ha sido depositada en la base de datos del *BioMagResBank*, BMRB (<http://www.bmrb.wisc.edu/>), con el número de acceso 15013.

6.9.4. Cálculo de la estructura

La determinación de la estructura tridimensional de una proteína en disolución consiste en la obtención de las coordenadas atómicas compatibles con las restricciones experimentales derivadas de los espectros de RMN. Como se menciona al principio de esta Sección, nosotros hemos utilizado restricciones de distancias interatómicas, derivadas del volumen de los NOEs observados en los espectros NOESY, y de ángulos diedros, obtenidos a partir de las medidas de la variación de los desplazamientos químicos experimentales con respecto a los tabulados para estructura desordenada. En general, habrá más de una estructura que cumpla estas restricciones y, por ello, la etapa de cálculo de estructuras en RMN consiste en encontrar el conjunto de estructuras que mejor satisfaga la información experimental y que sea estadísticamente representativo del espacio conformacional compatible con los datos de RMN. Por este motivo, cuando se representan estructuras determinadas por RMN se muestra una superposición de varias conformaciones. Para el cálculo se utilizó el programa CYANA 2.0 (**C**ombined assignment and **d**Ynamics **A**lgorithm for **N**MR **A**pplications), que usa una combinación de cálculo de estructuras por geometría de distancias, dinámica molecular en ángulos de torsión y asignación automática de espectros NOESY basado en el algoritmo CANDID (Herrmann *et al.*, 2002) (<http://quentert.gsc.riken.go.jp/Software/Cyana.html>).

En función de la aproximación matemática empleada existen diversos métodos de cálculo de estructuras a partir de la información experimental, que pueden dividirse

en dos grandes grupos: métodos de geometría de distancias y métodos de dinámica molecular restringida. El método empleado para realizar los cálculos presentados es un método de geometría de distancias que utiliza una *función blanco* variable (Guntert *et al.*, 1991; Guntert, 2004). El procedimiento consiste en la minimización de una *función blanco* positiva definida de manera que tome valor cero si se cumplen todas las restricciones experimentales y distinto de cero en cualquier otro caso. La definición exacta de esta *función blanco* (V) es:

$$V = \sum_{c=U,I,V} w_c \sum_{(\alpha,\beta) \in I_c} (d_{\alpha\beta} - b_{\alpha\beta})^2 + w_a \sum_{i \in I_a} \left[1 - \frac{1}{2} \left(\frac{\Delta_i}{\Gamma_i} \right)^2 \right] \Delta_i^2 \quad (6.9.3)$$

donde $b_{\alpha\beta}$ y $d_{\alpha\beta}$ son los límites superiores e inferiores de distancia, respectivamente, entre dos átomos α y β . I_u , I_i y I_v son el conjunto de parejas de átomos (α,β) con distancias de enlace superiores, inferiores y de van der Waals, respectivamente, e I_a es el conjunto de restricciones de ángulos diedros. w_u , w_i , w_v y w_a son los factores de carga de los distintos tipos de restricciones. $\Gamma = \pi - (\theta_i^{\max} - \theta_i^{\min})/2$ representa la anchura media del rango prohibido para los valores de los ángulos diedros y se consideran las restricciones de los ángulos diedros, θ_i , como intervalos permitidos $[\theta_i^{\min}, \theta_i^{\max}]$, y Δ_i es el tamaño de las violaciones en las restricciones de los ángulos diedros.

La *función blanco* depende de los ángulos de torsión de cada residuo, ya que el número de grados de libertad del problema es mucho menor que si la dependencia de la *función blanco* se define con respecto a las coordenadas cartesianas de los átomos. Esta estrategia permite fijar la estructura covalente de la molécula (distancias y ángulos de enlace, quiralidad, planaridad de anillos aromáticos, etc.) y la información experimental incluida en la función blanco son las restricciones de distancias y de ángulos diedros. Además, uno de los términos de la *función blanco* penaliza el solapamiento de los radios de van der Waals de los átomos que constituyen la molécula estudiada. Una vez definida la *función blanco*, el cálculo de estructura consiste en un proceso de minimización de esta función, ya que a las conformaciones que mejor satisfacen las restricciones experimentales les corresponde un menor valor de *función blanco*. Existen diversos algoritmos de minimización de funciones. Los más eficientes son los que se basan en métodos de “templado simulado” (*simulated annealing*), en los que el sistema se “calienta” para permitirle saltar barreras de potencial que de otra forma le dejarían atrapado en mínimos locales, y a continuación se “enfía” lentamente hacia el mínimo absoluto que satisface todas las restricciones. En general, para evitar los mínimos locales en lugar de minimizar directamente la *función blanco*, lo que se hace es minimizar en pasos sucesivos funciones de forma análoga a la *función blanco*, pero en las que cambia el peso de los diferentes términos. Así, en los primeros pasos de la minimización sólo se incorporan las restricciones intrarresiduales y secuenciales, mientras que las restricciones no secuenciales se incluyen en los siguientes pasos, en los que se va aumentando progresivamente el peso de las interacciones de más largo alcance.

6.9.5. Análisis y calidad de las estructuras

Para analizar la calidad de las estructuras se utilizó PROCHECK-NMR (Laskowski *et al.*, 1996). La visualización de las estructuras se llevó a cabo con los programas MOLMOL (Koradi *et al.*, 1996) y PYMOL (<http://www.pymol.org/>). La precisión de las estructuras también se evaluó mediante MOLMOL. En primer lugar, se superpusieron las estructuras con menor valor de la función blanco y que, por tanto,

son las que mejor cumplen las restricciones experimentales. A continuación, se calculó la desviación cuadrática media (RMSD) entre las coordenadas atómicas de todos los pares de estructuras calculadas. Con el programa PROCHECK-NMR se determinaron, además, los ángulos entre los distintos elementos de estructura secundaria y el mapa de Ramachandran para los valores de ángulos diedros ϕ y ψ . Las coordenadas estructurales experimentalmente derivadas de las restricciones utilizadas para el cálculo han sido depositadas en el *RCSB Protein Data Bank*, PDB (<http://www.rcsb.org/pdb/>) con el número de acceso 2JMC.

El área superficial accesible al disolvente (ASA) de la quimera SPCp41 se calculó utilizando el programa NACCESS (Hubbard y Thornton, 1993) utilizando los parámetros por defecto para el radio del agua (1.4 Å), el parámetro *z-slice* (0.05 Å) y los radios de van der Waals. Para calcular la superficie de área accesible al disolvente enterrada en la interfase de unión de la quimera se calculó la diferencia entre el ASA de la quimera SPCp41 y la suma de las ASAs del dominio SH3 (aminoácidos 1-67) y del segmento p41 de la quimera (aminoácidos 68-77) por separado. Dado que se definen diferentes tipos de superficies moleculares, es conveniente especificar a qué nos referimos exactamente con el término ASA; entendemos por tal el área de superficie que describe el centro de una sonda esférica al rodar en contacto permanente con las esferas rígidas de van der Waals de los átomos de la proteína.

6.10. Dinámica de proteínas por RMN

Una visión estática de la estructura de una proteína resulta insuficiente para explicar las características funcionales de las proteínas, ya que los procesos dinámicos son cruciales para que estas moléculas desarrollen su función biológica. Por ejemplo, el proceso de reconocimiento que implica la interacción entre una proteína y su sustrato, un antígeno y su anticuerpo, o un inhibidor y su receptor, necesita de cierto grado de flexibilidad de determinadas partes de la proteína para permitir un reconocimiento eficaz, rápido y selectivo. Por otra parte, es preciso que cada uno de los componentes del proceso sea capaz de modificar su conformación y adoptar la requerida para poder llevar a cabo su función. No todas las partes de una misma proteína comparten las mismas propiedades dinámicas. De hecho la actividad química de las mismas se localiza, en la mayoría de las ocasiones, en la superficie de la molécula, donde reside la parte de mayor movilidad del sistema. Así se puede decir que, aunque la estructura de una molécula es de gran interés para el conocimiento del sistema, el estudio de su dinámica es imprescindible para la comprensión global de su actividad.

La naturaleza de los procesos dinámicos en las proteínas es muy diversa, y abarca un rango de frecuencias que va desde 10^{+14} s^{-1} hasta 10^{-2} s^{-1} (procesos del orden de segundos) y un rango de amplitudes que oscila desde décimas hasta cientos de ángstrom. En la Figura 6.10.1 se muestra la escala temporal de los movimientos moleculares en proteínas, junto con las técnicas que se pueden emplear para su estudio, destacándose los parámetros que pueden medirse por RMN.

A grandes rasgos, las proteínas en disolución pueden presentar dos tipos de movimientos. Los *movimientos globales*, afectan a la molécula en su conjunto y pueden estudiarse considerando a la molécula como un sólido rígido. Entre estos movimientos se encuentran los procesos de difusión rotacional y translacional. La forma de la molécula resulta ser un factor crucial a la hora de definir la naturaleza de este tipo de movimientos. Así, en proteínas globulares de forma aproximadamente

esférica, no existen ejes preferentes de rotación por lo que el proceso de difusión rotacional es isotrópico. Este movimiento se caracteriza por un único tiempo de correlación global, que es el tiempo promedio que tardan las moléculas en girar un radián alrededor de cualquier eje.

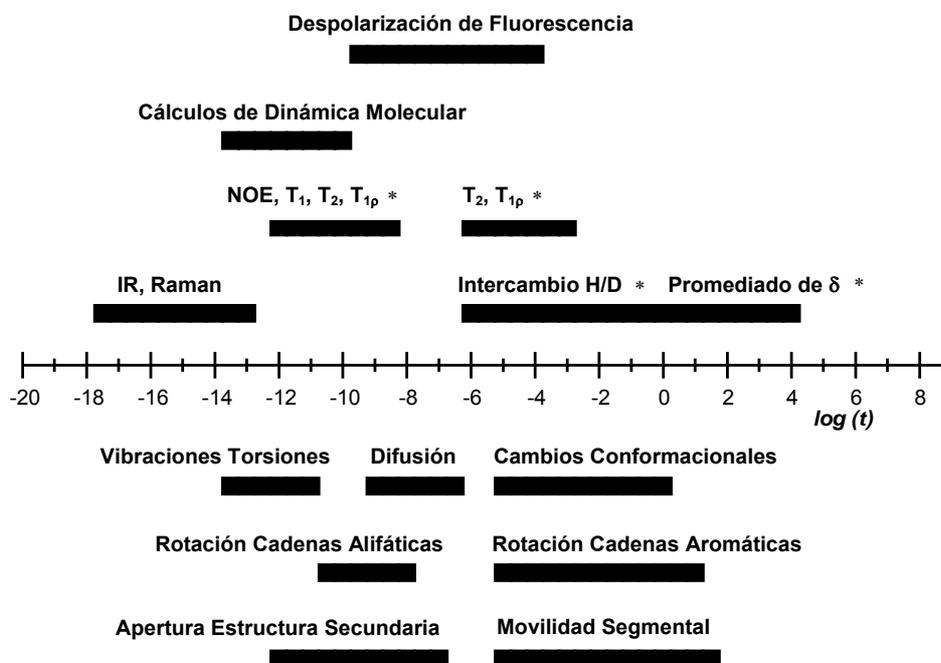


Figura 6.10.1. Escalas temporales de los movimientos moleculares en proteínas (parte inferior de la figura) y técnicas experimentales usadas para estudiarlas (parte superior). * Parámetros que pueden ser medidos por Resonancia Magnética Nuclear.

En muchos otros casos, las proteínas muestran comportamientos hidrodinámicos más complejos, con mayor o menor grado de anisotropía. En estos casos, hay ejes preferentes para la rotación de la molécula y tiempos de correlación distintos para las rotaciones alrededor de dichos ejes.

Otro tipo de movimientos que podemos encontrar en las proteínas son los *movimientos internos*, que implican translaciones y rotaciones de unas partes de la molécula respecto de otras. Estos movimientos internos presentan amplitudes muy diversas que oscilan desde los movimientos locales (0.01 Å–5 Å), como las vibraciones atómicas y las rotaciones de cadenas laterales, hasta cambios conformacionales extensos (decenas de Å). Estos movimientos pueden ocurrir a escalas de tiempo muy distintas, tan rápidas como unos pocos nanosegundos, o tan lentas como días. Hay que destacar que no existe necesariamente correlación entre la amplitud del movimiento y la velocidad del mismo. Así, entre los movimientos locales podemos encontrarnos desde procesos tan rápidos como los vibracionales y torsionales (10^{-11} – 10^{-13} segundos), o moderadamente lentos como la rotación de anillos aromáticos (~ 1 – 10^{-5} segundos). La rotación de las cadenas laterales alifáticas ocurre en una escala de tiempos intermedia (10^{-8} – 10^{-10} segundos).

Los movimientos de pequeña amplitud, en la escala de milisegundos a microsegundos, han demostrado ser esenciales para la función biológica de las proteínas. La dinámica interna de las proteínas no es sólo importante para su función

biológica, sino que también lo es en gran medida para su estabilidad. En el estado desplegado las proteínas presentan un contenido entrópico muy elevado, ya que no existen factores estructurales que restrinjan la conformación de determinados enlaces. El proceso de plegamiento introduce este tipo de restricciones que conllevan una pérdida de entropía. Este coste entrópico se mitiga en parte, gracias a que muchas cadenas laterales siguen manteniendo una amplia libertad conformacional en el estado plegado.

6.10.1. Medidas de relajación de ^{15}N en NH amídicos

En el término relajación se engloban todos aquellos fenómenos fisicoquímicos mediante los cuales un sistema perturbado evoluciona hacia la situación de equilibrio térmico. En RMN, la relajación tiene lugar mediante procesos de transferencia de energía entre el núcleo de la molécula y su entorno (Pérez-Cañadillas y Bruix, 2003). Al contrario de lo que sucede en otras espectroscopias, en RMN los niveles excitados suelen tener vidas medias muy largas (milisegundos a segundos), dificultando la vuelta del sistema al equilibrio térmico. Para que tenga lugar la relajación, han de existir campos magnéticos fluctuantes de carácter local en el entorno del núcleo estudiado y con la frecuencia e intensidad adecuadas: el enlace N-H presenta naturaleza propia polar que, debido a los movimientos moleculares tanto globales como internos, se comporta como un pequeño dipolo generando el tipo de campo mencionado. Es posible deducir expresiones teóricas para los distintos parámetros de relajación si se conocen los mecanismos operativos en cada caso y asumiendo un determinado modelo dinámico. En estas expresiones, la dinámica molecular viene incluida en términos de la función de densidad espectral $J(\omega)$, que es una medida de la intensidad de los movimientos moleculares a cada frecuencia ω .

En proteínas, los núcleos que se suelen usar en estudios dinámicos son el ^{13}C y el ^{15}N , por lo que los parámetros de relajación que se miden por RMN son de relajación heteronuclear: R_1 ($1/T_1$), o velocidad de relajación longitudinal, R_2 ($1/T_2$), o velocidad de relajación transversal, y el NOE heteronuclear. La estructura de una proteína se describe de manera adecuada mediante sus coordenadas atómicas, pero la descripción de su estado dinámico es más compleja. Los experimentos de relajación son los que aportan más información aunque sólo permiten describir algunos aspectos de la movilidad. Los estudios de relajación de ^{15}N proporcionan información del movimiento global de la proteína, pudiendo determinar el tipo de movimiento global (isótropo o anisótropo) que mejor define la dinámica del sistema y el tiempo de correlación global de la molécula, τ_m , en la escala de los nanosegundos. La información sobre la dinámica interna en la escala de los subnanosegundos y de los mili-microsegundos deriva principalmente del análisis “*model-free*” (sin modelo) propuesto por Lipari y Szabo (Lipari y Szabo, 1982a; Lipari y Szabo, 1982b) y ampliado posteriormente por otros autores.

Las velocidades de relajación R_1 , R_2 y el NOE heteronuclear de un ^{15}N amídico están dominados por la interacción dipolar (DD) del núcleo ^{15}N con el protón al que se encuentra unido y por la anisotropía de desplazamiento químico (CSA) (Abragam, 1961):

$$R_1 = \frac{1}{T_1} = d^2 [J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H + \omega_N)] + c^2 J(\omega_N) \quad (6.10.1)$$

$$R_2 = \frac{1}{T_2} = \frac{1}{2}d^2[4J(0) + J(\omega_H - \omega_N) + 3J(\omega_N) + 6J(\omega_H) + 6J(\omega_H + \omega_N)] + \frac{c^2}{6}[3J(\omega_N) + 4J(0)] \quad (6.10.2)$$

$$NOE = 1 + \left(\frac{\gamma_H}{\gamma_N}\right)d^2[6J(\omega_H + \omega_N) - J(\omega_H - \omega_N)]T_1 \quad (6.10.3)$$

Las constantes d^2 y c^2 se definen como,

$$d^2 = 0.1\gamma_H^2\gamma_N^2h^2 / (4\pi^2) \langle 1/r_{NH}^3 \rangle^2 \quad (6.10.4)$$

$$c^2 = (2/15)\gamma_N^2H_0^2(\sigma_{\parallel} - \sigma_{\perp})^2 \quad (6.10.5)$$

donde γ_H y γ_N son los radios giromagnéticos de los núcleos ^1H y ^{15}N , respectivamente, ω_H y ω_N son las frecuencias de Larmor de ^1H y ^{15}N , r_{NH} es la distancia internuclear ^1H - ^{15}N (1.02 Å), H_0 es la intensidad de campo magnético, y σ_{\parallel} y σ_{\perp} son respectivamente las componentes paralela y perpendicular, del tensor de desplazamiento químico de ^{15}N , que se supone axialmente simétrico.

La función de densidad espectral se puede representar usando el formalismo “*model-free*” antes mencionado (Lipari y Szabo, 1982a; Lipari y Szabo, 1982b), que es el que nosotros hemos considerado para todos los estudios de relajación realizados en esta Memoria. Este formalismo utiliza un número mínimo de parámetros para describir el movimiento isotrópico global de una macromolécula y los movimientos internos del vector ^1H - ^{15}N , en la siguiente expresión:

$$J(\omega) = [S^2\tau_m / (1 + \omega^2\tau_m^2)] + [(1 - S^2)\tau / (1 + \omega^2\tau^2)] \quad (6.10.6)$$

El parámetro de orden S^2 describe el grado de restricción espacial del movimiento interno del vector H-N; τ_m es el tiempo de correlación como resultado del movimiento isotrópico global de la molécula entera. Si no hay restricción de movimientos internos S^2 es cero, mientras que vale uno cuando la restricción del vector N-H es total. El tiempo de correlación efectivo resultante de los movimientos internos se define como τ_e , donde $1/\tau = 1/\tau_m + 1/\tau_e$.

La única premisa necesaria para aplicar la aproximación de Lipari-Szabo es suponer que los movimientos globales e internos son independientes y, en principio, no es necesario hacer ninguna suposición acerca del tipo de movimientos o de las escalas temporales de los mismos. No obstante, la interpretación de los parámetros de relajación de RMN, en los términos de la aproximación de Lipari-Szabo sólo es útil para caracterizar movimientos internos al menos de 10 a 100 veces más rápidos que la difusión rotacional global.

En algunas ocasiones la aproximación “*model-free*” no reproduce satisfactoriamente los parámetros de relajación. Esto ocurre cuando no es posible la descripción del movimiento interno del vector N-H mediante una única escala de tiempo. En este caso, es necesario incluir un tiempo de correlación y un parámetro de orden adicionales, en lo que se conoce como “*model-free extendido*” (Clare *et al.*, 1990). Los movimientos internos pueden dividirse en dos escalas diferentes: movimientos internos lentos, y movimientos internos rápidos, (ambos más rápidos que

en el movimiento de rotación global). A estos movimientos se les asocian los correspondientes parámetros de orden, S_s^2 (movimientos lentos) y S_f^2 (movimientos rápidos) de manera que $S^2 = S_s^2 S_f^2$; y τ_s y τ_f que son los tiempos de correlación efectivos de los dos movimientos internos considerados (lentos y rápidos respectivamente). La forma completa de la función de densidad espectral asociada a este modelo extendido se representa a continuación:

$$J(\omega) = S_f^2 S_s^2 \frac{\tau_m}{1 + \tau_m^2 \omega^2} + (1 - S_s^2) \frac{\tau'}{1 + \tau'^2 \omega^2} + (S_f^2 - S_s^2 S_f^2) \frac{\tau''}{1 + \tau''^2 \omega^2} \quad (6.10.7)$$

donde $1/\tau' = 1/\tau_s + 1/\tau_m$ y $1/\tau'' = 1/\tau_f + 1/\tau_m$. El valor de τ_f suele ser tan pequeño que es posible simplificar la ecuación omitiendo el último término y la ecuación se simplifica a lo que se conoce como función de densidad espectral reducida (“*model-free reducido*”):

$$J(\omega) = S_f^2 S_s^2 \frac{\tau_m}{1 + \tau_m^2 \omega^2} + (1 - S_s^2) \frac{\tau'}{1 + \tau'^2 \omega^2} \quad (6.10.8)$$

Cuando en el modelo se considera la relajación transversal, se requiere un término adicional, R_{ex} , que tiene en cuenta las contribuciones debidas a procesos de intercambio químico, además de la DD y la CSA. En muchos casos, estas contribuciones son el resultado de promediar un cambio conformacional. El término adicional, R_{ex} , se puede incluir para modificar el valor calculado para R_2 usando la siguiente relación,

$$R_2 = R_{2(DD)} + R_{2(CSA)} + R_{ex} \quad (6.10.9)$$

donde los sufijos DD y CSA indican las contribuciones de la interacción dipolar y la anisotropía de desplazamiento químico a la relajación transversal.

6.10.2. Registro de los espectros

Todas las medidas de relajación heteronuclear de ^{15}N se realizaron en muestras de proteína enriquecida con ^{15}N a pH 3.5 y a 30°C para la quimera SPCp41 y a 25°C para la quimera SPCp41 y el mutante P74A. A 800 MHz se obtuvo el NOE heteronuclear para la quimera y a 600 MHz se obtuvieron R_1 , $R_{1\rho}$ y NOE para SPCp41 y para el mutante P74A. Los valores de R_2 se calcularon a partir de las medidas de R_1 y $R_{1\rho}$ como se describe más abajo. Las secuencias de pulsos utilizadas para la medida de los parámetros de relajación se basan principalmente en la de correlación bidimensional heteronuclear HSQC (Bodenhausen y Ruben, 1980). La aplicación de estas secuencias permite adquirir espectros de similares características al HSQC convencional, pero con señales de correlación cuya intensidad está modulada en función de la pérdida de magnetización debida a la relajación (Farrow *et al.*, 1994).

La determinación del NOE heteronuclear se ha realizado mediante dos experimentos prácticamente idénticos, uno de ellos con saturación de protón durante la preparación (experimento con NOE) y el otro sin ella (experimentos sin NOE). La saturación del protón en el experimento con NOE se consigue mediante un tren de pulsos aplicado durante 3 segundos inmediatamente antes del primer pulso en el heteronúcleo.

Para la determinación de las velocidades de relajación se aplicaron las secuencias de pulsos correspondientes, incrementando el número de módulos de relajación (n). Los valores de n para las medidas de R_1 fueron (en ms): 10, 100, 200, 300, 400, 500, 600, 700, 850, 1000, 1150, 1300, 1500, y para las medidas de $R_{1\rho}$ fueron (en ms): 10, 20, 40, 60, 80, 100, 130, 160, 200, 240, 280, 330, 400.

Después de cada espectro se registró un espectro ^{15}N -HSQC, para corregir la pérdida de intensidad debida a la lenta agregación experimentada por el mutante P74A. En la Tabla 6.10.1 se indican el número de puntos registrados para cada dimensión de los espectros y el número de barridos realizados, que debido al problema de agregación, fueron el doble para todos los experimentos del mutante.

Tabla 6.10.1. Parámetros de adquisición y procesado de los distintos experimentos de relajación para la quimera SPCp41 y el mutante P74A.

Experimento	Anchura espectral (ppm)	Número de puntos complejos	Número de barridos	^a Número total de puntos
SPCp41 (800 MHz)				
$^{15}\text{N}\{^1\text{H}\}$ NOE con saturación de ^1H	$^1\text{H} \rightarrow 13$ $^{15}\text{N} \rightarrow 30$	$^1\text{H} \times ^{15}\text{N}$ 1024 x 128	40	1024 x 1024
$^{15}\text{N}\{^1\text{H}\}$ NOE sin saturación de ^1H	$^1\text{H} \rightarrow 13$ $^{15}\text{N} \rightarrow 30$	$^1\text{H} \times ^1\text{H}$ 1024 x 128	40	1024 x 1024
SPCp41 (600 MHz)				
R_1	$^1\text{H} \rightarrow 14.1$ $^{15}\text{N} \rightarrow 36.2$	$^1\text{H} \times ^{15}\text{N}$ 1024 x 64	32	2048 x 256
$R_{1\rho}$	$^1\text{H} \rightarrow 14.1$ $^{15}\text{N} \rightarrow 36.2$	$^1\text{H} \times ^{15}\text{N}$ 1024 x 64	32	2048 x 256
$^{15}\text{N}\{^1\text{H}\}$ NOE con saturación de ^1H	$^1\text{H} \rightarrow 14.1$ $^{15}\text{N} \rightarrow 36.2$	$^1\text{H} \times ^{15}\text{N}$ 1024 x 64	32	2048 x 256
$^{15}\text{N}\{^1\text{H}\}$ NOE sin saturación de ^1H	$^1\text{H} \rightarrow 14.1$ $^{15}\text{N} \rightarrow 36.2$	$^1\text{H} \times ^{15}\text{N}$ 1024 x 64	32	2048 x 256
P74A (600 MHz)				
R_1	$^1\text{H} \rightarrow 14.1$ $^{15}\text{N} \rightarrow 36.2$	$^1\text{H} \times ^{15}\text{N}$ 1024 x 64	32	2048 x 256
$R_{1\rho}$	$^1\text{H} \rightarrow 14.1$ $^{15}\text{N} \rightarrow 36.2$	$^1\text{H} \times ^{15}\text{N}$ 1024 x 64	32	2048 x 256
$^{15}\text{N}\{^1\text{H}\}$ NOE con saturación de ^1H	$^1\text{H} \rightarrow 14.1$ $^{15}\text{N} \rightarrow 36.2$	$^1\text{H} \times ^{15}\text{N}$ 1024 x 64	32	2048 x 256
$^{15}\text{N}\{^1\text{H}\}$ NOE sin saturación de ^1H	$^1\text{H} \rightarrow 14.1$ $^{15}\text{N} \rightarrow 36.2$	$^1\text{H} \times ^{15}\text{N}$ 1024 x 64	32	2048 x 256

^aDespués del procesado.

6.10.3. Procesado y análisis de los datos

Los espectros bidimensionales se procesaron con el programa NMRPipe (Delaglio *et al.*, 1995) y fueron analizados con el programa SPARKY (Goddard y Kneller). Las intensidades de los picos de cruce en los espectros bidimensionales se han medido utilizando la subrutina de integración del programa SPARKY. Los valores de R_1 y $R_{1\rho}$ se han determinado ajustando los decaimientos de la intensidad de los picos a una función exponencial decreciente de la forma:

$$I(t) = I_0 \exp(-R_{1\rho}t) \quad (6.10.10)$$

donde $I(t)$ es la intensidad tras un tiempo de relajación t y I_0 es la intensidad a $t = 0$, utilizando para ello el programa ORIGIN 6.1 (Microcal Software Inc.).

El valor de R_2 se ha calculado a partir de $R_{1\rho}$ teniendo en cuenta que:

$R_{1\rho} = R_1 \cos^2(\theta) + R_2 \sin^2(\theta)$, donde θ es el ángulo entre el campo magnético efectivo, B_{efectivo} y el campo magnético principal, B_0 . Esto significa que para cada pico del espectro se necesita calcular la diferencia que existe entre la frecuencia del ^{15}N y la del *offset* del espectro: O (Hz) = (ppm_{15N} (pico) - ppm_{15N} (*offset*)) * ν_{N} . En los experimentos realizados en el espectrómetro de 600 MHz la frecuencia del nitrógeno es: $\nu_{\text{N}} = 60.7690010071$ y ppm_{15N} (*offset*) = 118.258 ppm. Además se necesita saber cual es la potencia del pulso *spinlock* (B_1). B_1 se puede estimar como: B_1 (Hz) $\sim 1/4$ (4 * pulso de 90° usado en el *spinlock*), siendo en nuestros experimentos $B_1 = 6250$ Hz. Así, al final podemos calcular θ como: $\tan(\theta) = B_1/O$.

Los valores de NOE heteronuclear se han determinado mediante la siguiente relación:

$$^{15}\text{N}\{^1\text{H}\}\text{NOE} = \frac{I_{\text{sat}}}{I_0} \quad (6.10.11)$$

donde I_{sat} e I_0 representan las correspondientes intensidades de las señales en los espectros con y sin saturación del protón respectivamente. En el caso del mutante P74A, para corregir la pérdida de intensidad debida a la agregación experimentada por éste, las intensidades I_{sat} fueron determinadas haciendo la media de las intensidades de los espectros con saturación recogidos antes y después de los experimentos sin saturación.

El tiempo de correlación global y la relación R_2/R_1 para la quimera SPCp41 se han calculado con el programa Fast-HYDRONMR (Ortega y Garcia de la Torre, 2005), utilizando para construir el fichero de entrada para el programa una temperatura de 298°K, una viscosidad de 1.1 cP y un radio de 2.6 Å para los elementos esféricos usados.

6.11. Calorimetría isotérmica de titulación (CIT)

La calorimetría isotérmica de titulación (CIT) es una técnica que nos permite medir directamente el calor asociado a un proceso de interacción, el cual, a presión constante, equivale a la entalpía de este proceso (Wiseman *et al.*, 1989). Si el efecto térmico neto es lo suficientemente grande, se pueden determinar parámetros tales como la constante de equilibrio de asociación, K_a , el cambio de entalpía de unión, ΔH_a ,

así como la estequiometría de la interacción, n . Como consecuencia, se pueden calcular la energía de Gibbs y la entropía de formación del complejo (ΔG_a , ΔS_a). Además, si se realizan experimentos a diferentes temperaturas es posible determinar el cambio en la capacidad calorífica del proceso, ΔC_p .

Los experimentos de titulación del dominio S19P20s con el péptido p41 se realizaron a 25°C, utilizando un microcalorímetro isotérmico de titulación MCS-ITC (Microcal INC, Northampton, MA). La disolución de S19P20s se dializó frente al tampón de realización del experimento, glicina 20 mM, pH 3.5. Las disoluciones ya dializadas se filtraron, se desgasificaron para evitar la formación de burbujas y se equilibraron a la temperatura experimental. La disolución de proteína en la célula de muestra de concentración 0.44 mM fue titulada con el ligando p41, cuya disolución se preparó a 10.83 mM, disolviendo el péptido liofilizado en el tampón de diálisis y reajustando el pH. Debido a la baja afinidad de este ligando por el dominio, los experimentos de titulación se realizaron inyectando sucesivamente volúmenes variables (perfil de volúmenes variable) de la disolución de ligando (de 2 μ L a 13 μ L) para poder definir mejor la curva de titulación. De estos experimentos obtenemos el termograma correspondiente a la interacción. Del mismo modo, es necesario obtener el termograma de dilución para lo que titulamos el ligando sobre la disolución tampón en ausencia de proteína utilizando el mismo perfil de volúmenes. En ambos casos, el calor desarrollado tras cada inyección de péptido fue medido a partir de la integral de la señal calorimétrica. Una vez corregida la línea base de ambos experimentos, la isoterma de unión neta para la interacción proteína-ligando se obtiene restando para cada inyección los calores de dilución a los calores del experimento principal. La constante de afinidad, K_a , y el cambio de entalpía aparente de unión, ΔH_a , se determinaron analizando las isotermas de unión resultantes mediante ajustes no lineal de mínimos cuadrados al modelo de un conjunto de sitios idénticos e independientes, tal y como se describe en (Casares *et al.*, 2007).

Los experimentos de CIT de desplazamiento se realizaron titulando una disolución de 0.44 mM de S19P20s en presencia del decapeptido APSYSPAPPP (P7Ap) con una disolución de 10.83 mM de p41. Se hicieron dos experimentos, uno con una concentración de 1.32 mM de P7Ap y otro con 5.17 mM de P7Ap. Las isotermas de unión resultantes se analizaron como describe Sigurskjold (Sigurskjold, 2000), fijando los parámetros termodinámicos de la interacción S19P20s:p41 (ver Figura 4.7.1 del Capítulo 4). El análisis de los datos se realizó utilizando el programa ORIGIN 7 (*Microcal Software Inc.*). Los péptidos sintéticos fueron adquiridos de la empresa PEPSCAN.

6.12. Estudio de la cinética de plegamiento-desplegamiento de proteínas mediante la técnica de flujo detenido (*stopped-flow*)

Las proteínas pueden desnaturalizarse mediante la adición de agentes químicos como urea o guanidina, de ácidos o bases, mediante la unión a ligandos determinados o mediante la variación de la temperatura o la presión. Dentro de los estudios cinéticos de plegamiento y desplegamiento, llevados a cabo con proteínas, el diseño experimental más frecuente usado en el uso de las técnicas de flujo, mediante las que se consigue mezclar de forma rápida una disolución de proteína con la disolución de agente desnaturante para seguir posteriormente el proceso utilizando la fluorescencia intrínseca de las proteínas o el dicroísmo circular como observables.

Para el caso de pequeñas proteínas globulares, las cinéticas de las reacciones de plegamiento y desplegamiento inducidas por un agente desnaturante, suelen ser curvas monoexponenciales, hecho que indica la existencia de una única barrera de energía de activación para la reacción. En el caso de algunas proteínas, particularmente aquellas constituidas por varios dominios y/o conformaciones alternativas, las cinéticas pueden presentar respuestas más complicadas, dando lugar generalmente a cinéticas multiexponenciales. Estos casos se producen como consecuencia de la existencia de múltiples estados desplegados que se interconvierten lentamente o de la existencia de varias barreras de energía que el sistema ha de vencer para pasar de un estado inicial a otro final.

La sensibilidad del método de flujo detenido con la fluorescencia como observable permite utilizar disoluciones de proteína en el intervalo de concentración micromolar, por lo que no es necesario disponer de grandes cantidades de muestra, minimizando así los problemas asociados a la agregación y baja solubilidad de la misma. En estas condiciones, el instrumento obtiene una buena relación señal-ruido, con una resolución temporal limitada sólo por el tiempo muerto de mezcla del sistema. Pueden aplicarse también en estos casos otros observables físicos, como el dicróismo circular o la espectroscopia de absorción.

La ventaja del uso de la fluorescencia se basa en que la presencia de los grupos fluoróforos en las proteínas permite obtener información acerca de la exposición de estos grupos al disolvente, movimientos locales y globales en la estructura, distancias entre restos y existencia de superficies hidrofóbicas. Los cambios en la fluorescencia que se producen pueden seguirse midiendo parámetros (Eftink, 1994) tales como la intensidad de la fluorescencia a distintas longitudes de onda de excitación y emisión, el rendimiento cuántico, la longitud de onda del máximo de emisión, el tiempo de vida o la anisotropía, que proporciona información de la libertad conformacional en distintos estados de la proteína.

6.12.1. Preparación y realización de los experimentos

Rutinariamente, un experimento cinético para evaluar el plegamiento de una proteína consta de dos aspectos, por una parte la evaluación cinética de la reacción de plegamiento y por otra la de desplegamiento, que se llevan a cabo en experimentos independientes. El equipo de flujo detenido utilizado para realizar los experimentos es el modelo SX.18MV-R de la casa comercial APPLIED PHOTOPHYSICS.

Para realizar los experimentos cinéticos de plegamiento, se preparan 12 mL de proteína de concentración aproximadamente 50 μ M, disuelta en tampón fosfato sódico 50 mM, pH 7.0, conteniendo urea a una concentración 8 M. Además, se preparan una serie de tampones en las mismas condiciones de pH y fuerza iónica, conteniendo diferentes concentraciones de urea, de tal manera que tras la mezcla la concentración final de ésta en el experimento varía en el intervalo comprendido entre 0.75 y 8 M incrementándose cada 0.25 unidades. La realización de los experimentos cinéticos de desplegamiento, se lleva a cabo preparando 12 mL de proteína de forma análoga a los descritos anteriormente, a diferencia de la ausencia de urea en el tampón de trabajo y de que en este caso las disoluciones de urea se preparan en el intervalo de concentraciones comprendido entre 3 y 8 M.

Previo a la realización del experimento todas las disoluciones se termostatan en un baño a la temperatura a la que se deseen llevar a cabo los experimentos, ya que

es muy importante mantener la temperatura constante, debido a que tanto la intensidad de fluorescencia como la velocidad de los procesos son sensibles a cambios en la temperatura. En nuestro caso los experimentos se han realizado a 25 °C para la quimera SPCp41 y sus mutantes de Pro-Ala y el permutante circular S19P20s y a 5, 15, 25 y 35°C para el estudio de las cinéticas en función de la temperatura realizado para la quimera SPCp41. Antes de comenzar el experimento se realiza un lavado del sistema con agua y con el tampón de trabajo (fosfato sódico 50 mM, pH 7.0), y luego se van llenando las jeringas con la muestra de proteína y de tampón a diferentes concentraciones de urea, según el experimento que se esté realizando y se termostatan durante aproximadamente 5 minutos. La cinética se registra inmediatamente después de la mezcla rápida de estas dos disoluciones de diferente volumen (en relación 1:10).

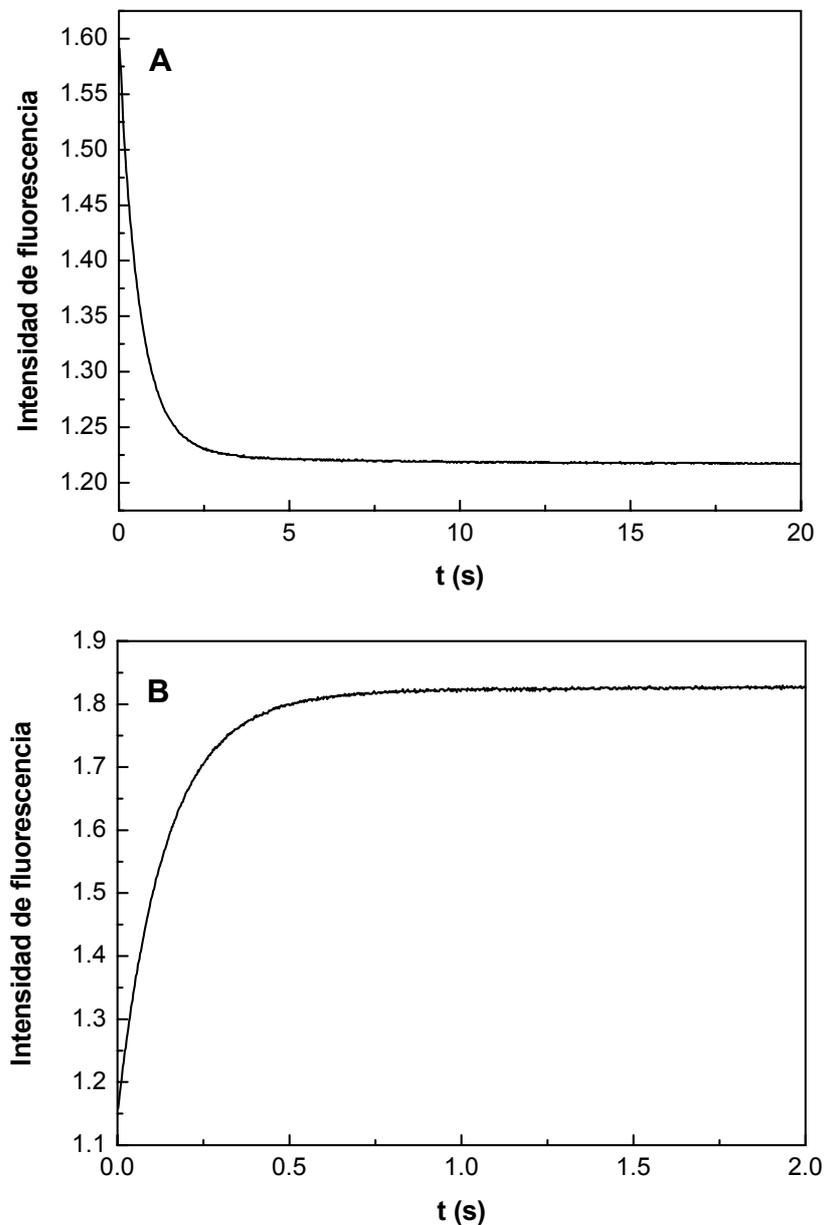


Figura 6.12.1. Perfiles cinéticos de la quimera SPCp41 registrados a 25°C. **A.** Cinética de desplegamiento de una mezcla de proteína a concentración 50 mM en tampón fosfato sódico 50 μ M, pH 7.0 con urea 6.5 M en el mismo tampón. **B.** Cinética de plegamiento de una mezcla formada por la proteína a concentración 50 μ M en tampón fosfato sódico 50 mM, pH 7.0 y urea 8 M, y urea 1.5 M en el mismo tampón.

Cada una de las curvas obtenidas es el resultado de realizar una inyección y medir la fluorescencia de una mezcla que contiene una concentración determinada de proteína y desnaturalizante. Es necesario ajustar la escala de tiempo en función de la velocidad del proceso, para conseguir curvas en las que se observe claramente la región de intensidad constante o “plateau” (Figura 6.12.1). La cinética definitiva para una mezcla concreta es el resultado del promedio de 5 ó 6 curvas obtenidas en diferentes inyecciones idénticas.

Al finalizar el experimento se determina la concentración de las diferentes disoluciones de urea, midiendo el índice de refracción correspondiente a cada caso, de la forma descrita en la Sección 6.5.2. También se miden los índices de refracción de las disoluciones de proteína en urea y tampón, para calcular las concentraciones de urea presentes en cada mezcla utilizada.

En todos los experimentos la longitud de onda de excitación utilizada fue de 298 nm, y el filtro usado de 320 nm.

6.12.2. Determinación de las constantes de velocidad

Las constantes de velocidad y las amplitudes de las cinéticas de plegamiento y desplegamiento correspondientes a una concentración de urea determinada se calculan a partir del ajuste no lineal de las curvas promediadas obtenidas utilizando una ecuación multiexponencial, según sea necesario. Para determinar de forma precisa estos parámetros, especialmente las constantes de velocidad, es muy importante estudiar el gráfico de las desviaciones de residuos de los datos experimentales sobre la curva resultante del análisis. La amplitud promedio de los puntos obtenidos de dicho gráfico debe ser muy próxima a cero y es un indicador del grado en que la ecuación cinética aplicada es válida para la descripción de los datos experimentales. Las posibles desviaciones de los residuos dependen en gran medida de la corrección del modelo cinético utilizado. En nuestro caso, la amplitud promedio del gráfico de las desviaciones de los residuos para el ajuste de las curvas a una ecuación monoexponencial, no estaba próximo a cero, mientras que para una ecuación biexponencial sí, lo que nos indicaba que teníamos dos fases cinéticas tanto en la reacción de desplegamiento como en la de plegamiento, como se describe con más detalle en el Capítulo 5 de esta Memoria. Para realizar estos análisis se utilizó el programa del propio instrumento.

El observable físico seguido fue la intensidad de fluorescencia, por lo que el valor de dicha intensidad, I_t , para un tiempo determinado viene descrito por la siguiente ecuación biexponencial de velocidad:

$$I_t = I_\infty + A_1 \cdot \exp(-k_1 \cdot t) + A_2 \cdot \exp(-k_2 \cdot t) \quad (6.12.1)$$

donde k_1 y k_2 son dos constantes de velocidad que corresponden a fases cinéticas distintas, k_p o k_d según se trate de los procesos de desplegamiento o plegamiento respectivamente a una concentración dada de urea, A_1 y A_2 son las amplitudes de cada fase exponencial, asociadas al cambio de fluorescencia que ocurre durante el proceso e I_∞ , es la intensidad de fluorescencia una vez finalizado el proceso. Una cinética de desplegamiento, como la que se representa en la Figura 6.12.1.A, presenta un valor positivo de amplitud, debido a la disminución de la intensidad de fluorescencia con el tiempo. Por el contrario, en una cinética de plegamiento (Figura 6.12.1.B) la amplitud toma valores negativos por el aumento de la señal con el tiempo,

aunque esto no ocurre siempre de la misma manera, ya que depende del sistema de estudio.

6.12.3. Análisis de las cinéticas de plegamiento-desplegamiento según la teoría del estado de transición

El estado de transición en el proceso de plegamiento y desplegado de una proteína que sigue el modelo de dos estados se define de forma sencilla como el punto de mayor energía a lo largo del camino que conduce del estado inicial (nativo) al estado final (desnaturalizado), en la representación de la energía de Gibbs en función de la coordenada de reacción (Figura 6.12.2). En el caso de un proceso en el que se dé la presencia de intermedios, cada etapa cinética lleva asociado un correspondiente estado de transición.

La reacción transcurre a través de una superficie de energía potencial que comprende una zona para el estado inicial y una zona para el estado final, separadas ambas por una barrera energética. La teoría del estado de transición (TET) supone que existe un equilibrio entre el estado de partida y el estado de transición (ET^\ddagger) (Laidler y King, 1983):



Por tanto, partiendo del estado nativo N de la proteína, existe un equilibrio entre este estado y el estado de transición $[ET^\ddagger]$, cuya constante de equilibrio sería K^\ddagger , mientras que el proceso cinético para llegar al estado desplegado D, es el que viene determinado por la constante de velocidad $k_{\ddagger-N}$. En el caso de partir del estado desplegado, la reacción, con una constante de velocidad $k_{\ddagger-D}$, sería la siguiente:



Para llevar a cabo el desarrollo matemático de esta teoría hay que considerar otras suposiciones previas. En primer lugar, que todas las moléculas que alcanzan la barrera energética llegan al estado final (Figura 6.12.2). En segundo lugar, que durante la reacción se mantiene la distribución de Boltzmann para las energías de las moléculas del estado inicial. Por último, también se supone que las moléculas que cruzan la barrera de energía desde el estado de partida, tienen una distribución de Boltzmann de energías correspondiente a la temperatura del estado inicial. Con estas premisas, se obtiene la ecuación de Eyring:

$$k = (k_B \cdot T / h) \cdot K^\ddagger \tag{6.12.2}$$

donde k representa la constante de velocidad $k_{\ddagger-N}$ ó $k_{\ddagger-D}$ según los dos casos descritos, k_B es la constante de Boltzmann, h es la constante de Planck y T es la temperatura absoluta. La constante K^\ddagger representa la constante del equilibrio químico que tiene lugar entre el estado inicial (D ó N) y el estado de transición ($[ET^\ddagger]$). En un sentido riguroso, este equilibrio implicaría que en el estado de transición existirá un número de

moléculas, algunas de las cuales, después de un tiempo determinado evolucionarán para pasar al estado final, con una constante de velocidad k , mientras que otras volverán al estado de inicial. Sin embargo, este equilibrio no puede interpretarse de esta manera, ya que el estado de transición, $[ET^\ddagger]$, tan solo implica a aquellas moléculas que alcanzan la barrera energética desde el estado inicial y, según se ha indicado anteriormente, todas las que lo consiguen pasan al estado final, con lo que el estado de transición no puede dar lugar rigurosamente a una auténtica reacción de equilibrio químico con el estado inicial, sea éste cualquiera de los dos posibles en los que la proteína exista (D ó N).

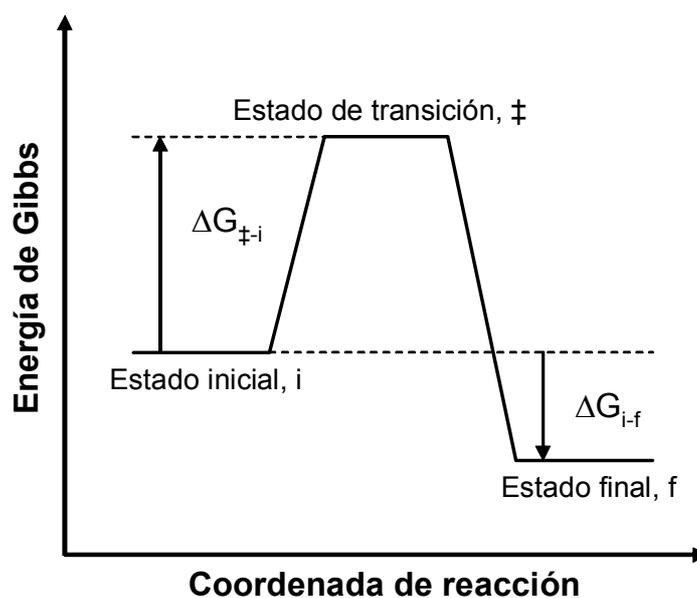


Figura 6.12.2. Teoría del estado de transición. El estado de transición es el correspondiente al estado de mayor energía a lo largo de la ruta que conduce del estado inicial al estado final. La altura de la barrera de energía del estado de transición determina en gran medida la magnitud de la constante de velocidad.

Sabiendo que la energía de Gibbs del proceso de activación puede expresarse también como:

$$\Delta G^\ddagger = \Delta H^\ddagger - T \cdot \Delta S^\ddagger = R \cdot T \cdot \ln K^\ddagger \quad (6.12.3)$$

siendo R la constante de los gases, sustituyendo en la ecuación de Eyring (6.12.2) se obtendría:

$$k = (k_B \cdot T / h) \cdot \exp(-\Delta G^\ddagger / R \cdot T) \quad (6.12.4)$$

de donde:

$$k = (k_B \cdot T / h) \cdot \exp(\Delta S^\ddagger / R) \cdot \exp(-\Delta H^\ddagger / R \cdot T) \quad (6.12.5)$$

6.12.4. Gráficos de Chevron

Cuando representamos en un gráfico, a una temperatura dada, la dependencia entre el logaritmo neperiano de las constantes de velocidad aparentes obtenidas a cada concentración de urea en función de ésta obtenemos un gráfico que tiene forma de "V" y que se denomina "Chevron", como el que se muestra en la Figura 6.12.3. Existen en la literatura numerosos estudios que detallan el análisis de este tipo de representación (Beasty *et al.*, 1986; Hurle *et al.*, 1987; Matthews y Hurle, 1987; Chen, B. L. *et al.*, 1989; Chen, B. L. y Schellman, 1989; Perry *et al.*, 1989; Chen, B. L. *et al.*, 1992; Ibarra-Molero y Sanchez-Ruiz, 1996; Martinez y Serrano, 1999).

En un gráfico de Chevron (Figura 6.12.3) el mínimo corresponde al punto en el que se cumple que el valor de la constante de equilibrio de plegamiento, K , es 1, que viene definida como el cociente entre las constantes de velocidad de plegamiento y desplegamiento a una concentración determinada de urea: $K = k_{\ddagger-D,urea} / k_{\ddagger-N,urea}$. Por tanto, a partir de este mínimo puede definirse el valor $[urea]_{1/2}$ como la concentración de urea en la que se cumple que el cambio de energía libre de Gibbs se iguala a cero según la ecuación:

$$\Delta G = -R \cdot T \cdot \ln K \quad (6.12.6)$$

A altas concentraciones de desnaturalizante (rama de la derecha de la curva de la Figura 6.12.3), cuando se produce el fenómeno de desplegamiento de la proteína, la constante que se mide es aproximadamente $k_{\ddagger-N,urea}$. A la pendiente de dicha rama se la designa por $m_{\ddagger-N}$. De la extrapolación de esta rama hasta igualar la concentración de urea al valor cero, se puede calcular el valor de la constante de desplegamiento en ausencia de urea: $k_{\ddagger-N}$. A bajas concentraciones de desnaturalizante (rama de la izquierda de la Figura 6.12.3), en la reacción de plegamiento de la proteína, se mide aproximadamente $k_{\ddagger-D,urea}$, correspondiendo el valor de la pendiente de dicha rama al parámetro $m_{\ddagger-D}$. En este caso, también se puede obtener por extrapolación el valor de la constante de velocidad de plegamiento en ausencia de urea, $k_{\ddagger-D}$.

En la Figura 6.12.4 se presenta un esquema en el que se distinguen los estados predominantes existentes en cada una de las ramas del Chevron.

La ecuación matemática general que describe ambas ramas del gráfico de Chevron se deduce a partir de la relación $k_{urea} = k_{\ddagger-N,urea} + k_{\ddagger-D,urea}$, de forma que:

$$\ln k_{urea} = \ln(k_{\ddagger-N,urea} + k_{\ddagger-D,urea}) \quad (6.12.7)$$

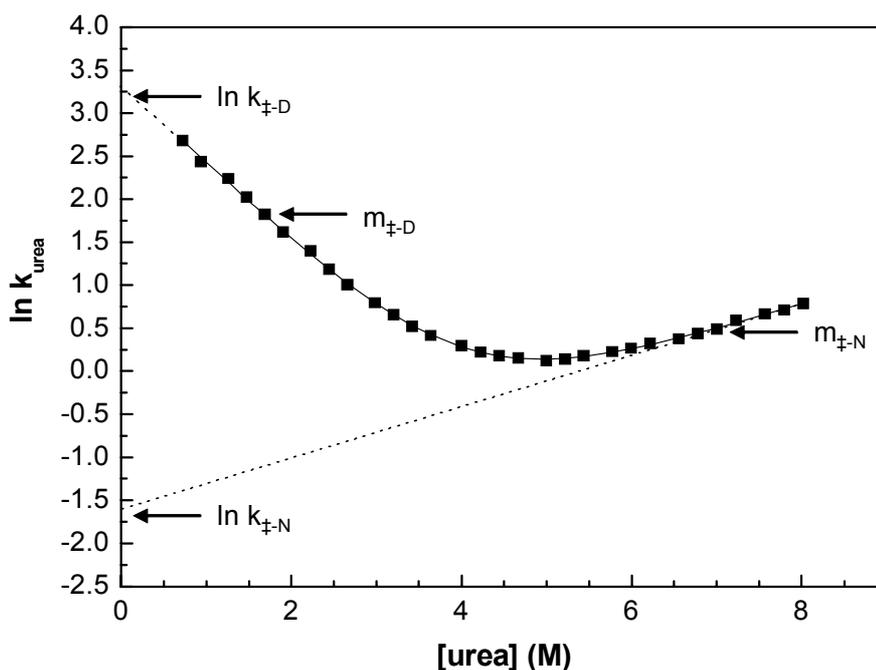


Figura 6.12.3. Gráfico Chevron de SPCp41 realizado a 25°C en tampón fosfato sódico 50 mM, pH 7.0 en el que se observan las ramas de desdoblamiento (izquierda) y plegamiento (derecha) y la zona de la transición (mínimo de la curva). Los símbolos representan los datos obtenidos experimentalmente a partir de la realización de los experimentos cinéticos en las direcciones de plegamiento y desdoblamiento, analizados según la ecuación 6.12.1. La línea continua corresponde al ajuste de los datos al modelo de dos estados descrito mediante la ecuación 6.12.10.

La dependencia lineal de $\ln k([urea])$ con la concentración de urea que se observa en las dos ramas del gráfico de Chevron, implica una dependencia lineal de la energía de Gibbs de activación con dicha concentración para una temperatura, según la ecuación de Eyring definida en esta Sección (ecuación 6.12.4). Las expresiones generales que relacionan $k_{\pm N, urea}$ y $k_{\pm D, urea}$ con la concentración de urea son las siguientes:

$$\ln k_{\pm N, urea} = \ln k_{\pm N} + m_{\pm N} \cdot [urea] \quad (6.12.8)$$

$$\ln k_{\pm D, urea} = \ln k_{\pm D} + m_{\pm D} \cdot [urea] \quad (6.12.9)$$

Con estas dos relaciones y la ecuación 3.5.6, la expresión global que define el gráfico de Chevron es la siguiente:

$$\ln k_{urea} = \ln [k_{\pm D} \cdot \exp(m_{\pm D} \cdot [urea]) + k_{\pm N} \cdot \exp(m_{\pm N} \cdot [urea])] \quad (6.12.10)$$



Figura 6.12.4. Diagramas de energía relativos a los estados nativo, desplegado y de transición, representados de forma simplificada, para el caso hipotético de una proteína que despliega siguiendo el modelo de dos estados. A baja concentración de desnaturalizante (rama de la izquierda del gráfico de Chevron, como podemos ver en la Figura 6.12.3) el estado energéticamente más favorecido es el nativo. A alta concentración de desnaturalizante, (rama de la derecha en el gráfico de Chevron, como podemos ver en la Figura 6.12.3), por el contrario, el estado más favorecido es el desnaturalizado. En la región del estado de transición (mínimo del gráfico de Chevron) las constantes de velocidad de los procesos de plegamiento y desplegamiento se igualan entre sí.

6.12.5. Estados de transición y cálculo de los valores de ϕ mediante el método de ingeniería de proteínas

Mediante el método de ingeniería de proteínas desarrollado por Fersht (1995), que se describirá a continuación es posible determinar, usando los valores de las constantes de velocidad determinadas a partir de los gráficos de Chevron, las energías de interacción que se producen en la molécula de la proteína en su estado de transición. La idea básica de este análisis consiste en determinar la extensión en la que una determinada interacción establecida por cierto aminoácido, que existe en el estado nativo, se encuentra presente en el estado de transición. En la práctica, esto se lleva a cabo realizando una mutación del aminoácido en cuestión por otro de cadena lateral más corta (a menudo Ala), con el objeto de eliminar dicha interacción. Con relación a esto, es necesario suponer *a priori* que la mutación realizada no alterará la ruta de plegamiento, que no cambiará las estructuras de los estados nativo y desnaturalizado y que la nueva cadena lateral no introducirá nuevas interacciones. Esta mutación introducida va a producir diferencias en el comportamiento cinético entre la especie mutada y la especie natural, permitiendo dichas diferencias determinar la relevancia de esta interacción en el estado de transición.

Desde un punto de vista experimental, el parámetro que determina el grado de formación de cada interacción en el estado de transición se designa por $\phi_{\ddagger-D}$, y se define como el cociente entre el cambio producido por una mutación en la energía de Gibbs de activación del proceso de plegamiento y el cambio en la energía de Gibbs de plegamiento de equilibrio. Este cociente se obtiene experimentalmente a través de las diferencias entre dichos parámetros para la especie de referencia y la especie mutada en esa determinada interacción (Fersht, 1995):

$$\phi_{\pm-D} = \Delta\Delta G_{\pm-D} / \Delta\Delta G_{N-D} \quad (6.12.11)$$

En primer lugar, para determinar el valor de $\Delta\Delta G_{\pm-D}$ hay que tener en cuenta que las constantes cinéticas medidas experimentalmente para los procesos de plegamiento y desplegamiento pueden expresarse en función de cambios en la energía de Gibbs. Para ello, se despeja el valor de K^\ddagger de la ecuación 6.12.2 y se sustituye en la ecuación 6.12.3, dando lugar a la expresión general:

$$\Delta G^\ddagger = -R \cdot T \cdot \ln(k \cdot h / k_B \cdot T) \quad \text{ó} \quad \Delta G^\ddagger = -R \cdot T \cdot (\ln k - \ln(h / k_B \cdot T)) \quad (6.12.12)$$

En el caso del proceso de plegamiento, y siguiendo la ecuación anterior, podemos escribir las expresiones correspondientes para la especie de referencia y el mutante, $(\Delta G_{\pm-D})_{ref}$ y $(\Delta G_{\pm-D})_{mut}$, como:

$$(\Delta G_{\pm-D})_{ref} = -R \cdot T \cdot \ln(k_{\pm-D,ref} - \ln(h / k_B \cdot T)) \quad (6.12.13)$$

$$(\Delta G_{\pm-D})_{mut} = -R \cdot T \cdot \ln(k_{\pm-D,mut} - \ln(h / k_B \cdot T)) \quad (6.12.14)$$

donde $k_{\pm-D,ref}$ y $k_{\pm-D,mut}$ se refieren a los valores calculados como se detalla en la Sección 6.12.4 para $k_{\pm-D}$ en el caso de la especie como referencia (ref) y del mutante de la misma (mut) respectivamente.

A partir de las dos ecuaciones anteriores, el valor de $\Delta\Delta G_{\pm-D}$ se expresa de la siguiente forma:

$$\Delta\Delta G_{\pm-D} = (\Delta G_{\pm-D})_{ref} - (\Delta G_{\pm-D})_{mut} = -R \cdot T \cdot \ln(k_{\pm-D,ref} / k_{\pm-D,mut}) \quad (6.12.15)$$

En el caso del proceso de desplegamiento, el cambio en la energía del estado de transición entre la especie de referencia y el mutante relativo al estado nativo, $\Delta\Delta G_{\pm-N}$, se expresa igualmente como:

$$\Delta\Delta G_{\pm-N} = (\Delta G_{\pm-N})_{ref} - (\Delta G_{\pm-N})_{mut} = -R \cdot T \cdot \ln(k_{\pm-N,ref} / k_{\pm-N,mut}) \quad (6.12.16)$$

En esta ecuación, los valores de $(\Delta G_{\pm-N})_{ref}$ y $(\Delta G_{\pm-N})_{mut}$ guardan relaciones equivalentes a las definidas en las ecuaciones 6.12.2 y 6.12.3 mientras que $k_{\pm-N,ref}$ y $k_{\pm-N,mut}$ corresponden a las constantes de velocidad $k_{\pm-N}$ calculadas como se describe en la Sección 6.12.4 para la referencia (ref) y el mutante (mut) respectivamente.

En segundo lugar, el valor de la energía de desestabilización inducida por la mutación, $\Delta\Delta G_{N-D}$, se define en función del cambio de energía de Gibbs entre el estado nativo y el desnaturalizado para la referencia y el mutante, como:

$$\Delta\Delta G_{N-D} = (\Delta G_{N-D})_{ref} - (\Delta G_{N-D})_{mut} = \Delta\Delta G_{\pm-D} - \Delta\Delta G_{\pm-N} \quad (6.12.17)$$

En la Figura 6.12.5 se muestra gráficamente el valor de dicha energía.

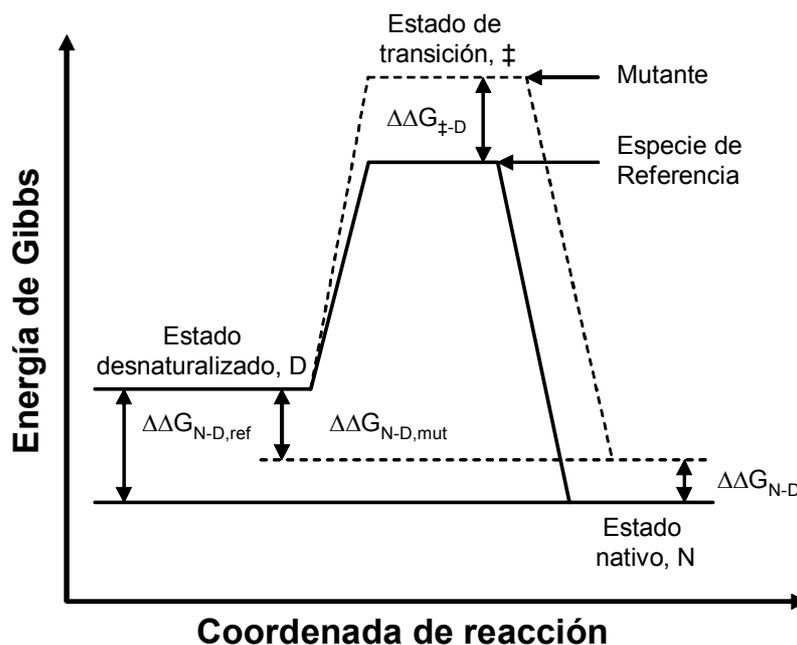


Figura 6.12.5. Diagrama de energías de Gibbs correspondiente a un caso hipotético de una proteína y un mutante de la misma que desnaturalizan siguiendo el modelo de dos estados y presentan, por tanto, un único estado de transición. Los parámetros señalados en este esquema se describen en el texto y permiten realizar el análisis de ϕ .

A partir de las definiciones dadas para las funciones anteriores se deduce que éstas pueden calcularse mediante los datos cinéticos experimentales. El parámetro $\phi_{‡-D}$, calculado según la ecuación 6.12.11, representa la pérdida de energía en el estado de transición relativa al estado nativo, inducida por una mutación puntual en la proteína.

A su vez, podemos expresar este parámetro en función del cambio de energía en el estado de transición respecto al estado desplegado debida a dicha mutación, como:

$$\phi_{‡-N} = \Delta\Delta G_{‡-N} / \Delta\Delta G_{N-D} \quad (6.12.18)$$

De esta forma, $\phi_{‡-D}$ se relaciona con el valor de $\phi_{‡-N}$ mediante la expresión:

$$\phi_{‡-D} = 1 - \phi_{‡-N} \quad (6.12.19)$$

De esta última expresión se deduce fácilmente que los valores de $\phi_{‡-D}$ y $\phi_{‡-N}$ están comprendidos entre 0 y 1. Existe una relación entre $\phi_{‡-D}$ y el grado de formación de estructura nativa durante el plegamiento en la región que envuelve a la cadena lateral del resto mutado, esto es, si está completamente plegada, parcialmente plegada o desplegada en el estado de transición. Así, teniendo en cuenta que ha de cumplirse que $0 \leq \phi_{‡-D} \leq 1$ se pueden encontrar las siguientes situaciones según el valor calculado para $\phi_{‡-D}$ (Fersht, 1995):

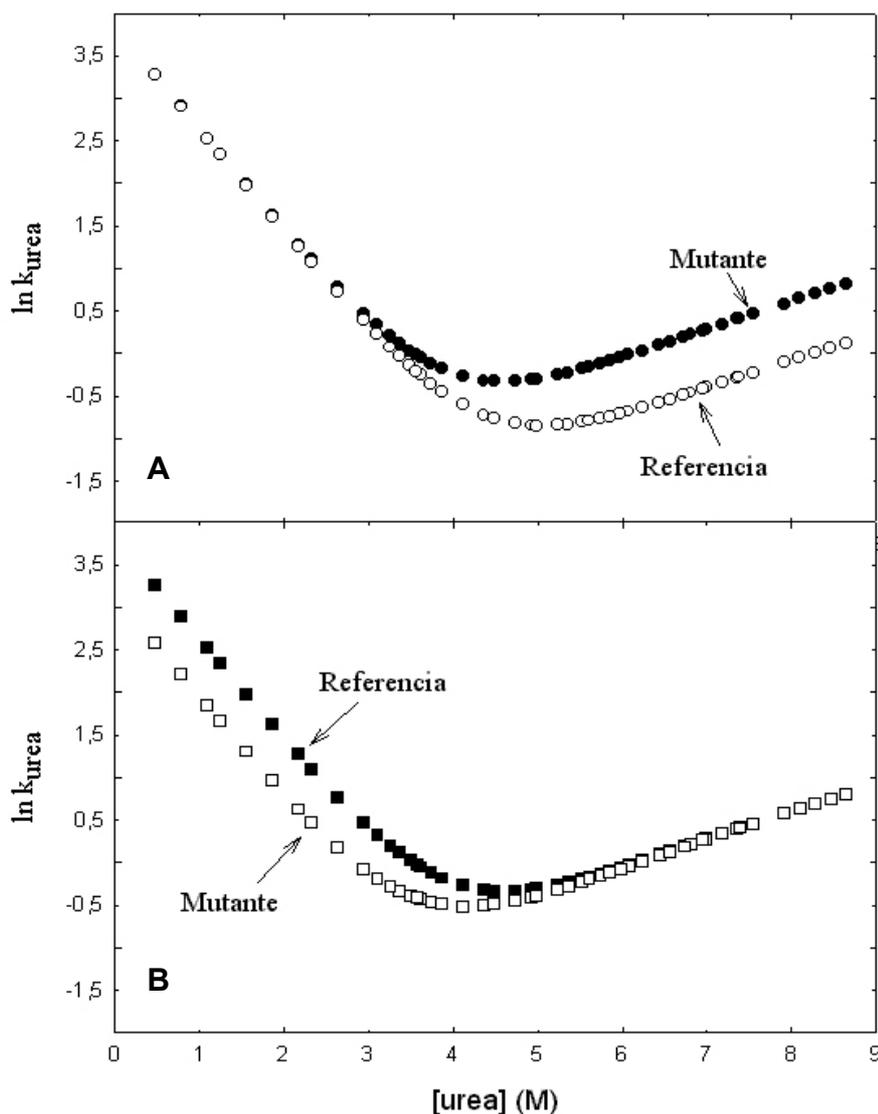


Figura 6.12.6. Gráficos de Chevron del caso hipotético de una proteína (referencia) y un mutante de la misma. **A.** $\phi_{+-D} = 0$ ($\phi_{+-N} = 1$). En este caso las ramas de la cinética en la dirección de plegamiento de la referencia (círculos rellenos) y del mutante (círculos huecos) coinciden ($k_{+-D,ref} = k_{+-D,mut}$) indicando que no ha tenido lugar la formación de la interacción en el estado de la transición respecto al estado desplegado (es decir, que hay una completa ruptura de la estructura en el estado de transición respecto al estado nativo). **B.** $\phi_{+-N} = 1$ ($\phi_{+-D} = 0$). Las ramas de desplegamiento de la referencia (cuadrados negros) y mutante (cuadrados huecos) coinciden cumpliéndose que $k_{+-N,ref} = k_{+-N,mut}$. En tal caso, la mutación afecta exclusivamente a la cinética de plegamiento, indicando que la interacción involucrada está igualmente presente en el estado de transición y en el nativo (es decir, que existe una completa formación de la estructura en el estado de transición respecto al estado desnaturalizado).

En primer lugar, si $\phi_{+-D} = 0$ (ó $\phi_{+-N} = 1$) en la ecuación 6.12.11 se cumple que $\Delta\Delta G_{+-D} = 0$ y, si esto es así, según la ecuación 6.12.15 se da que $k_{+-D,ref} = k_{+-D,mut}$. Según esta última relación, se observará en el gráfico de Chevron correspondiente que las ramas de plegamiento de la referencia y mutante coinciden, indicando que la mutación no afecta a la cinética de plegamiento (Figura 6.12.6.A) y por tanto, la interacción que implica la mutación no estará involucrada en la formación del estado de transición. Es decir, la región de la proteína que contiene la mutación está completamente desplegada en el estado de transición.

En segundo lugar, si $\phi_{\ddagger-D} = 1$ (ó $\phi_{\ddagger-N} = 0$) según la relación (6.12.17), se obtiene que $\phi_{\ddagger-N} = 0$ y si esto es así, en la ecuación 6.12.17 se cumple que $\Delta\Delta G_{\ddagger-N} = 0$. Según esto, en la ecuación 6.12.16 se tiene que dar que $k_{\ddagger-N,ref} = k_{\ddagger-N,mut}$. Esto último implica que en los gráficos Chevron del mutante y la referencia las ramas del desplegamiento coinciden, es decir, que la mutación afecta exclusivamente a las cinéticas de plegamiento cuyas constantes, $k_{\ddagger-D,ref}$, $k_{\ddagger-D,mut}$ adquieren valores distintos y sus ramas no coinciden (Figura 6.12.6.B). Dicho de otra forma, en este caso la región de la proteína que contiene la mutación está expuesta al disolvente en igual extensión en los estados de transición y desnaturalizado (está completamente desplegada en estos dos estados) tanto en el mutante como en la referencia. Es decir, la interacción está igualmente presente en el estado de transición y en el estado nativo.

Puede ser posible también que $\phi_{\ddagger-D}$ adopte valores intermedios entre los dos límites descritos: $0 < \phi_{\ddagger-D} < 1$. A pesar de que la relación entre $\phi_{\ddagger-D}$ y la formación de la estructura a lo largo de la ruta de plegamiento no sigue una dependencia de tipo lineal, puede decirse que valores elevados de $\phi_{\ddagger-D}$ indican que la estructura está prácticamente formada aunque no estén presentes todas las interacciones existentes en el estado nativo. La interpretación de los valores intermedios de $\phi_{\ddagger-D}$ resulta compleja, debido a la existencia de una serie de razones que dan lugar a que este parámetro adopte valores comprendidos en el intervalo anteriormente definido (Fersht, 1995). Algunas de estas causas se resumen a continuación:

En primer lugar, la formación de interacciones de carácter débil de tipo no covalente en un único estado o la existencia de rutas de plegamiento paralelas que dan lugar a una mezcla de estados de transición, en algunos de los cuales las interacciones están completamente formadas mientras que en otros no. Realizando una serie de mutaciones en los mismos elementos de estructura, se ha comprobado que estos valores intermedios de $\phi_{\ddagger-D}$ no resultan de rutas paralelas de plegamiento con una mezcla de estados de transición, unos completamente formados y otros desplegados, sino de la formación de interacciones débiles en un único estado de transición (Fersht *et al.*, 1994; Otzen *et al.*, 1994; Martínez y Serrano, 1999).

Otra causa posible de los valores intermedios de $\phi_{\ddagger-D}$ pueden deberse a que la cadena lateral del aminoácido mutado es capaz de interactuar con diferentes partes de la estructura, que a su vez siguen distintas rutas de desplegamiento (Matouschek *et al.*, 1992; Serrano *et al.*, 1992).

En tercer lugar, puede producirse una distorsión en la estructura del mutante en el estado plegado que varía en el estado de transición, dando lugar a una nueva contribución a la energía de Gibbs introducida por la mutación y debida a esta reorganización en la estructura. La existencia de estos artefactos se comprobó mediante el análisis de una serie de mutaciones en una región determinada de la proteína barnasa, llegándose a la conclusión de que en el estado de transición de su ruta de desplegamiento los valores de $\phi_{\ddagger-D}$ varían gradualmente dentro del intervalo en el que se ha definido este parámetro, en función a la región donde estén situados dichos residuos (Fersht, 1993).

Además de los casos descritos hasta el momento sobre los valores que $\phi_{\ddagger-D}$ (ó $\phi_{\ddagger-N}$) puede tomar, hay que añadir que valores negativos de este parámetro (ó valores por encima de 1 de $\phi_{\ddagger-D}$) implican que $\Delta\Delta G_{\ddagger-D} > \Delta\Delta G_{N-D}$, indicando que la mutación estabiliza el estado de transición, desestabilizando el estado nativo y, más aún, que la

cadena lateral del resto involucrado en la mutación está originando contactos en el estado de transición que no tienen lugar en el estado nativo (contactos no nativos).

CAPÍTULO 7: **RESUMEN Y CONCLUSIONES**

CAPÍTULO 7: RESUMEN Y CONCLUSIONES

De forma general, el trabajo presentado en esta Memoria tenía como objetivo fundamental el diseño de proteínas quiméricas que integraran en una única cadena polipeptídica la secuencia de un dominio SH3 y la de un ligando peptídico y que reprodujeran en su estructura terciaria las interacciones típicas entre los dominios SH3 y sus ligandos ricos en prolina. Este objetivo se consiguió mediante la fusión del permutante circular S19P20s y el deca péptido p41 a través de una secuencia de conexión de tres residuos, obteniéndose una proteína quimérica denominada SPCp41. Se caracterizó su estructura a alta resolución, su estabilidad térmica, la termodinámica de las interacciones intramoleculares SH3:ligando en su seno, así como su mecanismo de plegamiento, utilizando un conjunto de técnicas como la resonancia magnética nuclear (RMN), la calorimetría diferencial de barrido (CDB), el dicroísmo circular (DC) y la espectroscopia de fluorescencia y técnicas cinéticas de flujo detenido.

De forma más detallada el trabajo se resume en:

- i) Se clonaron y purificaron un conjunto de proteínas quiméricas mediante la fusión entre el permutante circular S19P20s del dominio Spc-SH3 y la secuencia peptídica del ligando p41 utilizando varias secuencias de conexión cortas. Los estudios se centraron en la proteína quimérica que resultó ser más estable térmicamente, y a la que denominamos SPCp41, la cual se caracterizó mediante calorimetría diferencial de barrido, dicroísmo circular y fluorescencia en comparación con el permutante S19P20s. Estos resultados han sido recogidos en una publicación científica en la revista *FEBS Letters* (Martin-Sierra *et al.*, 2003).
- ii) Para continuar nuestra investigación se determinó la estructura tridimensional a alta resolución de la quimera SPCp41, mediante técnicas de RMN heteronuclear multidimensional, lo que nos ha permitido una mejor interpretación de las magnitudes termodinámicas en los estudios posteriores y ayudará a optimizar las interacciones en el sitio de unión utilizando esta información estructural para futuros procesos de diseño de ligandos. Estos resultados se han publicado en la revista *FEBS Letters* (Candel *et al.*, 2007).
- iii) Con el fin de demostrar que es posible utilizar la quimera SPCp41 como una herramienta apropiada para evaluar la energética de la unión entre dominios SH3 y ligandos ricos en prolina, y para elucidar, desde un punto de vista termodinámico, la importancia en la afinidad de unión de cada uno de los residuos prolina del ligando p41, se ha sustituyó cada prolina por alanina en la quimera SPCp41 y se caracterizó la estabilidad estructural de la quimera y de estos mutantes mediante fluorescencia y DC. Se llevó a cabo el análisis profundo del desplegamiento térmico de la quimera SPCp41 y los mutantes Pro-Ala mediante CDB, según un modelo de tres estados que nos permitió obtener un juego consistente de parámetros termodinámicos tanto para la estabilidad del permutante como para la interacción intramolecular con el ligando p41. Este análisis se completó con un estudio comparativo de las propiedades estructurales y dinámicas mediante RMN de la quimera SPCp41 y uno de los mutantes. Los resultados de la caracterización termodinámica de la estabilidad estructural de la quimera y los mutantes Pro-Ala han sido objeto de una publicación en la revista *Journal of Molecular Biology* (Candel *et al.*).

- iv) Por último, se caracterizó la cinética de plegamiento de la quimera SPCp41 y sus mutantes Pro-Ala, en comparación con el permutante circular S19P20s para analizar en profundidad su mecanismo de plegamiento.

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

1. Hemos conseguido diseñar, clonar y purificar un conjunto de proteínas quiméricas de fusión entre el permutante circular S19P20s del dominio Spc-SH3 y el ligando peptídico p41, eligiéndose para un estudio en mayor profundidad la quimera que resultó más estable térmicamente, a la que se denominó SPCp41.
2. La caracterización preliminar de su estructura y estabilidad mediante CDB y espectroscopias de dicroísmo circular y fluorescencia bajo distintas condiciones experimentales nos permitió confirmar que estas proteínas quiméricas poseen una mayor estabilidad que el permutante circular debido a la interacción intramolecular que se establece entre la secuencia del ligando p41 y el sitio de unión característico del dominio SH3.
3. La estructura tridimensional en disolución de la quimera SPCp41, obtenida mediante RMN, demuestra que esta quimera imita correctamente en su estructura terciaria las interacciones típicas que se observan en los complejos entre dominios SH3 y péptidos ricos en prolina y es muy similar al complejo formado por el dominio Spc-SH3 y el ligando p41, con una conformación de hélice de poliprolina II para p41 y una inmovilización parecida de los residuos aromáticos en la interfase de unión.
4. El análisis estructural de la quimera SPCp41 y de los mutantes Pro-Ala mediante CD y fluorescencia, junto con las medidas de relajación por RMN indicó la existencia de flexibilidad conformacional en la superficie de unión, que se ve fuertemente acentuada por las diferentes mutaciones Pro-Ala que debilitan la interacción intramolecular SH3:p41.
5. Las magnitudes termodinámicas resultantes del análisis de los termogramas de CDB utilizando un modelo de desplegamiento de tres estados nos permitió clasificar las prolina del péptido según su importancia en la interacción, que está de acuerdo con el modelo de Lim para la interacción entre dominios SH3 y péptidos ricos en prolina.
6. Los parámetros termodinámicos de la interacción intramolecular SH3:p41 son similares a los que normalmente se obtienen para este tipo de interacción, tanto en su magnitud como en su signo, lo que confirma la validez de la quimera como una herramienta de diseño.
7. Se observaron variaciones apreciables en la entalpía de la interacción SH3:p41 con las diferentes mutaciones Pro-Ala, compensadas parcialmente por variaciones entrópicas. Esta compensación entalpía-entropía aparece como consecuencia de un aumento de la flexibilidad conformacional a medida que se produce un debilitamiento de las interacciones en la interfase de unión.
8. La caracterización de la cinética de plegamiento de la quimera SPCp41 y de sus mutantes Pro-Ala puso de manifiesto la existencia de dos fases cinéticas en el plegamiento y en el desplegamiento, una rápida y otra lenta, siendo la fase rápida representativa del plegamiento y del desplegamiento intrínseco de la quimera SPCp41 y de los mutantes Pro-Ala, mientras que la fase lenta se puede atribuir a un proceso de isomerización de prolina, posiblemente acoplado a la formación de algún estado intermedio minoritario.
9. La magnitud de los parámetros de activación de la quimera SPCp41, similar a los obtenidos para el permutante circular S19P20s, y los valores de $\phi_{\pm D}$

cercanos a cero indican la escasa o nula contribución energética del segmento p41 al estado de transición de plegamiento, manteniéndose por consiguiente el núcleo de plegamiento en el dominio SH3, al que posteriormente se uniría el péptido p41 a través de un proceso de unión intramolecular, similar al descrito para el péptido libre, lo que argumenta claramente a favor de la efectividad de esta quimera de diseño para abordar estudios de unión SH3-péptido.

BIBLIOGRAFÍA

BIBLIOGRAFÍA

- Abragam, A. (1961). The principles of Nuclear Magnetism, Clarendon Press, Oxford Akke.
- Aghazadeh, B. y Rosen, M. K. (1999). "Ligand recognition by SH3 and WW domains: the role of N-alkylation in PPII helices." Chemistry & Biology **6**(9): R241-246.
- Agrawal, V. y Kishan, K. V. R. (2002). "Promiscuous binding nature of SH3 domains to their target proteins." Protein and Peptide Letters **9**(3): 185-193.
- Alexandropoulos, K., Cheng, G. y Baltimore, D. (1995). "Proline-rich sequences that bind to Src homology 3 domains with individual specificities." Proc. Natl. Acad. Sci. U S A **92**(8): 3110-3114.
- Anderson, D. R. y Weber, G. R. (1969). "Fluorescence Polarization of the complexes of 1-anilino-8-naphtalene-sulfonate with bovine serum albumin. Evidence for preferential orientation of the ligand." Biochemistry **8**: 371-377.
- Andreotti, A. H., Bunnell, S. C., Feng, S., Berg, L. J. y Schreiber, S. L. (1997). "Regulatory intramolecular association in a tyrosine kinase of the Tec family." Nature **385**(6611): 93-97.
- Argos, P. (1988). "An Investigation of Protein Subunit and Domain Interfaces." Protein Engineering **2**(2): 101-113.
- Arold, S., O'Brien, R., Franken, P., Strub, M. P., Hoh, F., Dumas, C. y Ladbury, J. E. (1998). "RT loop flexibility enhances the specificity of Src family SH3 domains for HIV-1 Nef." Biochemistry **37**(42): 14683-14691.
- Arold, S. T., Ulmer, T. S., Mulhern, T. D., Werner, J. M., Ladbury, J. E., Campbell, I. D. y Noble, M. E. M. (2001). "The role of the Src homology 3-Src homology 2 interface in the regulation of Src kinases." Journal of Biological Chemistry **276**(20): 17199-17205.
- Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. y Struhl, K. (1992). Short protocols in molecular biology. New York (USA), Wiley & Sons.
- Ayrapetov, M. K., Wang, Y. H., Lin, X. F., Gu, X. F., Parang, K. y Sun, G. Q. (2006). "Conformational basis for SH2-Tyr(P)(527) binding in Src inactivation." Journal of Biological Chemistry **281**(33): 23776-23784.
- Bach, D. y Chapman, D. (1980). Biological Microcalorimetry. Press, A. London (UK): 265.
- Bachmann, A. y Kiefhaber, T. (2005). Kinetic Mechanisms in Protein Folding. Protein Folding Handbook. Part I. Buchner, J. y Kiefhaber, T., WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim. **I**: 379-410.
- Bachmann, A., Segel, D. y Kiefhaber, T. (2002). "Test for cooperativity in the early kinetic intermediate in lysozyme folding." Biophysical Chemistry **96**(2-3): 141-151.
- Ball, L. J., Kuhne, R., Schneider-Mergener, J. y Oschkinat, H. (2005). "Recognition of Proline-Rich Motifs by Protein-Protein-Interaction Domains." Angew. Chem. Int. Ed. Engl. **44**(19): 2852-2869.
- Barila, D. y Superti-Furga, G. (1998). "An intramolecular SH3-domain interaction regulates c-Abl activity." Nature Genetics **18**(3): 280-282.
- Bax, A. y Davis, D. G. (1985). "Mlev-17-Based Two-Dimensional Homonuclear Magnetization Transfer Spectroscopy." Journal of Magnetic Resonance **65**(2): 355-360.
- Beasty, A. M., Hurle, M. R., Manz, J. T., Stackhouse, T., Onuffer, J. J. y Matthews, C. R. (1986). "Effects of the Phenylalanine-22- Leucine, Glutamic Acid-49- Methionine, Glycine-234- Aspartic Acid, and Glycine-234- Lysine Mutations on

- the Folding and Stability of the Alpha-Subunit of Tryptophan Synthase from Escherichia-Coli." Biochemistry **25**(10): 2965-2974.
- Berjanskii, M. V., Neal, S. y Wishart, D. S. (2006). "PREDITOR: a web server for predicting protein torsion angle restraints." Nucleic Acids Research **34**(Web Server issue): W63-9.
- Bevington, P. R. y Robinson, D. K. (2003). Data Reduction and Error Analysis for the Physical Sciences, McGraw-Hill, New York.
- Blanco, F. J., Ortiz, A. R. y Serrano, L. (1997). "¹H and ¹⁵N NMR assignment and solution structure of the SH3 domain of spectrin: comparison of unrefined and refined structure sets with the crystal structure." Journal of Biomolecular NMR **9**(4): 347-357.
- Bodenhausen, G. y Ruben, D. J. (1980). "Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy." Chemical Physics Letters **69**(1): 185-189.
- Boggon, T. J. y Eck, M. J. (2004). "Structure and regulation of Src family kinases." Oncogene **23**(48): 7918-7927.
- Brabek, J., Mojzita, D., Novotny, M., Puta, F. y Folk, P. (2002). "The SH3 domain of Src can downregulate its kinase activity in the absence of the SH2 domain-pY527 interaction." Biochemical and Biophysical Research Communications **296**(3): 664-670.
- Bradley, C. M., Barrick, D., Mello, C. C. y Tripp, K. W. (2005). "Effect of multiple prolyl isomerization reactions on the stability and folding kinetics of the notch ankyrin domain: experiment and theory
Experimental characterization of the folding kinetics of the notch ankyrin domain." Journal of Molecular Biology **352**(2): 253-265.
- Brahms, S. y Brahms, J. (1980). "Determination of Protein Secondary Structure in Solution by Vacuum Ultraviolet Circular-Dichroism." Journal of Molecular Biology **138**(2): 149-178.
- Brandts, J. F., Halvorson, H. R. y Brennan, M. (1975). "Consideration of Possibility That Slow Step in Protein Denaturation Reactions Is Due to Cis-Trans Isomerism of Proline Residues." Biochemistry **14**(22): 4953-4963.
- Brandts, J. F. y Lin, L. N. (1990). "Study of Strong to Ultratight Protein Interactions Using Differential Scanning Calorimetry." Biochemistry **29**(29): 6927-6940.
- Brannetti, B., Via, A., Cestra, G., Cesareni, G. y Helmer-Citterich, M. (2000). "SH3-SPOT: an algorithm to predict preferred ligands to different members of the SH3 gene family." Journal of Molecular Biology **298**(2): 313-328.
- Brasher, B. B., Roumiantsev, S. y Van Etten, R. A. (2001). "Mutational analysis of the regulatory function of the c-Abl Src homology 3 domain." Oncogene **20**(53): 7744-7752.
- Candel, A. M., Conejero-Lara, F., Martinez, J. C., van Nuland, N. A. J. y Bruix, M. (2007). "The high-resolution NMR structure of a single-chain chimeric protein mimicking a SH3-peptide complex." FEBS Letters **581**(4): 687-692.
- Candel, A. M., Nuland, N. A. J. v., Martin-Sierra, F. M., Martinez, J. C. y Conejero-Lara, F. "Analysis of the thermodynamics of binding of an SH3 domain to proline-rich peptides using a chimeric fusion protein." Journal of Molecular Biology
Aceptado para su publicación: doi: 10.1016/j.jmb.2007.11.060.
- Cantor, C. R. y Schimmel, P. R. (1980). Techniques for the Study of Biological Structure and Function. Biophysical Chemistry, W.H. Freeman. **2**.
- Capaldi, A. P., Kleanthous, C. y Radford, S. E. (2002). "Im7 folding mechanism: misfolding on a path to the native state." Nature Structural Biology **9**(3): 209-216.
- Casares, S., Eiso, A. B., Eshuis, H., Lopez-Mayorga, O., van Nuland, N. A. J. y Conejero-Lara, F. (2007). "The high-resolution NMR structure of the R21A Spc-

- SH3 : P41 complex: Understanding the determinants of binding affinity by comparison with Abl-SH3." BMC Structural Biology **7**(22): doi:10.1186/1472-6807-7-22.
- Casares, S., Sadqi, M., Lopez-Mayorga, O., Martinez, J. C. y Conejero-Lara, F. (2003). "Structural cooperativity in the SH3 domain studied by site-directed mutagenesis and amide hydrogen exchange." FEBS Letters **539**(1-3): 125-130.
- Cesareni, G., Panni, S., Nardelli, G. y Castagnoli, L. (2002). "Can we infer peptide recognition specificity mediated by SH3 domains?" FEBS Letters **513**(1): 38-44.
- Chaffotte, A. F., Cadieux, C., Guillou, Y. y Goldberg, M. E. (1992). "A Possible Initial Folding Intermediate - the C-Terminal Proteolytic Domain of Tryptophan Synthase-Beta Chains Folds in Less Than 4 Milliseconds into a Condensed State with Non-Native-Like Secondary Structure." Biochemistry **31**(17): 4303-4308.
- Chen, B. L., Baase, W. A., Nicholson, H. y Schellman, J. A. (1992). "Folding Kinetics of T4 Lysozyme and 9 Mutants at 12-Degrees-C." Biochemistry **31**(5): 1464-1476.
- Chen, B. L., Baase, W. A. y Schellman, J. A. (1989). "Low-Temperature Unfolding of a Mutant of Phage-T4 Lysozyme .2. Kinetic Investigations." Biochemistry **28**(2): 691-699.
- Chen, B. L. y Schellman, J. A. (1989). "Low-Temperature Unfolding of a Mutant of Phage-T4 Lysozyme .1. Equilibrium Studies." Biochemistry **28**(2): 685-691.
- Chen, J. K., Lane, W. S., Brauer, A. W., Tanaka, A. y Schreiber, S. L. (1993). "Biased combinatorial libraries: novel ligands for the SH3 domain of phosphatidylinositol 3-kinase." Journal of the American Chemical Society **115**(26): 12591-12592.
- Chen, R. F. (1967). "Fluorescence Quantum Yield of tryptophan and tyrosine." Analytical Letters **1**(1): 35-42.
- Clore, G. M. y Gronenborn, A. M. (1991). "Applications of three- and four-dimensional heteronuclear NMR spectroscopy to protein structure determination." Progress in Nuclear Magnetic Resonance Spectroscopy **23**(1): 43-92.
- Clore, G. M., Szabo, A., Bax, A., Kay, L. E., Driscoll, P. C. y Gronenborn, A. M. (1990). "Deviations from the simple two-parameter model-free approach to the interpretation of nitrogen-15 nuclear magnetic relaxation of proteins." Journal of the American Chemical Society **112**(12): 4989-4991.
- Cobos, E. S. (2002). Tesis Doctoral: Estabilidad, plegamiento y estructura de la proteína circular AS-48 y de mutantes del dominio SH3. Departamento de Química Física, Universidad de Granada: 253.
- Cobos, E. S., Candel, A. M. y Martinez, J. C. "An error analysis for two-state protein folding kinetic parameters and phi values: progress towards precision by exploring pH dependencies on Leffler plots." Biophysical Journal **Aceptado para su publicación**.
- Cobos, E. S., Filimonov, V. V., Vega, M. C., Mateo, P. L., Serrano, L. y Martinez, J. C. (2003). "A thermodynamic and kinetic analysis of the folding pathway of an SH3 domain entropically stabilised by a redesigned hydrophobic core." Journal of Molecular Biology **328**(1): 221-233.
- Combs, A. P., Kapoor, T. M., Feng, S., Chen, J. K., Daude-Snow, L. F. y Schreiber, S. L. (1996). "Protein Structure-Based Combinatorial Chemistry: Discovery of Non-Peptide Binding Elements to Src SH3 Domain." Journal of the American Chemical Society **118**(1): 287-288.
- Compton, L. A. y Johnson, W. C. (1986). "Analysis of Protein Circular-Dichroism Spectra for Secondary Structure Using a Simple Matrix Multiplication." Analytical Biochemistry **155**(1): 155-167.
- Condie, C. C. y Quay, S. C. (1983). "Molecular Studies of Protein-Lipid Interactions .3. Conformational Studies of Aqueous Melittin - Characteristics of a Fluorescent-Probe Binding-Site." Journal of Biological Chemistry **258**(13): 8231-8234.

- Cooper, A., Johnson, C. M., Lakey, J. H. y Nollmann, M. (2001). "Heat does not come in different colours: entropy-enthalpy compensation, free energy windows, quantum confinement, pressure perturbation calorimetry, solvation and the multiple causes of heat capacity effects in biomolecular interactions." *Biophysical Chemistry* **93**(2-3): 215-230.
- Creamer, T. P. (1998). "Left-handed polyproline II helix formation is (very) locally driven." *Proteins-Structure Function and Genetics* **33**(2): 218-226.
- Creighton, T. E. (1993). *Proteins: Structures and Molecular Properties*. New York (USA), WH Freeman and Co.
- Crowley, P., Ubbink, M. y Otting, G. (2000). "Phi; Angle Restraints in Protein Backbones from Dipole-Dipole Cross-Correlation between ^1HN - ^{15}N and ^1HN - ^1H ; Vectors." *Journal of the American Chemical Society* **122**(12): 2968-2969.
- Dalgarno, D. C., Botfield, M. C. y Rickles, R. J. (1997). "SH3 domains and drug design: Ligands, structure, and biological function." *Biopolymers* **43**(5): 383-400.
- Daniel, E. y Weber, G. (1966). "Cooperative Effects in Binding by Bovine Serum Albumin .I. Binding of 1-Anilino-8-Naphthalenesulfonate. Fluorimetric Titrations." *Biochemistry* **5**(6): 1893-1896.
- Daquino, J. A., Gomez, J., Hilser, V. J., Lee, K. H., Amzel, L. M. y Freire, E. (1996). "The magnitude of the backbone conformational entropy change in protein folding." *Proteins-Structure Function and Genetics* **25**(2): 143-156.
- De Los Rios, M. A., Muralidhara, B. K., Wildes, D., Sosnick, T. R., Marqusee, S., Wittung-Stafshede, P., Plaxco, K. W. y Ruczinski, I. (2006). "On the precision of experimentally determined protein folding rates and phi-values." *Protein Science* **15**(3): 553-563.
- Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. y Bax, A. (1995). "NMRPipe - a Multidimensional Spectral Processing System Based on Unix Pipes." *Journal of Biomolecular NMR* **6**(3): 277-293.
- Demchenko, A. P. (1981). *Ultraviolet Spectroscopy of Proteins*. New York (USA), Springer-Verlag.
- Dosset, P., Hus, J. C., Blackledge, M. y Marion, D. (2000). "Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data." *Journal of Biomolecular NMR* **16**(1): 23-28.
- Duerst, R. y Merbach, A. (1965). "Accurate NMR Temperature Measurements." *Review of Scientific Instruments* **36**(12): 1896-1897.
- Dunitz, J. D. (1995). "Win Some, Lose Some - Enthalpy-Entropy Compensation in Weak Intermolecular Interactions." *Chemistry & Biology* **2**(11): 709-712.
- Eaton, W. A., Munoz, V., Thompson, P. A., Chan, C. K. y Hofrichter, J. (1997). "Submillisecond kinetics of protein folding." *Current Opinion in Structural Biology* **7**(1): 10-14.
- Eftink, M. R. (1991). "Fluorescence Techniques for Studying Protein-Structure." *Methods of Biochemical Analysis* **35**: 127-205.
- Eftink, M. R. (1994). "The Use of Fluorescence Methods to Monitor Unfolding Transitions in Proteins." *Biophysical Journal* **66**(2): 482-501.
- Eftink, M. R., Anusiem, A. C. y 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**(16): 3884-3896.
- Emsley, L. (1994). Selective Pulses and Their Applications to Assignment and Structure Determination in Nuclear-Magnetic-Resonance. *Nuclear Magnetic Resonance, Pt C*. **239**: 207-246.

- Fanghanel, J. y Fischer, G. (2003). "Thermodynamic characterization of the interaction of human cyclophilin 18 with cyclosporin A." *Biophysical Chemistry* **100**(1-3): 351-366.
- Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Formankay, J. D. y Kay, L. E. (1994). "Backbone Dynamics of a Free and a Phosphopeptide-Complexed Src Homology-2 Domain Studied by N-15 NMR Relaxation." *Biochemistry* **33**(19): 5984-6003.
- Feng, S., Chen, J. K., Yu, H., Simon, J. A. y 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**(5188): 1241-1247.
- Feng, S., Kapoor, T. M., Shirai, F., Combs, A. P. y Schreiber, S. L. (1996). "Molecular basis for the binding of SH3 ligands with non-peptide elements identified by combinatorial synthesis." *Chemistry & Biology* **3**(8): 661-670.
- Ferreon, J. C. y Hilser, V. J. (2003). "The effect of the polyproline II (PPII) conformation on the denatured state entropy." *Biophysical Journal* **84**(2): 482A-483A.
- Ferreon, J. C. y Hilser, V. J. (2004). "Thermodynamics of binding to SH3 domains: The energetic impact of polyproline II (P-II) helix formation." *Biochemistry* **43**(24): 7787-7797.
- Fersht, A. R. (1993). "Protein-Folding and Stability - the Pathway of Folding of Barnase." *FEBS Letters* **325**(1-2): 5-16.
- Fersht, A. R. (1995). "Characterizing Transition-States in Protein-Folding - an Essential Step in the Puzzle." *Current Opinion in Structural Biology* **5**(1): 79-84.
- Fersht, A. R. (2000). "Transition-state structure as a unifying basis in protein-folding mechanisms: contact order, chain topology, stability, and the extended nucleus mechanism." *Proc. Natl. Acad. Sci. U S A* **97**(4): 1525-1529.
- Fersht, A. R., Itzhaki, L. S., elMasry, N. F., Matthews, J. M. y Otzen, D. E. (1994). "Single versus parallel pathways of protein folding and fractional formation of structure in the transition state." *Proc. Natl. Acad. Sci. U S A* **91**(22): 10426-10429.
- Fersht, A. R. y Sato, S. (2004). "Phi-value analysis and the nature of protein-folding transition states." *Proc. Natl. Acad. Sci. U S A* **101**(21): 7976-7981.
- Filimonov, V. V., Azuaga, A. I., Viguera, A. R., Serrano, L. y Mateo, P. L. (1999). "A thermodynamic analysis of a family of small globular proteins: SH3 domains." *Biophysical Chemistry* **77**(2-3): 195-208.
- Fischer, G. (1994). "Peptidyl-prolyl cis/trans isomerases and their effectors." *Angew. Chem. Int. Ed. Engl.* **33**: 1415-1436.
- Fisher, H. F., Colen, A. H. y Medary, R. T. (1981). "Temperature-dependent delta C_{0p} generated by a shift in equilibrium between macrostates of an enzyme." *Nature* **292**(5820): 271-272.
- Freire, E., van Osdol, W. W., Mayorga, O. L. y Sanchez-Ruiz, J. M. (1990). "Calorimetrically determined dynamics of complex unfolding transitions in proteins." *Annual Review of Biophysics and Bioengineering* **19**: 159-188.
- Freund, C., Kuhne, R., Park, S., Thiemke, K., Reinherz, E. L. y Wagner, G. (2003). "Structural investigations of a GYF domain covalently linked to a proline-rich peptide." *Journal of Biomolecular NMR* **27**(2): 143-149.
- Freund, C., Kuhne, R., Yang, H. L., Park, S., Reinherz, E. L. y Wagner, G. (2002). "Dynamic interaction of CD2 with the GYF and the SH3 domain of compartmentalized effector molecules." *EMBO Journal* **21**(22): 5985-5995.
- Garbay, C., Liu, W. Q., Vidal, M. y Roques, B. P. (2000). "Inhibitors of Ras signal transduction as antitumor agents." *Biochem Pharmacol* **60**(8): 1165-1169.
- Garel, J. R. y Baldwin, R. L. (1973). "Both the fast and slow refolding reactions of ribonuclease A yield native enzyme." *Proc. Natl. Acad. Sci. U S A* **70**(12): 3347-3351.

- Garel, J. R. y Baldwin, R. L. (1975). "A physical difference between the fast- and slow-refolding forms of nitrotyrosyl ribonuclease A: the pK values of the nitrotyrosyl groups." Journal of Molecular Biology **94**(4): 621-632.
- Garel, J. R., Nall, B. T. y Baldwin, R. L. (1976). "Guanidine-unfolded state of ribonuclease A contains both fast- and slow-refolding species." Proc. Natl. Acad. Sci. U S A **73**(6): 1853-1857.
- Geen, H. y Freeman, R. (1991). "Band-Selective Radiofrequency Pulses." Journal of Magnetic Resonance **93**(1): 93-141.
- Gill, S. C. y Vonhippel, P. H. (1989). "Calculation of Protein Extinction Coefficients from Amino-Acid Sequence Data." Analytical Biochemistry **182**(2): 319-326.
- Gmeiner, W. H., Xu, J. Z., Horita, D. A., Smithgall, T. E., Engen, J. R., Smith, D. L. y Byrd, R. A. (2001). "Intramolecular binding of a proximal PPII helix to an SH3 domain in the fusion protein SH3Hck : PPIIhGAP." Cell Biochemistry and Biophysics **35**(2): 115-126.
- Goddard, T. D. y Kneller, D. G. SPARKY 3. San Francisco, University of California.
- Golbik, R., Fischer, G. y Fersht, A. R. (1999). "Folding of barstar C40A/C82A/P27A and catalysis of the peptidyl-prolyl cis/trans isomerization by human cytosolic cyclophilin (Cyp18)." Protein Science **8**(7): 1505-1514.
- Guijarro, J. I., Morton, C. J., Plaxco, K. W., Campbell, I. D. y Dobson, C. M. (1998). "Folding kinetics of the SH3 domain of PI3 kinase by real-time NMR combined with optical spectroscopy." Journal of Molecular Biology **276**(3): 657-667.
- Guntert, P. (2004). "Automated NMR structure calculation with CYANA." Methods in Molecular Biology **278**: 353-378.
- Guntert, P., Braun, W. y Wuthrich, K. (1991). "Efficient Computation of 3-Dimensional Protein Structures in Solution from Nuclear-Magnetic-Resonance Data Using the Program Diana and the Supporting Programs Caliba, Habas and Glomsa." Journal of Molecular Biology **217**(3): 517-530.
- Hagel, P., Gerding, J. J. T., Fieggen, W. y Bloemend.H (1971). "Cyanate Formation in Solutions of Urea .1. Calculation of Cyanate Concentrations at Different Temperature and Ph." Biochimica Et Biophysica Acta **243**(3): 366-369.
- Hamburger, J. B., Ferreón, J. C., Whitten, S. T. y Hilser, V. J. (2004). "Thermodynamic mechanism and consequences of the polyproline II (P-II) structural bias in the denatured states of proteins." Biochemistry **43**(30): 9790-9799.
- Herrmann, T., Guntert, P. y Wuthrich, K. (2002). "Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA." Journal of Molecular Biology **319**(1): 209-227.
- Hilser, V. J., Gomez, J. y Freire, E. (1996). "Enthalpy change in protein folding and binding: Refinement of parameters for structure-based calculations." Proteins-Structure Function and Genetics **26**(2): 123-133.
- Hubbard, S. J. y Thornton, J. M. (1993). NACCESS, Computer Program. London: Department of Biochemistry and Molecular Biology, University College.
- Hurle, M. R., Anderson, S. y Kuntz, I. D. (1991). "Confirmation of the Predicted Source of a Slow Folding Reaction - Proline-8 of Bovine Pancreatic Trypsin-Inhibitor." Protein Engineering **4**(4): 451-455.
- Hurle, M. R. y Matthews, C. R. (1987). "Proline Isomerization and the Slow Folding Reactions of the Alpha-Subunit of Tryptophan Synthase from Escherichia-Coli." Biochimica Et Biophysica Acta **913**(2): 179-184.
- Hurle, M. R., Michelotti, G. A., Crisanti, M. M. y Matthews, C. R. (1987). "Characterization of a Slow Folding Reaction for the Alpha-Subunit of Tryptophan Synthase." Proteins-Structure Function and Genetics **2**(1): 54-63.

- Ibarra-Molero, B. y Sanchez-Ruiz, J. M. (1996). "A model-independent, nonlinear extrapolation procedure for the characterization of protein folding energetics from solvent-denaturation data." Biochemistry **35**(47): 14689-14702.
- Ikura, M., Kay, L. E. y Bax, A. (1990). "A Novel-Approach for Sequential Assignment of H-1, C-13, and N-15 Spectra of Larger Proteins - Heteronuclear Triple-Resonance 3-Dimensional NMR-Spectroscopy - Application to Calmodulin." Biochemistry **29**(19): 4659-4667.
- Jacob, J., Krantz, B., Dothager, R. S., Thiyagarajan, P. y Sosnick, T. R. (2004). "Early collapse is not an obligate step in protein folding." Journal of Molecular Biology **338**(2): 369-382.
- Jarymowycz, V. A. y Stone, M. J. (2006). "Fast time scale dynamics of protein backbones: NMR relaxation methods, applications, and functional consequences." Chemical Review **106**(5): 1624-1671.
- Jee, J. y GÅ¼ntert, P. (2003). "Influence of the completeness of chemical shift assignments on NMR structures obtained with automated NOE assignment." Journal of Structural and Functional Genomics **4**(2): 179-189.
- Jelen, F., Oleksy, A., Smietana, K. y Otlewski, J. (2003). "PDZ domains - common players in the cell signaling." Acta Biochimica Polymers **50**(4): 985-1017.
- Johnson, B. A. y Blevins, R. A. (1994). "NMR View - a Computer-Program for the Visualization and Analysis of NMR Data." Journal of Biomolecular NMR **4**(5): 603-614.
- Johnson, W. C. J. (1990). Proteins: Structure, Function, and Genetics **7**: 205-214.
- Kamen, D. E. y Woody, R. W. (2002a). "Folding kinetics of the protein pectate lyase C reveal fast-forming intermediates and slow proline isomerization." Biochemistry **41**(14): 4713-4723.
- Kamen, D. E. y Woody, R. W. (2002b). "Identification of proline residues responsible for the slow folding kinetics in pectate lyase C by mutagenesis." Biochemistry **41**(14): 4724-4732.
- Kami, K., Takeya, R., Sumimoto, H. y Kohda, D. (2002). "Diverse recognition of non-PxxP peptide ligands by the SH3 domains from p67(phox), Grb2 and Pex13p." EMBO Journal **21**(16): 4268-4276.
- Kardinal, C., Konkol, B., Schulz, A., Posern, G., Lin, H., Adermann, K., Eulitz, M., Estrov, Z., Talpaz, M., Arlinghaus, R. B. y Feller, S. M. (2000). "Cell-penetrating SH3 domain blocker peptides inhibit proliferation of primary blast cells from CML patients." FASEB Journal **14**(11): 1529-1538.
- Kardinal, C., Posern, G., Zheng, J., Knudsen, B. S., Moarefi, I. y Feller, S. M. (1999). "Rational development of cell-penetrating high affinity SH3 domain binding peptides that selectively disrupt the signal transduction of Crk family adapters. Amgen Peptide Technology Group." Annals of the New York Academy of Sciences **886**: 289-292.
- Kay, B. K., Williamson, M. P. y Sudol, P. (2000). "The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains." FASEB Journal **14**(2): 231-241.
- Kay, L. E., Xu, G. Y., Singer, A. U., Muhandiram, D. R. y Formankay, J. D. (1993). "A Gradient-Enhanced HCCH-TOCSY Experiment for Recording Side-Chain 1H and 13C Correlations in H2O Samples of Proteins." Journal of Magnetic Resonance, Series B **101**(3): 333-337.
- Kay, L. E., Xu, G. Y. y Yamazaki, T. (1994). "Enhanced-Sensitivity Triple-Resonance Spectroscopy with Minimal H2O Saturation." Journal of Magnetic Resonance, Series A **109**(1): 129-133.
- Kelly, M. A., Chellgren, B. W., Rucker, A. L., Troutman, J. M., Fried, M. G., Miller, A. F. y Creamer, T. P. (2001). "Host-guest study of left-handed polyproline II helix formation." Biochemistry **40**(48): 14376-14383.

- Kiefhaber, T., Kohler, H. H. y Schmid, F. X. (1992). "Kinetic Coupling between Protein Folding and Prolyl Isomerization .1. Theoretical-Models." Journal of Molecular Biology **224**(1): 217-229.
- Kiefhaber, T., Quaas, R., Hahn, U. y Schmid, F. X. (1990). "Folding of Ribonuclease-T1 .1. Existence of Multiple Unfolded States Created by Proline Isomerization." Biochemistry **29**(12): 3053-3061.
- Kiefhaber, T. y Schmid, F. X. (1992). "Kinetic Coupling between Protein Folding and Prolyl Isomerization .2. Folding of Ribonuclease-a and Ribonuclease-T1." Journal of Molecular Biology **224**(1): 231-240.
- Kiyokawa, E., Mochizuki, N., Kurata, T. y Matsuda, M. (1997). "Role of Crk oncogene product in physiologic signaling." Critical Review in Oncology **8**(4): 329-342.
- Klejman, A., Schreiner, S. J., Nieborowska-Skorska, M., Slupianek, A., Wilson, M., Smithgall, T. E. y Skorski, T. (2002). "The Src family kinase Hck couples BCR/ABL to STAT5 activation in myeloid leukemia cells." EMBO Journal **21**(21): 5766-5774.
- Koch, C. A., Anderson, D., Moran, M. F., Ellis, C. y Pawson, T. (1991). "SH2 and SH3 domains: elements that control interactions of cytoplasmic signaling proteins." Science **252**(5006): 668-674.
- Konev, S. V. (1967). Fluorescence and Phosphorescence of Proteins and Nucleic Acids. New York (USA), Plenum Press.
- Koradi, R., Billeter, M. y Wuthrich, K. (1996). "MOLMOL: a program for display and analysis of macromolecular structures." Journal of Molecular Graphics and Modelling **14**(1): 51-55, 29-32.
- Korzhev, D. M., Billeter, M., Arseniev, A. S. y Orekhov, V. Y. (2001). "NMR studies of Brownian tumbling and internal motions in proteins." Progress in Nuclear Magnetic Resonance Spectroscopy **38**(3): 197-266.
- Krantz, B. A., Mayne, L., Rumbley, J., Englander, S. W. y Sosnick, T. R. (2002). "Fast and slow intermediate accumulation and the initial barrier mechanism in protein folding." Journal of Molecular Biology **324**(2): 359-371.
- Krantz, B. A. y Sosnick, T. R. (2000). "Distinguishing between two-state and three-state models for ubiquitin folding." Biochemistry **39**(38): 11696-11701.
- Krieger, F., Fierz, B., Bieri, O., Drewello, M. y Kiefhaber, T. (2003). "Dynamics of unfolded polypeptide chains as model for the earliest steps in protein folding." Journal of Molecular Biology **332**(1): 265-274.
- Krishnamurthy, V. M., Bohall, B. R., Semetey, V. y Whitesides, G. M. (2006). "The paradoxical thermodynamic basis for the interaction of ethylene glycol, glycine, and sarcosine chains with bovine carbonic anhydrase II: An unexpected manifestation of enthalpy/entropy compensation." Journal of the American Chemical Society **128**(17): 5802-5812.
- Kumar, A., Ernst, R. R. y Wuthrich, K. (1980). "A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules." Biochemical and Biophysical Research Communications **95**(1): 1-6.
- Kuriyan, J. y Cowburn, D. (1997). "Modular peptide recognition domains in eukaryotic signaling." Annual Review in Biophysical and Biochemical Communication **26**: 259-288.
- Kuwajima, K., Garvey, E. P., Finn, B. E., Matthews, C. R. y Sugai, S. (1991). "Transient Intermediates in the Folding of Dihydrofolate-Reductase as Detected by Far-Ultraviolet Circular-Dichroism Spectroscopy." Biochemistry **30**(31): 7693-7703.
- Laederach, A., Cradic, K. W., Fulton, D. B. y Andreotti, A. H. (2003). "Determinants of intra versus intermolecular self-association within the regulatory domains of RIK and Itk." Journal of Molecular Biology **329**(5): 1011-1020.

- Laemmli, U. K. (1970). "Cleavage of structural proteins during the assembly of the head of bacteriophage T4." Nature **227**(5259): 680-685.
- Laidler, K. J. y King, M. C. (1983). "The Development of Transition-State Theory." Journal of Physical Chemistry **87**(15): 2657-2664.
- Lapidus, L. J., Yao, S. H., McGarrity, K. S., Hertzog, D. E., Tubman, E. y Bakajin, O. (2007). "Protein hydrophobic collapse and early folding steps observed in a microfluidic mixer." Biophysical Journal **93**(1): 218-224.
- Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R. y Thornton, J. M. (1996). "AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR." Journal of Biomolecular NMR **8**(4): 477-486.
- Lawrence, D. S. (2005). "Signaling protein inhibitors via the combinatorial modification of peptide scaffolds." Biochimica Et Biophysica Acta **1754**(1-2): 50-57.
- Lee, K. Y., Hyeok Yoon, J. H., Kim, M., Roh, S., Lee, Y. S., Seong, B. L. y Kim, K. (2002). "A dipalmitoyl peptide that binds SH3 domain, disturbs intracellular signal transduction, and inhibits tumor growth in vivo." Biochemical and Biophysical Research Communications **296**(2): 434-442.
- Lehto, V. P., Wasenius, V. M., Salven, P. y Saraste, M. (1988). "Transforming and membrane proteins." Nature **334**(6181): 388.
- Lewitzky, M., Harkiolaki, M., Domart, M. C., Jones, E. Y. y Feller, S. M. (2004). "Mona/Gads SH3C binding to hematopoietic progenitor kinase 1 (HPK1) combines an atypical SH3 binding motif, R/KXXK, with a classical PXXP motif embedded in a polyproline type II (PPII) helix." Journal of Biological Chemistry **279**(27): 28724-28732.
- Lewitzky, M., Kardinal, C., Gehring, N. H., Schmidt, E. K., Konkol, B., Eulitz, M., Birchmeier, W., Schaeper, U. y Feller, S. M. (2001). "The C-terminal SH3 domain of the adapter protein Grb2 binds with high affinity to sequences in Gab1 and SLP-76 which lack the SH3-typical P-x-x-P core motif." Oncogene **20**(9): 1052-1062.
- Li, J. S., Shinjo, M., Matsumura, Y., Morita, M., Baker, D., Ikeguchi, M. y Kihara, H. (2007). "An alpha-helical burst in the src SH3 folding pathway." Biochemistry **46**(17): 5072-5082.
- Lim, W. A., Richards, F. M. y Fox, R. O. (1994). "Structural determinants of peptide-binding orientation and of sequence specificity in SH3 domains." Nature **372**(6504): 375-379.
- Lipari, G. y Szabo, A. (1982a). "Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity." Journal of the American Chemical Society **104**(17): 4546-4559.
- Lipari, G. y Szabo, A. (1982b). "Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results." Journal of the American Chemical Society **104**(17): 4559-4570.
- Liu, J., Albers, M. W., Chen, C. M., Schreiber, S. L. y Walsh, C. T. (1990). "Cloning, expression, and purification of human cyclophilin in Escherichia coli and assessment of the catalytic role of cysteines by site-directed mutagenesis." Proc. Natl. Acad. Sci. U S A **87**(6): 2304-2308.
- Lumry, R. y Rajender, S. (1970). "Enthalpy-Entropy Compensation Phenomena in Water Solutions of Proteins and Small Molecules - a Ubiquitous Property of Water." Biopolymers **9**(10): 1125-1128.
- Luque, I. y Freire, E. (1998). Structure-based prediction of binding affinities and molecular design of peptide ligands. Energetics of Biological Macromolecules, Pt B. **295**: 100-127.
- Luque, I. y Freire, E. (2000). "Structural stability of binding sites: consequences for binding affinity and allosteric effects." Proteins Suppl **4**: 63-71.

- Luque, I. y Freire, E. (2002). "Structural parameterization of the binding enthalpy of small ligands." Proteins-Structure Function and Genetics **49**(2): 181-190.
- Luque, I., Leavitt, S. A. y Freire, E. (2002). "The linkage between protein folding and functional cooperativity: Two sides of the same coin?" Annual Review of Biophysics and Biomolecular Structure **31**: 235-256.
- Luque, I., Todd, M. J., Gomez, J., Semo, N. y Freire, E. (1998). "Molecular basis of resistance to HIV-1 protease inhibition: A plausible hypothesis." Biochemistry **37**(17): 5791-5797.
- Mabrey, S., Mateo, P. L. y Sturtevant, J. M. (1978). "High-Sensitivity Scanning Calorimetric Study of Mixtures of Cholesterol with Dimyristoylphosphatidylcholines and Dipalmitoylphosphatidylcholines." Biochemistry **17**(12): 2464-2468.
- Macarthur, M. W. y Thornton, J. M. (1991). "Influence of Proline Residues on Protein Conformation." Journal of Molecular Biology **218**(2): 397-412.
- Magg, C., Kubelka, J., Holtermann, G., Haas, E. y Schmid, F. X. (2006). "Specificity of the initial collapse in the folding of the cold shock protein." Journal of Molecular Biology **360**(5): 1067-1080.
- Makhatadze, G. I. y Privalov, P. L. (1990). "Heat-Capacity of Proteins .1. Partial Molar Heat-Capacity of Individual Amino-Acid-Residues in Aqueous-Solution - Hydration Effect." Journal of Molecular Biology **213**(2): 375-384.
- Makhatadze, G. I. y Privalov, P. L. (1995). Energetics of protein structure. Advances in Protein Chemistry, Vol 47. **47**: 307-425.
- Maki, K., Cheng, H., Dolgikh, D. A. y Roder, H. (2007). "Folding kinetics of staphylococcal nuclease studied by tryptophan engineering and rapid mixing methods." Journal of Molecular Biology **368**(1): 244-255.
- Manavalan, P. y Johnson, W. C. (1983). "Identification of Tertiary Structural Types All-Alpha, All-Beta, Alpha+Beta and Alpha/Beta from Circular-Dichroism of Proteins." Biophysical Journal **41**(2): A267-A267.
- Mann, C. J. y Matthews, C. R. (1993). "Structure and Stability of an Early Folding Intermediate of Escherichia-Coli Trp Aporepressor Measured by Far-Uv Stopped-Flow Circular-Dichroism and 8-Anilino-1-Naphthalene Sulfonate Binding." Biochemistry **32**(20): 5282-5290.
- Manning, M. C. y Woody, R. W. (1991). "Theoretical Cd Studies of Polypeptide Helices - Examination of Important Electronic and Geometric Factors." Biopolymers **31**(5): 569-586.
- Martinez, J. C., Elharrous, M., Filimonov, V. V., Mateo, P. L. y Fersht, A. R. (1994). "A Calorimetric Study of the Thermal-Stability of Barnase and Its Interaction with 3'gmp." Biochemistry **33**(13): 3919-3926.
- Martinez, J. C., Filimonov, V. V., Mateo, P. L., Schreiber, G. y Fersht, A. R. (1995). "A Calorimetric Study of the Thermal-Stability of Barstar and Its Interaction with Barnaset." Biochemistry **34**(15): 5224-5233.
- Martinez, J. C., Pisabarro, M. T. y Serrano, L. (1998). "Obligatory steps in protein folding and the conformational diversity of the transition state." Nature Structural Biology **5**(8): 721-729.
- Martinez, J. C. y Serrano, L. (1999). "The folding transition state between SH3 domains is conformationally restricted and evolutionarily conserved." Nature Structural Biology **6**(11): 1010-1016.
- Martinez, J. C., Viguera, A. R., Berisio, R., Wilmanns, M., Mateo, P. L., Filimonov, V. V. y 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**(2): 549-559.

- Martin-Sierra, F. M., Candel, A. M., Casares, S., Filimonov, V. V., Martinez, J. C. y Conejero-Lara, F. (2003). "A binding event converted into a folding event." FEBS Letters **553**(3): 328-332.
- Mateu, M. G. N., J.L.; (2003). Plegamiento de proteínas y enfermedades relacionadas. Estructura de proteínas. Barcelona, Editorial Ariel, S.A.: 201-231.
- Matouschek, A., Serrano, L. y Fersht, A. R. (1992). "The Folding of an Enzyme .4. Structure of an Intermediate in the Refolding of Barnase Analyzed by a Protein Engineering Procedure." Journal of Molecular Biology **224**(3): 819-835.
- Matthews, C. R. y Hurle, M. R. (1987). "Mutant Sequences as Probes of Protein Folding Mechanisms." Bioessays **6**(6): 254-257.
- Maxwell, K. L., Wildes, D., Zarrine-Afsar, A., de los Rios, M. A., Brown, A. G., Friel, C. T., Hedberg, L., Horng, J. C., Bona, D., Miller, E. J., Vallee-Belisle, A., Main, E. R. G., Bemporad, F., Qiu, L. L., Teilum, K., Vu, N. D., Edwards, A. M., Ruczinski, I., Poulsen, F. M., Kragelund, B. B., Michnick, S. W., Chiti, F., Bai, Y. W., Hagen, S. J., Serrano, L., Oliveberg, M., Raleigh, D. P., Wittung-Stafshede, P., Radford, S. E., Jackson, S. E., Sosnick, T. R., Marqusee, S., Davidson, A. R. y Plaxco, K. W. (2005). "Protein folding: Defining a "standard" set of experimental conditions and a preliminary kinetic data set of two-state proteins." Protein Science **14**(3): 602-616.
- Mayer, B. J. (2001). "SH3 domains: complexity in moderation." Journal of Cell Science **114**(7): 1253-1263.
- Mayne, L., Paterson, Y., Cerasoli, D. y Englander, S. W. (1992). "Effect of Antibody-Binding on Protein Motions Studied by Hydrogen-Exchange Labeling and 2-Dimensional NMR." Biochemistry **31**(44): 10678-10685.
- McPherson, P. S. (1999). "Regulatory role of SH3 domain-mediated protein-protein interactions in synaptic vesicle endocytosis." Cell Signal **11**(4): 229-238.
- Mello, C. C., Bradley, C. M., Tripp, K. W. y Barrick, D. (2005). "Experimental characterization of the folding kinetics of the notch ankyrin domain." Journal of Molecular Biology **352**(2): 266-281.
- Miyazaki, K., Senga, T., Matsuda, S., Tanaka, M., Machida, E., Takenouchi, Y., Nimura, Y. y Hamaguchi, M. (1999). "Critical amino acid substitutions in the Src SH3 domain that convert c-Src to be oncogenic." Biochemical and Biophysical Research Communications **263**(3): 759-764.
- Mullis, K. B. y Faloona, F. A. (1987). "Specific Synthesis of DNA Invitro Via a Polymerase-Catalyzed Chain-Reaction." Methods in Enzymology **155**: 335-350.
- Musacchio, A. (2002). "How SH3 domains recognize proline." Advances in Protein Chemistry **61**: 211-268.
- Musacchio, A. (2003). How SH3 domains recognize proline. Protein Modules and Protein-Protein Interactions. **61**: 211-268.
- Musacchio, A., Gibson, T., Lehto, V. P. y Saraste, M. (1992a). "SH3 - an Abundant Protein Domain in Search of a Function." FEBS Letters **307**(1): 55-61.
- Musacchio, A., Noble, M., Pauptit, R., Wierenga, R. y Saraste, M. (1992b). "Crystal structure of a Src-homology 3 (SH3) domain." Nature **359**(6398): 851-855.
- Musacchio, A., Saraste, M. y Wilmanns, M. (1994). "High-Resolution Crystal-Structures of Tyrosine Kinase SH3 Domains Complexed with Proline-Rich Peptides." Nature Structural Biology **1**(8): 546-551.
- Myers, J. K., Pace, C. N. y Scholtz, J. M. (1995). "Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding." Protein Science **4**(10): 2138-2148.
- Nall, B. T., Garel, J. R. y Baldwin, R. L. (1978). "Test of the extended two-state model for the kinetic intermediates observed in the folding transition of ribonuclease A." Journal of Molecular Biology **118**(3): 317-330.

- Nguyen, J. T. y Lim, W. A. (1997). "How Src exercises self-restraint." Nature Structural Biology **4**(4): 256-260.
- Nguyen, J. T., Porter, M., Amoui, M., Miller, W. T., Zuckermann, R. N. y Lim, W. A. (2000). "Improving SH3 domain ligand selectivity using a non-natural scaffold." Chemistry & Biology **7**(7): 463-473.
- Nguyen, J. T., Turck, C. W., Cohen, F. E., Zuckermann, R. N. y Lim, W. A. (1998). "Exploiting the Basis of Proline Recognition by SH3 and WW Domains: Design of N-Substituted Inhibitors 10.1126/science.282.5396.2088." Science **282**(5396): 2088-2092.
- Oneyama, C., Nakano, H. y Sharma, S. V. (2002). "UCS15A, a novel small molecule, SH3 domain-mediated protein-protein interaction blocking drug." Oncogene **21**(13): 2037-2050.
- Ortega, A. y Garcia de la Torre, J. (2005). "Efficient, accurate calculation of rotational diffusion and NMR relaxation of globular proteins from atomic-level structures and approximate hydrodynamic calculations." Journal of the American Chemical Society **127**(37): 12764-12765.
- Otzen, D. E., Itzhaki, L. S., elMasry, N. F., Jackson, S. E. y Fersht, A. R. (1994). "Structure of the transition state for the folding/unfolding of the barley chymotrypsin inhibitor 2 and its implications for mechanisms of protein folding." Proc. Natl. Acad. Sci. U S A **91**(22): 10422-10425.
- Pace, C. N., Shirley, B. A. y Thomson, J. A. (1989). Measuring the conformational stability of a protein. Protein Structure, a practical approach. Creighton, T. E., IRL Press at Oxford University Press, Oxford: 311-330.
- Palencia, A., Cobos, E. S., Mateo, P. L., Martinez, J. C. y Luque, I. (2004). "Thermodynamic dissection of the binding energetics of proline-rich peptides to the Abl-SH3 domain: Implications for rational ligand design." Journal of Molecular Biology **336**(2): 527-537.
- Palmer, A. G., Cavanagh, J., Byrd, R. A. y Rance, M. (1992). "Sensitivity Improvement in 3-Dimensional Heteronuclear Correlation NMR-Spectroscopy." Journal of Magnetic Resonance **96**(2): 416-424.
- Panni, S., Dente, L. y Cesareni, G. (2002). "In vitro evolution of recognition specificity mediated by SH3 domains reveals target recognition rules." Journal of Biological Chemistry **277**(24): 21666-21674.
- Patt, S. L. (1992). "Single-Frequency-Shifted and Multiple-Frequency-Shifted Lamina Pulses." Journal of Magnetic Resonance **96**(1): 94-102.
- Pawson, T. (1995). "Protein modules and signalling networks." Nature **373**(6515): 573-580.
- Pawson, T., Raina, M. y Nash, P. (2002). "Interaction domains: from simple binding events to complex cellular behavior." FEBS Letters **513**(1): 2-10.
- Pawson, T. y Schlessinger, J. (1993). "SH2 and SH3 Domains." Current Biology **3**(7): 434-442.
- Pawson, T. y Scott, J. D. (1997). "Signaling through scaffold, anchoring, and adaptor proteins." Science **278**(5346): 2075-2080.
- Penzer, G. R. (1980). An introduction to spectroscopy for biochemistry. Brown, S. B. e. London, Academic Press. **70**.
- Peranen, J., Rikkinen, M., Hyvonen, M. y Kaariainen, L. (1996). "T7 vectors with modified T7lac promoter for expression of proteins in Escherichia coli." Analytical Biochemistry **236**(2): 371-373.
- Pérez-Cañadillas, J. M. y Bruix, M., ; (2003). Dinámica de proteínas por RMN. Estructura de Proteínas. Barcelona, España, Editorial Ariel, S.A.: 333-355.
- Perl, D., Welker, C., Schindler, T., Schroder, K., Marahiel, M. A., Jaenicke, R. y Schmid, F. X. (1998). "Conservation of rapid two-state folding in mesophilic,

- thermophilic and hyperthermophilic cold shock proteins." Nature Structural Biology **5**(3): 229-235.
- Permyakov, E. A. (1993). Luminiscent Spectroscopy of Proteins. Boca Raton (Florida, USA), CRC Press.
- Perry, K. M., Onuffer, J. J., Gittelman, M. S., Barmat, L. y Matthews, C. R. (1989). "Long-Range Electrostatic Interactions Can Influence the Folding, Stability, and Cooperativity of Dihydrofolate-Reductase." Biochemistry **28**(19): 7961-7968.
- Piotto, M., Saudek, V. y Sklenar, V. (1992). "Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions." Journal of Biomolecular NMR **2**(6): 661-665.
- Pisabarro, M. T., Ortiz, A. R., Viguera, A. R., Gago, F. y Serrano, L. (1994). "Molecular modeling of the interaction of polyproline-based peptides with the Abl-SH3 domain: rational modification of the interaction." Protein Engineering **7**(12): 1455-1462.
- Pisabarro, M. T. y Serrano, L. (1996). "Rational design of specific high-affinity peptide ligands for the Abl-SH3 domain." Biochemistry **35**(33): 10634-10640.
- Pisabarro, M. T., Serrano, L. y Wilmanns, M. (1998). "Crystal structure of the Abl-SH3 domain complexed with a designed high-affinity peptide ligand: Implications for SH3-ligand interactions." Journal of Molecular Biology **281**(3): 513-521.
- Plaxco, K. W., Guijarro, J. I., Morton, C. J., Pitkeathly, M., Campbell, I. D. y Dobson, C. M. (1998a). "The folding kinetics and thermodynamics of the Fyn-SH3 domain." Biochemistry **37**(8): 2529-2537.
- Plaxco, K. W., Simons, K. T. y Baker, D. (1998b). "Contact order, transition state placement and the refolding rates of single domain proteins." Journal of Molecular Biology **277**(4): 985-994.
- Plotnikov, V. V., Brandts, J. M., Lin, L. N. y Brandts, J. F. (1997). "A new ultrasensitive scanning calorimeter." Analytical Biochemistry **250**(2): 237-244.
- Polshakov, V. I., Birdsall, B. y Feeney, J. (2006). "Effects of co-operative ligand binding on protein amide NH hydrogen exchange." Journal of Molecular Biology **356**(4): 886-903.
- Pople, J. A. S., W.G. & Bernstein, H.J. (1959). High Resolution NMR. New York (USA), McGraw Hill.
- Prabhu, N. V. y Sharp, K. A. (2005). "Heat capacity in proteins." Annual Review of Physical Chemistry **56**(1): 521-548.
- Prajapati, R. S., Das, M., Sreeramulu, S., Sirajuddin, M., Srinivasan, S., Krishnamurthy, V., Ranjani, R., Ramakrishnan, C. y Varadarajan, R. (2007). "Thermodynamic effects of proline introduction on protein stability." Proteins-Structure Function and Bioinformatics **66**(2): 480-491.
- Privalov, P. L. (1979). "Stability of proteins: small globular proteins." Advance in Protein Chemistry **33**: 167-241.
- Privalov, P. L. y Khechinashvili, N. N. (1974). "A thermodynamic approach to the problem of stabilization of globular protein structure: a calorimetric study." Journal of Molecular Biology **86**(3): 665-684.
- Privalov, P. L. y Makhatadze, G. I. (1990). "Heat-Capacity of Proteins .2. Partial Molar Heat-Capacity of the Unfolded Polypeptide-Chain of Proteins - Protein Unfolding Effects." Journal of Molecular Biology **213**(2): 385-391.
- Privalov, P. L. y Potekhin, S. A. (1986). "Scanning microcalorimetry in studying temperature-induced changes in proteins." Methods in Enzymology **131**: 4-51.
- Privalov, P. L., Tiktopulo, E. I., Venyaminov, S. Y., Griko, Y. V., Makhatadze, G. I. y Khechinashvili, N. N. (1989). "Heat-Capacity and Conformation of Proteins in the Denatured State." Journal of Molecular Biology **205**(4): 737-750.

- Rabanal, F., Ludevid, M. D., Pons, M. y Giralt, E. (1993). "Cd of Proline-Rich Polypeptides - Application to the Study of the Repetitive Domain of Maize Glutelin-2." Biopolymers **33**(7): 1019-1028.
- Ren, R., Mayer, B. J., Cicchetti, P. y Baltimore, D. (1993). "Identification of a ten-amino acid proline-rich SH3 binding site." Science **259**(5098): 1157-1161.
- Renzoni, D. A., Pugh, D. J. R., Siligardi, G., Das, P., Morton, C. J., Rossi, C., Waterfield, M. D., Campbell, I. D. y 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**(49): 15646-15653.
- Robert, C. H., Gill, S. J. y Wyman, J. (1988). "Quantitative-Analysis of Linkage in Macromolecules When One Ligand Is Present in Limited Total Quantity." Biochemistry **27**(18): 6829-6835.
- Roberts, G. C. (1999). "NMR spectroscopy in structure-based drug design." Current Opinion in Biotechnology **10**(1): 42-47.
- Roder, H. y Colon, W. (1997). "Kinetic role of early intermediates in protein folding." Current Opinion in Structural Biology **7**(1): 15-28.
- Roder, H., Maki, K. y Cheng, H. (2006). "Early events in protein folding explored by rapid mixing methods." Chemical Review **106**(5): 1836-1861.
- Rucker, A. L., Pager, C. T., Campbell, M. N., Qualls, J. E. y Creamer, T. P. (2003). "Host-guest scale of left-handed polyproline II helix formation." Proteins-Structure Function and Genetics **53**(1): 68-75.
- Sadqi, M. (2000). Tesis Doctoral: Estudio Termodinámico de los estados parcialmente plegados del dominio SH3 de alfa espectrina. Departamento de Química Física, Universidad de Granada: 347.
- Sadqi, M., Casares, S., Abril, M. A., Lopez-Mayorga, O., Conejero-Lara, F. y Freire, E. (1999). "The native state conformational ensemble of the SH3 domain from alpha-spectrin." Biochemistry **38**(28): 8899-8906.
- Sadqi, M., Casares, S., Lopez-Mayorga, O. y Conejero-Lara, F. (2002a). "The temperature dependence of the hydrogen exchange in the SH3 domain of alpha-spectrin." FEBS Letters **527**(1-3): 86-90.
- Sadqi, M., Casares, S., Lopez-Mayorga, O., Martinez, J. C. y Conejero-Lara, F. (2002b). "pH dependence of the hydrogen exchange in the SH3 domain of alpha-spectrin." FEBS Letters **514**(2-3): 295-299.
- Sanchez, I. E. y Kiefhaber, T. (2003). "Origin of unusual Phi-values in protein folding: Evidence against specific nucleation sites." Journal of Molecular Biology **334**(5): 1077-1085.
- Sanchez-Ruiz, J. M. (1995). "Differential scanning calorimetry of proteins." Subcellular Biochemistry **24**: 133-176.
- Sauder, J. M., MacKenzie, N. E. y Roder, H. (1996). "Kinetic mechanism of folding and unfolding of Rhodobacter capsulatus cytochrome c(2)." Biochemistry **35**(51): 16852-16862.
- Schmid, F. X. (1989). Spectral methods of characterizing protein conformation and conformational changes. Protein Structure, a practical approach. Creighton, T. E., IRL Press at Oxford University Press, Oxford: 251-286.
- Schmid, F. X. (1992). Kinetics of unfolding and refolding of single-domain proteins. Protein folding. Creighton, T. E. New York, W.H. Freeman and Company: 197-241.
- Schmid, F. X. (2005). Prolyl Isomerization in Protein Folding. Protein Folding Handbook. Buchner, J. y Kiefhaber, T., WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim. **2**: 916-964.

- Schmid, F. X. y Baldwin, R. L. (1978). "Acid catalysis of the formation of the slow-folding species of RNase A: evidence that the reaction is proline isomerization." Proc. Natl. Acad. Sci. U S A **75**(10): 4764-4768.
- Schmid, F. X. y Baldwin, R. L. (1979). "Detection of an Early Intermediate in the Folding of Ribonuclease-a by Protection of Amide Protons against Exchange." Journal of Molecular Biology **135**(1): 199-215.
- Schmid, F. X., Mayr, L. M., Mucke, M. y Schonbrunner, E. R. (1993). Prolyl Isomerases - Role in Protein-Folding. Advances in Protein Chemistry **44**: 25-66.
- Schumacher, T. N., Mayr, L. M., Minor, D. L., Jr., Milhollen, M. A., Burgess, M. W. y Kim, P. S. (1996). "Identification of D-peptide ligands through mirror-image phage display." Science **271**(5257): 1854-1857.
- Serrano, L., Matouschek, A. y Fersht, A. R. (1992). "The Folding of an Enzyme .3. Structure of the Transition-State for Unfolding of Barnase Analyzed by a Protein Engineering Procedure." Journal of Molecular Biology **224**(3): 805-818.
- Shastry, M. C. R. y Roder, H. (1998). "Evidence for barrier-limited protein folding kinetics on the microsecond time scale." Nature Structural Biology **5**(5): 385-392.
- Shrake, A. y Ross, P. D. (1990). "Ligand-Induced Biphasic Protein Denaturation." Journal of Biological Chemistry **265**(9): 5055-5059.
- Sigurskjold, B. W. (2000). "Exact analysis of competition ligand binding by displacement isothermal titration calorimetry." Analytical Biochemistry **277**(2): 260-266.
- Sinha, K. K. y Udgaonkar, J. B. (2007). "Dissecting the non-specific and specific components of the initial folding reaction of barstar by multi-site FRET measurements." Journal of Molecular Biology **370**(2): 385-405.
- 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. y Calabretta, B. (1998). "The SH3 domain contributes to BCR/ABL-dependent leukemogenesis in vivo: role in adhesion, invasion, and homing." Blood **91**(2): 406-418.
- Spolar, R. S. y Record, M. T., Jr. (1994). "Coupling of local folding to site-specific binding of proteins to DNA." Science **263**(5148): 777-784.
- Stark, G. R. (1965). "Reactions of Cyanate with Functional Groups of Proteins .4. Inertness of Aliphatic Hydroxyl Groups . Formation of Carbamyl- and Acylhydantoins." Biochemistry **4**(11): 2363-2366.
- Stauffer, T. P., Martenson, C. H., Rider, J. E., Kay, B. K. y Meyer, T. (1997). "Inhibition of Lyn function in mast cell activation by SH3 domain binding peptides." Biochemistry **36**(31): 9388-9394.
- Stein, R. (1998). "SH2 and SH3 domains. Unraveling signaling networks with peptide antagonists." Methods in Molecular Biology **88**: 187-195.
- Straume, M. y Freire, E. (1992). "2-Dimensional Differential Scanning Calorimetry - Simultaneous Resolution of Intrinsic Protein Structural Energetics and Ligand-Binding Interactions by Global Linkage Analysis." Analytical Biochemistry **203**(2): 259-268.
- Sudol, M. (1998). "From Src Homology domains to other signaling modules: proposal of the 'protein recognition code'." Oncogene **17**(11 Reviews): 1469-1474.
- Takei, J., Chu, R. A. y Bai, Y. (2000). "Absence of stable intermediates on the folding pathway of barnase." Proc. Natl. Acad. Sci. U S A **97**(20): 10796-10801.
- Thein, S. L. W., R.B.; (1986). Human Genetic Diseases, a practical approach, IRL Press, Virginia.
- 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. y Cesareni, G. (2002). "A combined

- experimental and computational strategy to define protein interaction networks for peptide recognition modules." *Science* **295**(5553): 321-324.
- Vega, M. C., Martínez, J. C. y Serrano, L. (2000). "Thermodynamic and structural characterization of Asn and Ala residues in the disallowed II' region of the Ramachandran plot." *Protein Science* **9**(12): 2322-2328.
- Velazquez Campoy, A. y Freire, E. (2005). "ITC in the post-genomic era...? Priceless." *Biophysical Chemistry* **115**(2-3): 115-124.
- Velazquez-Campoy, A. y Freire, E. (2001). "Incorporating target heterogeneity in drug design." *Journal of Cellular Biochemistry Supplement* **Suppl 37**: 82-88.
- Velazquez-Campoy, A., Luque, I. y Freire, E. (2001). "The application of thermodynamic methods in drug design." *Thermochimica Acta* **380**(2): 217-227.
- Velazquez-Campoy, A., Luque, I., Todd, M. J., Milutinovich, M., Kiso, Y. y Freire, E. (2000). "Thermodynamic dissection of the binding energetics of KNI-272, a potent HIV-1 protease inhibitor." *Protein Science* **9**(9): 1801-1809.
- Vidal, M., Gigoux, V. y Garbay, C. (2001). "SH2 and SH3 domains as targets for anti-proliferative agents." *Critical Reviews in Oncology Hematology* **40**(2): 175-186.
- Viguera, A. R., Arrondo, J. L. R., Musacchio, A., Saraste, M. y Serrano, L. (1994a). "Characterization of the Interaction of Natural Proline-Rich Peptides with 5 Different SH3 Domains." *Biochemistry* **33**(36): 10925-10933.
- Viguera, A. R., Blanco, F. J. y Serrano, L. (1995). "The Order of Secondary Structure Elements Does Not Determine the Structure of a Protein but Does Affect Its Folding Kinetics." *Journal of Molecular Biology* **247**(4): 670-681.
- Viguera, A. R., Martínez, J. C., Filimonov, V. V., Mateo, P. L. y Serrano, L. (1994b). "Thermodynamic and Kinetic-Analysis of the SH3 Domain of Spectrin Shows a 2-State Folding Transition." *Biochemistry* **33**(8): 2142-2150.
- Viguera, A. R. y Serrano, L. (1997). "Loop length, intramolecular diffusion and protein folding." *Nature Structural Biology* **4**(11): 939-946.
- Viguera, A. R., Serrano, L. y Wilmanns, M. (1996). "Different folding transition states may result in the same native structure." *Nature Structural Biology* **3**(10): 874-880.
- Wallace, A. C., Laskowski, R. A. y Thornton, J. M. (1995). "Ligplot - a Program to Generate Schematic Diagrams of Protein Ligand Interactions." *Protein Engineering* **8**(2): 127-134.
- Wang, C. Y., Pawley, N. H. y Nicholson, L. K. (2001a). "The role of backbone motions in ligand binding to the c-Src SH3 domain." *Journal of Molecular Biology* **313**(4): 873-887.
- Wang, W., Lim, W. A., Jakalian, A., Wang, J., Luo, R., Bayly, C. I. y Kollman, P. A. (2001b). "An analysis of the interactions between the Sem-5 SH3 domain and its ligands using molecular dynamics, free energy calculations, and sequence analysis." *Journal of the American Chemical Society* **123**(17): 3986-3994.
- Warren, J. R. y Gordon, J. A. (1966). "On the Refractive Indices of Aqueous Solutions of Urea." *Journal of Physical Chemistry* **70**(1): 297-300.
- Weinryb, I. S., R.F. (1971). The Luminescence of aromatic amino acids. *Excited States of Proteins and Nucleic Acids*. Weinryb, R. F. S. a. I. New York (USA), Plenum Press: 277-318.
- Welker, E., Maki, K., Shastry, M. C., Juminaga, D., Bhat, R., Scheraga, H. A. y Roder, H. (2004). "Ultrarapid mixing experiments shed new light on the characteristics of the initial conformational ensemble during the folding of ribonuclease A." *Proc. Natl. Acad. Sci. U S A* **101**(51): 17681-17686.
- Wildes, D. y Marqusee, S. (2005). "Hydrogen exchange and ligand binding: Ligand-dependent and ligand-independent protection in the Src SH3 domain." *Protein Science* **14**(1): 81-88.

- Williams, D. H., Stephens, E., O'Brien, D. P. y Zhou, M. (2004). "Understanding noncovalent interactions: ligand binding energy and catalytic efficiency from ligand-induced reductions in motion within receptors and enzymes." Angew. Chem. Int. Ed. Engl. **43**(48): 6596-616.
- Wiseman, T., Williston, S., Brandts, J. F. y Lin, L. N. (1989). "Rapid Measurement of Binding Constants and Heats of Binding Using a New Titration Calorimeter." Analytical Biochemistry **179**(1): 131-137.
- Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S. y Sykes, B. D. (1995). "H-1, C-13 and N-15 Random Coil NMR Chemical-Shifts of the Common Amino-Acids .1. Investigations of Nearest-Neighbor Effects (Vol 5, Pg 67, 1995)." Journal of Biomolecular NMR **5**(3): 332-332.
- Wishart, D. S. y Sykes, B. D. (1994). "The C-13 Chemical-Shift Index - a Simple Method for the Identification of Protein Secondary Structure Using C-13 Chemical-Shift Data." Journal of Biomolecular NMR **4**(2): 171-180.
- Witter, D. J., Famiglietti, S. J., Cambier, J. C. y Castelhana, A. L. (1998). "Design and synthesis of SH3 domain binding ligands: modifications of the consensus sequence XPpXP." Bioorganic & Medicinal Chemistry Letters **8**(22): 3137-3142.
- Woody, R. W. (1994). "Contributions of tryptophan side chains to the far-ultraviolet circular dichroism of proteins." European Biophysical Journal **23**(4): 253-262.
- Wu, J., Yang, J. T. y Wu, C. S. C. (1992). "Beta-li Conformation of All-Beta Proteins Can Be Distinguished from Unordered Form by Circular-Dichroism." Analytical Biochemistry **200**(2): 359-364.
- Wu, Y., Vadrevu, R., Yang, X. Y. y Matthews, C. R. (2005). "Specific structure appears at the N terminus in the sub-millisecond folding intermediate of the alpha subunit of tryptophan synthase, a TIM barrel protein." Journal of Molecular Biology **351**(3): 445-452.
- Wüthrich, K. (1986). NMR of Proteins and Nucleic Acids. New York.
- Wuthrich, K. y Grathwoh, C. (1974). "Novel Approach for Studies of Molecular-Conformations in Flexible Polypeptides." FEBS Letters **43**(3): 337-340.
- Xu, W., Harrison, S. C. y Eck, M. J. (1997). "Three-dimensional structure of the tyrosine kinase c-Src." Nature **385**(6617): 595-602.
- Yan, K. S., Kuti, M. y Zhou, M. M. (2002). "PTB or not PTB -- that is the question." FEBS Letters **513**(1): 67-70.
- Yu, H., Chen, J. K., Feng, S., Dalgarno, D. C., Brauer, A. W. y Schreiber, S. L. (1994). "Structural basis for the binding of proline-rich peptides to SH3 domains." Cell **76**(5): 933-945.
- Yu, H., Rosen, M. K., Shin, T. B., Seidel-Dugan, C., Brugge, J. S. y Schreiber, S. L. (1992). "Solution structure of the SH3 domain of Src and identification of its ligand-binding site." Science **258**(5088): 1665-1668.
- Yuzawa, S., Yokochi, M., Hatanaka, H., Ogura, K., Kataoka, M., Miura, K., Mandiyan, V., Schlessinger, J. y Inagaki, F. (2001). "Solution structure of Grb2 reveals extensive flexibility necessary for target recognition." Journal of Molecular Biology **306**(3): 527-537.
- Zaman, M. H., Berry, R. S. y Sosnick, T. R. (2002). "Entropic benefit of a cross-link in protein association." Proteins-Structure Function and Genetics **48**(2): 341-351.
- Zarrine-Afsar, A. y Davidson, A. R. (2004). "The analysis of protein folding kinetic data produced in protein engineering experiments." Methods **34**(1): 41-50.
- Zarrinpar, A., Bhattacharyya, R. P. y Lim, W. A. (2003). "The structure and function of proline recognition domains." Science Signaling **2003**(179): RE8.
- Zhou, H. X. (2006). "Quantitative relation between intermolecular and intramolecular binding of pro-rich peptides to SH3 domains." Biophysical Journal **91**(9): 3170-3181.

APÉNDICES

APÉNDICE I: TABLA DE DESPLAZAMIENTOS QUÍMICOS

Tabla A.1. Desplazamientos químicos en ppm de ^1H , ^{13}C (cursiva) y ^{15}N (cursiva) de SPCp41 en glicina-D5 20 mM, 93% H_2O :7% D_2O a pH 3.5 y 25 °C. ^aEste valor no pudo observarse.

Residuo	NH	CO	C_αH	C_βH	Otros
Gly1			3.849, 3.849		
Ala2	8.587 (123.863)		4.359	1.387	
Met3	8.460 (119.757)	(172.802)	4.534 (55.423)	2.015, 2.101 (32.969)	C_γH 2.568, 2.621 (31.107)
Gly4	8.472 (112.111)	(--- ^a)	4.034, 4.162 (44.822)		
Pro5		(177.247)	4.458 (63.418)	1.926, 2.282 (32.457)	C_γH 1.990, 1.990 (27.389) C_δH 3.612, 3.612 (50.152)
Arg6	8.495 (120.801)	(176.122)	4.383 (56.200)	1.795, 1.913 (30.942)	C_γH 1.654, 1.715 (27.187) C_δH 3.220, 3.220 (43.673) $\text{N}_\epsilon\text{H}$ 7.183 (115.328)
Glu7	8.078 (120.603)	(176.109)	4.593 (55.792)	1.950, 2.084 (55.792)	C_γH 2.393, 2.393 (33.743)
Val8	8.296 (120.603)	(174.660)	4.096 (62.866)	1.920 (33.474)	$\text{C}_\gamma\text{H}_3$ 0.753, 0.828 (21.125, 21.861)
Thr9	7.972 (120.018)	(173.345)	4.803 (61.409)	4.063 (70.989)	$\text{C}_\gamma\text{H}_3$ 1.136 (24.631)
Met10	9.097 (122.967)	(173.829)	4.667 (55.224)	1.992, 2.240 (36.409)	C_γH 2.458, 2.655 (32.477)
Lys11	8.688 (126.398)	(175.497)	4.871 (53.481)	1.671, 1.704 (34.507)	C_γH 1.382, 1.406 (24.793) C_δH --- ^a (--- ^a)
Lys12	8.501 (121.424)	(177.341)	3.295 (58.620)	1.456, 1.610 (32.722)	C_γH 1.100, 1.166 (24.545) C_δH 1.654, 1.654 (29.750) $\text{C}_\epsilon\text{H}$ --- ^a (--- ^a)
Gly13	8.834 (115.551)	(174.350)	3.522, 4.482 (45.024)		
Asp14	8.489 (121.772)	(174.595)	4.564 (55.453)	2.615, 2.857 (41.349)	
Ile15	8.118 (120.153)	(176.683)	5.071 (60.407)	1.766 (38.623)	C_γH 1.103, 1.669 (28.160) $\text{C}_\gamma\text{H}_3$ 0.888 (19.021) $\text{C}_\delta\text{H}_3$ 0.847 (12.954)
Leu16	9.280 (126.959)	(175.911)	4.966 (54.216)	1.425, 1.645 (44.419)	C_γH 1.584 (28.602) $\text{C}_{\delta 1}\text{H}_3$ 0.865 (--- ^a) $\text{C}_{\delta 2}\text{H}_3$ 0.893 (27.891)

Thr17	8.477 (117.328)	(173.933)	4.569 (63.734)	4.099 (69.962)	C _γ H ₃ 1.154 (--- ^a)
Leu18	8.787 (129.268)	(174.874)	4.407 (54.703)	1.122, 1.770 (43.415)	C _γ H 1.146 (26.964) C _{δ1} H ₃ 0.391 (25.507) C _{δ2} H ₃ 0.682 (24.015)
Leu19	8.973 (126.111)	(177.750)	4.494 (55.406)	1.155, 1.411 (43.213)	C _γ H 1.407 (27.092) C _{δ1} H ₃ 0.746 (25.978) C _{δ2} H ₃ 0.704 (25.978)
Asn20	7.526 (114.046)	(174.573)	4.710 (54.701)	2.651, 2.795 (41.354)	N _δ H 7.184, 7.916 (114.763)
Ser21	9.063 (123.165)	(173.723)	3.950 (57.066)	2.032, 2.878 (62.415)	
Thr22	8.166 (115.378)	(175.474)	3.973 (65.615)	4.273 (69.709)	C _γ H ₃ 1.307 (--- ^a)
Asn23	8.622 (122.914)	(174.922)	4.832 (53.471)	2.853, 3.832 (40.374)	N _δ H 7.484, 7.930 (114.775)
Lys24	8.485 (121.275)	(176.547)	4.265 (58.469)	1.796, 1.841 (32.784)	C _γ H 1.442, 1.442 (24.519) C _δ H 1.695, 1.695 (29.544) C _ε H 3.037, 3.073 (42.422)
Asp25	8.326 (115.342)	(176.916)	4.524 (55.596)	1.955, 2.801 (43.386)	
Trp26	8.104 (121.794)	(174.490)	5.132 (55.890)	2.865, 2.972 (33.085)	
Trp27	9.530 (124.771)	(174.485)	5.511 (54.170)	2.804, 2.967 (31.848)	
Lys28	8.723 (124.270)	(175.913)	4.415 (55.430)	1.104, 1.498 (34.684)	C _γ H 0.305, 0.948 (25.911) C _δ H 1.349, 1.394 (29.562) C _ε H 2.614, 2.684 (--- ^a)
Val29	9.176 (121.453)	(173.995)	5.334 (59.048)	2.114 (36.757)	C _γ H ₃ 0.753, 0.835 (18.876, 21.878)
Glu30	8.666 (118.996)	(174.821)	5.332 (54.514)	1.861, 1.861 (32.215)	C _γ H 2.292, 2.292 (32.215)
Val31	8.657 (124.548)	(174.801)	4.453 (61.217)	2.115 (33.928)	C _γ H ₃ 0.929, 1.045 (19.733, 21.563)
Asn32	9.425 (126.251)	(174.402)	4.334 (55.247)	2.861, 3.063 (37.740)	N _δ H 7.042, 7.586 (113.211)
Asp33	8.748 (112.100)	(174.748)	4.418 (54.934)	2.960, 3.077 (38.682)	
Arg34	8.106 (119.743)	(175.030)	4.691 (55.131)	1.795, 1.917 (32.315)	C _γ H 1.654, 1.716 (27.393) C _δ H 3.218, 3.218 (43.617) N _ε H 7.199 (114.751)

Gln35	8.465 (119.853)	(175.915)	5.333 (54.287)	1.739, 1.890 (31.572)	C _γ H 2.109, 2.253 (33.845) N _ε H 6.713, 7.304 (110.590)
Gly36	8.434 (107.483)	(170.828)	4.129, 4.129 (45.554)		
Phe37	8.492 (117.750)	(175.924)	5.592 (57.997)	2.761, 2.761 (42.204)	C _δ H 7.024, 7.024 (--- ^a) C _ε H 7.387, 7.387 (--- ^a) C _ζ H --- ^a (--- ^a)
Val38	9.262 (113.319)	(--- ^a)	4.919 (57.311)	1.825 (33.096)	C _γ H ₃ 0.711, 1.048 (17.975, 21.538)
Pro39		(177.902)	3.846 (61.946)	1.045, 1.327 (31.143)	C _γ H 0.104, 0.715 (27.733) C _δ H 2.097, 2.391 (50.297)
Ala40	7.651 (130.254)	(178.861)	2.629 (55.137)	-0.068 (16.049)	
Ala41	8.148 (113.328)	(177.864)	4.113 (52.897)	1.323 (18.002)	
Tyr42	7.920 (116.511)	(174.736)	4.793 (57.157)	3.163, 3.336 (37.725)	C _δ H 6.830, 6.830 (--- ^a) C _ε H 6.754, 6.754 (--- ^a)
Val43	7.472 (110.912)	(173.648)	5.483 (58.391)	1.937 (36.427)	C _γ H ₃ 0.775, 0.794 (--- ^a , 22.828)
Lys44	8.739 (118.908)	(176.280)	4.835 (54.240)	1.720, 1.720 (37.042)	C _γ H 1.373, 1.459 (--- ^a) C _δ H 1.642, 1.642 (30.018) C _ε H 2.958, 2.958 (--- ^a)
Lys45	9.190 (126.137)	(--- ^a)	4.500 (58.683)	1.838, 1.997 (33.291)	C _γ H 1.433, 1.712 (26.100) C _δ H 1.642, 1.719 (29.649) C _ε H 2.758, 2.846 (--- ^a)
Leu46	8.320 (125.332)	(176.707)	4.573 (54.088)	1.498, 1.565 (43.729)	C _γ H 1.643 (27.201) C _{δ1} H ₃ 0.867 (25.963) C _{δ2} H ₃ 0.867 (25.963)
Asp47	8.619 (121.309)	(175.795)	4.832 (53.777)	2.747, 2.881 (39.717)	
Ser48	8.312 (116.622)	(175.436)	4.560 (58.373)	3.817, 3.951 (64.689)	
Gly49	8.587 (111.630)	(174.886)	4.071, 4.071 (45.844)		
Thr50	8.145 (112.104)	(175.580)	4.408 (61.973)	4.364 (70.121)	C _γ H ₃ 1.196 (--- ^a)
Gly51	8.424 (110.990)	(174.068)	3.853, 4.085 (45.602)		

Lys52	7.724 (120.068)	(175.851)	4.343 (55.855)	1.649, 1.722 (33.698)	C _γ H 1.344, 1.381 (--- ^a) C _δ H 1.628, 1.628 (29.452) C _ε H 2.965, 2.965 (42.530)
Glu53	8.649 (122.758)	(174.641)	4.534 (55.780)	2.010, 2.122 (29.538)	C _γ H 2.350, 2.461 (33.837)
Leu54	8.438 (124.029)	(177.384)	5.412 (53.436)	1.375, 1.617 (46.225)	C _γ H 1.566 (--- ^a) C _{δ1} H ₃ 0.795 (23.625) C _{δ2} H ₃ 0.837 (25.645)
Val55	9.106 (124.029)	(177.384)	5.239 (58.106)	2.008 (36.203)	C _γ H ₃ 0.804, 1.057 (20.140, 24.276)
Leu56	9.015 (123.355)	(176.869)	5.106 (52.793)	1.388, 1.726 (46.801)	C _γ H 1.264 (--- ^a) C _{δ1} H ₃ 0.803 (26.541) C _{δ2} H ₃ 0.846 (24.357)
Ala57	9.214 (127.357)	(178.476)	4.549 (52.957)	1.611 (19.755)	
Leu58	9.238 (126.607)	(175.258)	3.912 (55.756)	0.526, 1.076 (43.079)	C _γ H 1.324 (--- ^a) C _{δ1} H ₃ 0.593 (21.950) C _{δ2} H ₃ 0.648 (25.551)
Tyr59	7.033 (111.462)	(173.398)	4.691 (54.556)	2.082, 3.075 (43.302)	C _δ H 6.640, 6.640 (--- ^a) C _ε H 6.586, 6.586 (--- ^a)
Asp60	8.281 (117.786)	(176.991)	4.920 (54.128)	2.742, 2.888 (41.044)	
Tyr61	8.877 (119.377)	(173.155)	4.732 (60.027)	2.751, 3.354 (43.365)	C _δ H 7.125, 7.125 (--- ^a) C _ε H 6.987, 6.987 (--- ^a)
Gln62	7.660 (126.124)	(173.764)	4.673 (54.838)	1.916, 1.916 (30.599)	C _γ H 1.994, 1.994 (--- ^a) N _ε H 6.759, 7.432 (112.103)
Glu63	8.180 (122.776)	(175.621)	4.380 (56.131)	2.145, 2.236 (30.130)	C _γ H 2.575, 2.652 (33.383)
Ser64	8.453 (117.598)	(174.943)	4.668 (58.191)	3.869, 3.869 (64.582)	
Gly65	8.360 (111.496)	(173.929)	4.063, 4.220 (45.286)		
Asp66	8.435 (117.852)	(175.372)	4.890 (53.966)	2.812, 2.812 (40.248)	
Asn67	8.053 (117.060)	(174.159)	4.723 (52.331)	2.566, 2.620 (40.204)	N _δ H 6.856, 7.432 (113.444)
Ala68	7.960 (125.697)	(--- ^a)	3.151 (50.551)	0.708 (17.177)	
Pro69		(176.218)	3.716 (62.623)	0.415, 1.027 (31.648)	C _γ H 0.901, 0.901 (27.569) C _δ H --- ^a (--- ^a)

Ser70	8.067 (114.298)	(173.872)	4.316 (58.165)	3.745, 3.888 (63.753)	
Tyr71	6.883 (118.415)	(174.887)	4.866 (55.083)	3.043, 3.253 (39.672)	C ₈ H 7.196, 7.196 (--- ^a) C ₆ H 6.844, 6.844 (--- ^a)
Ser72	8.554 (116.950)	(--- ^a)	4.719 (57.367)	3.817, 3.941 (62.951)	
Pro73		(--- ^a)	4.532 (61.507)	0.661, 1.322 (29.776)	C ₇ H 1.161, 1.686 (26.859) C ₈ H 3.302, 3.567 (50.001)
Pro74		(--- ^a)	4.741 (61.690)	1.999, 2.330 (30.166)	C ₇ H 2.001, 2.196 (27.776) C ₈ H 3.189, 3.871 (50.617)
Pro75		(--- ^a)	4.659 (61.405)	1.839, 2.274 (30.547)	C ₇ H 1.938, 2.001 (27.776) C ₈ H 3.607, 3.731 (--- ^a)
Pro76		(--- ^a)	3.906 (61.578)	1.401, 1.504 30.252	C ₇ H 1.672, 1.965 (27.719) C ₈ H 3.494, 3.703 (50.441)
Pro77		(--- ^a)	4.225 (--- ^a)	1.875, 2.153 (31.700)	C ₇ H 1.725, 1.877 (26.906) C ₈ H 2.055, 3.019 (50.136)

APÉNDICE II: ARTÍCULOS

A binding event converted into a folding event

F.M. Martín-Sierra^a, A.M. Candel^a, S. Casares^a, V.V. Filimonov^{a,b}, J.C. Martínez^{a,*},
F. Conejero-Lara^{a,*}

^aDepartamento de Química Física e Instituto de Biotecnología, Facultad de Ciencias, Universidad de Granada, 18071 Granada, Spain

^bInstitute of Protein Research of the Russian Academy of Sciences, Puschino, Moscow Region 142292, Russia

Received 21 July 2003; revised 9 September 2003; accepted 10 September 2003

First published online 25 September 2003

Edited by Thomas L. James

Abstract We have designed a chimeric protein by connecting a circular permutant of the α -spectrin SH3 domain to the proline-rich decapeptide APSYSPPPPP with a three-residue link. Our aim was to obtain a single-chain protein with a tertiary fold that would mimic the binding between SH3 domains and proline-rich peptides. A comparison of the circular-dichroism and fluorescence spectra of the purified chimera and the SH3 circular permutant showed that the proline-rich sequence occupies the putative SH3 binding site in a similar conformation and with comparable interactions to those found in complexes between SH3 and proline-rich peptides. Differential scanning calorimetry indicated that the interactions in the binding motif interface are highly cooperative with the rest of the structure and thus the protein unfolds in a two-state process. The chimera is more stable than the circular permutant SH3 by 6–8 kJ mol⁻¹ at 25°C and the difference in their unfolding enthalpy is approximately 32 kJ mol⁻¹, which coincides with the values found for the binding of proline-rich peptides to SH3 domains. This type of chimeric protein may be useful in designing SH3 peptide ligands with improved affinity and specificity.

© 2003 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

Key words: Src-homology region 3 domain; Proline-rich peptide; Protein–ligand binding; Protein folding; Protein stability; Differential scanning calorimetry

1. Introduction

Src-homology region 3 (SH3) domains are probably the most widely known modules for molecular recognition [1], which has resulted in their becoming important targets for ligand design. They interact with proline-rich sequences, with a preserved PxxP binding motif, adopting a poly-proline II (PPII) helical conformation. The binding site of SH3 domains consists of a hydrophobic surface containing three shallow pockets defined by preserved aromatic residues [2]. Two of the pockets accommodate each of the prolines, accompa-

nied by a hydrophobic residue in the PxxP motif. The third pocket, known as the ‘specificity pocket’, is flanked by the RT and n-src loops, which play an important role in both the affinity and the specificity of the interaction [1].

An understanding of the thermodynamic determinants of the affinity and specificity of the SH3 domains is of particular importance in being able to predict ligands for proteins of particular interest and, in some cases, to design ligands for the specific inhibition of interactions in vivo. The thermodynamic dissection of protein–ligand interactions involves laborious titration experiments with systematically modified ligands, usually carried out by isothermal titration calorimetry (ITC) [3]. In the case of SH3 domains this implies the use of considerable quantities of protein and expensive synthetic peptides, which may explain why only a few calorimetric studies into the binding between SH3 domains and proline-rich peptides have been undertaken [4–7].

The folding properties of SH3 domains are also well characterised both by experimental [8–14] and computational methods [15–18]. Although the folding kinetics of this family of small domains appears to be mainly determined by the chain topology of the native state [14,15,19,20], it has been shown that this topology can be changed in the SH3 domain of α -spectrin by circular permutation without affecting the overall fold [19,21]. The RT-loop of one of the circular permutants, S19P20s, is open (cf. Fig. 1A) and the new C-terminus is quite close to the binding site for proline-rich peptides, adjacent to the specificity pocket. This topology allows us to engineer chimeric variants of this circular permutant with an extended chain sequence that mimics in its fold the interactions with proline-rich peptides. On the basis of this approach it should be possible to design new interactions at the domain–ligand interface by simple site-directed mutagenesis in these chimeras and to characterise the effects produced by studying the thermodynamics of their folding–unfolding.

We describe here the design, isolation and characterisation of a chimeric protein obtained by connecting the circular permutant S19P20s of the SH3 domain of α -spectrin to the sequence of a proline-rich ligand designed previously for the Abl-SH3 domain [22]. The results highlight the common thermodynamic features between protein–ligand binding and protein folding.

2. Materials and methods

The gene encoding the chimeric proteins was obtained by polymerase chain reaction (PCR) [23], extending the 3'-end DNA sequence of the circular permutant S19P20s to obtain the following amino acid

*Corresponding author. Fax: (34)-958-272879.

E-mail addresses: jcmh@ugr.es (J.C. Martínez), conejero@ugr.es (F. Conejero-Lara).

Abbreviations: SH3, Src-homology region 3; p41, decapeptide of sequence APSYSPPPPP; S19P20s, circular permutant of the α -spectrin SH3 domain cut between positions S19 and P20 with its N- and C-ends linked; SPCp41, chimeric S19P20s elongated from its C-terminus with a three-residue link of sequence DGN plus the p41 sequence; DSC, differential scanning calorimetry; CD, circular dichroism

sequence: MGPREVTMKKGDILLLNSTNKDWWKVEVNDQRQ-GFVPAAYVKKLDSTGKELVLLALYDYQEKS_{xxx}APSYSPPPPP, where the underlined residues are the new amino acids added to the C-terminus of the original S19P20s sequence and xxx represents a two- or three-residue connection of variable composition. The DNA fragments were cloned into a pBAT4 plasmid [24] and expressed into *Escherichia coli* BL21/DE3 (Novagen). The proteins were purified as described elsewhere [25].

Prior to the experiments the samples were thoroughly dialysed against the appropriate buffers. Protein concentration was measured using extinction coefficients determined by Gill and von Hippel's method [26].

Differential scanning calorimetry (DSC) experiments were performed in a VP-DSC microcalorimeter (Microcal) at a scan rate of 90°C/h. The protein concentration in all DSC experiments was ≈1 mg/ml. The temperature dependence of the molar partial heat capacity (C_p) of SH3 was calculated from the DSC data and analysed using Origin 6.1 (OriginLab). C_p curves were fitted by a non-linear least-squares method using the two-state unfolding model as described elsewhere [19].

Titration experiments following the change in intensity of tryptophan fluorescence were carried out to monitor the binding of decapeptide p41 to the SH3 variants. Protein samples were thermostatised at 25°C in the cuvette of the fluorimeter at a concentration of ≈25 μM. Fluorescence spectra were recorded between 305 and 400 nm at different p41 concentrations from 0 to ≈1 mM. The excitation wavelength was 298 nm. The titration curves were analysed as described elsewhere [22].

Circular dichroism (CD) measurements were made in a temperature-controlled Jasco J-715 spectropolarimeter. Far-UV and near-UV CD spectra were recorded at 25°C in 20 mM glycine at pH 3.0 using cuvettes of 1 and 5 mm path length, respectively. Sample concentrations were ≈0.2–0.25 mg/ml for the far-UV and ≈2–2.5 mg/ml for the near-UV spectra. Baseline spectra obtained with pure buffer were subtracted for all spectra. The CD signal was expressed as molar ellipticity for the sake of comparison.

The thermal unfolding of S19P20s and SPCp41 was followed by recording the CD signal in both the far-UV (234 nm) and near-UV (295 nm) ranges at different temperatures between 5 and 95°C. The thermal profiles were corrected by subtracting the profiles obtained with pure buffers under identical conditions. The unfolding profiles were analysed according to the two-state model as reported elsewhere [27].

Computer analysis, structural alignment and modelling of protein structures were performed with the programme Swiss PDB viewer [28].

3. Results and discussion

The aim of this work was to design a single-chain chimeric protein that would mimic in its tertiary fold the typical interactions between proline-rich peptides and SH3 domains. Our first step was to make a computer model of this chimera using the crystal structure of the complex between the Abl-SH3 domain and the decapeptide p41 as a template [29]. Initially, we aligned the crystal structures of the Abl-SH3 domain in its complex with p41 (PDB: 1BBZ) and the circular permutant S19P20s (PDB: 1TUC) [21]. The root mean square deviation (RMSD) for all the backbone atoms was 1.37 Å. The main side chains in the consensus binding site of the SH3 domains were well conserved between Abl-SH3 and S19P20s (Tyr7, Phe9, Trp36, Trp47, Pro49 and Tyr52 for Abl-SH3, versus Tyr57, Tyr59, Trp24, Phe35, Pro37 and Tyr40 in S19P20s) and the spatial arrangement of the side chains was very similar for all of them (Fig. 1B). Subsequently we made a model of the complex between p41 and S19P20s by replacing the Abl-SH3 domain with S19P20s and finally we connected the S19P20s and p41 main chains with a short link of two to three residues, as illustrated in Fig. 1B. No clashes in the model were found. We did not attempt to model any additional

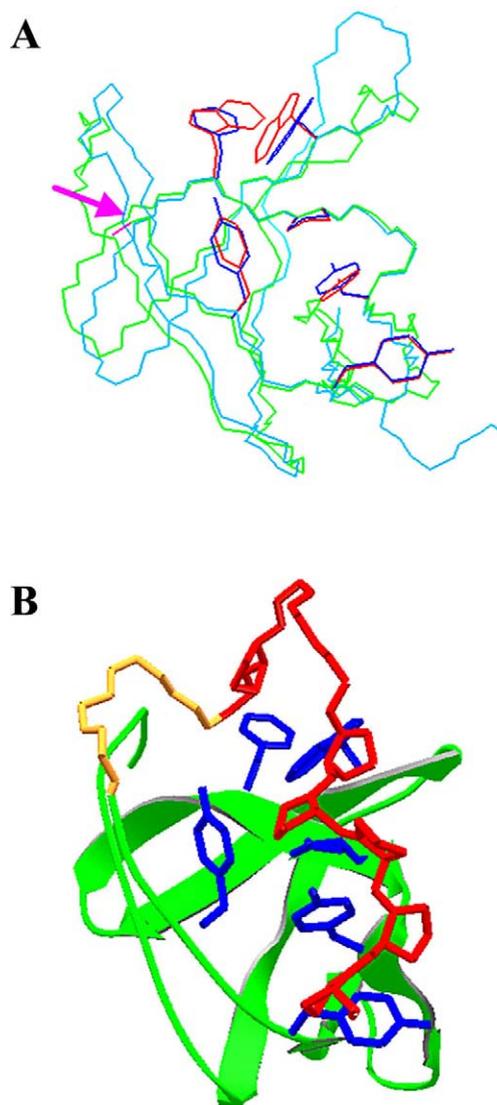


Fig. 1. A: Schematic view of the backbone chains in the alignment between the crystal structures of the Abl-SH3 domain in its complex with the decapeptide p41 (cyan) (PDB: 1BBZ) [29] and the circular permutant S19P20s (green) (PDB: 1TUC) [21]. The side chains forming the SH3 binding site have been highlighted for both domains (Abl-SH3 in red and S19P20s in blue). The arrow indicates the location of the C-terminal end in S19P20s. B: Schematic view showing a model of the structure of the chimeric domain SPCp41. The chain sequence corresponding to S19P20s is represented with ribbons. The decapeptide p41 and the backbone of the three-residue link connecting it with the S19P20s chain are coloured red and orange, respectively. The side chains of residues in the SH3 domain interacting with the p41 sequence are shown in blue.

interactions involving the two to three residues link. For this link we tried several sequences. The designed chimeras were cloned and expressed in *E. coli* and purified to homogeneity as stated in Section 2.

We made preliminary tests of the thermal stability of all the chimeric proteins compared to S19P20s by running single DSC experiments in 20 mM glycine at pH 3.5. All the chimeras were more stable than S19P20s by 5–10°C, suggesting the formation of stabilising interactions between the SH3 domain and the extended sequence. The most stable chimera contained the link GDN and the deletion of K62, close to the

C-terminus in the S19P20s sequence. This deletion was intended to release any possible constraints produced by charge–charge interactions in the region of the link. We named this chimera SPCp41.

To characterise the conformation of the extended chain sequence within SPCp41 we measured the CD spectra of SPCp41, S19P20s and the isolated p41 peptide at pH 3.0 and 25°C (Fig. 2) in both the far-UV (panel A) and near-UV (panel B) wavelength ranges, and calculated the difference spectra between SPCp41 and S19P20s. The far-UV CD spectrum of S19P20s is very similar to that of the wild-type α -spectrin SH3 domain [8], which accords with the similarity in their structures [21,30], whilst the spectrum of SPCp41 has a larger negative ellipticity. The far-UV CD spectrum of isolated p41 is quite similar to those reported elsewhere for other

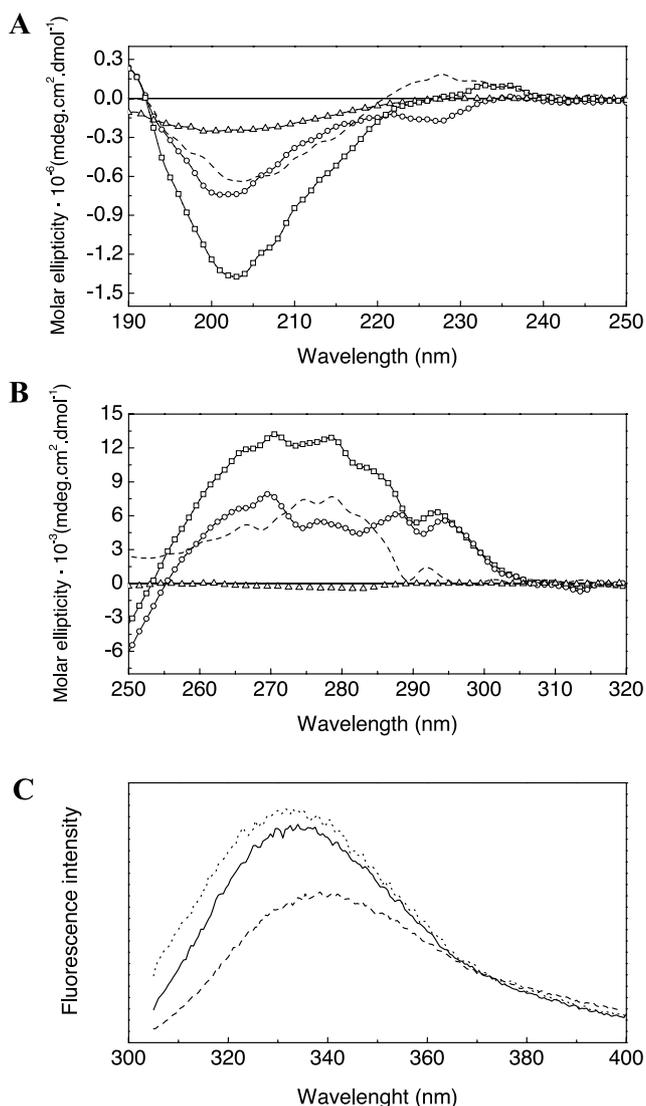


Fig. 2. A: Far-UV CD spectra of SPCp41 (squares), S19P20s (circles) and the isolated p41 decapeptide (triangles) in 20 mM glycine buffer, pH 3.0, 25°C. The dashed line represents the difference spectra between SPCp41 and S19P20s. B: Near-UV CD spectra for the same protein systems and conditions as those in panel A. Symbols and lines are also the same. C: Fluorescence spectra of SPCp41 (solid line), S19P20s (dashed line) and S19P20s in the presence of 1.5 mM of p41 (dotted line) under the same conditions as those of panels A and B. Excitation wavelength = 298 nm.

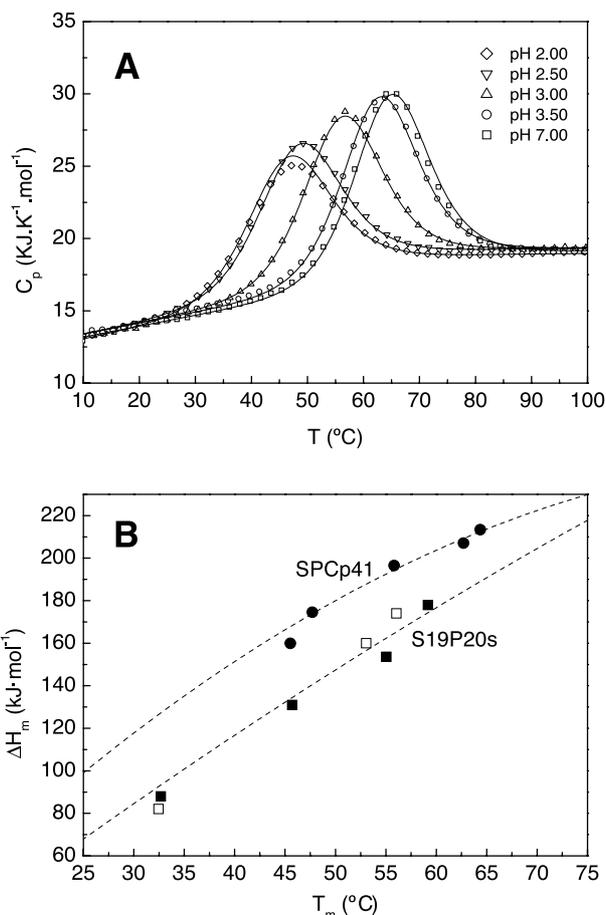


Fig. 3. A: Temperature dependence of the partial molar heat capacity, C_p , of SPCp41 at several pH values, as indicated in the plot. The experimental data are represented by open symbols. The solid lines represent the non-linear least-squares fitting to the whole set of C_p curves using the two-state unfolding model, in which the heat capacity functions for the native and the unfolded states (dashed lines) and the enthalpy change upon unfolding are considered to be independent of pH. B: Plot of the increase in enthalpy upon unfolding, ΔH_m , versus the unfolding temperature, T_m , for SPCp41 (circles) and S19P20s (squares). Open squares correspond to the data reported elsewhere for S19P20s [19]. The dashed lines represent the temperature dependence of the unfolding enthalpy obtained directly from the overall fittings using the two-state model.

proline-rich peptides in solution [31], compatible with a partial propensity to adopt a PPII helical conformation. In contrast, the difference spectrum between SPCp41 and S19P20s, which mainly characterises the backbone conformation of the chain extension, shows both in shape and intensity the typical characteristics of the CD spectrum of a PPII helix, with a positive band at 228 nm and a negative band at 204 nm [32,33]. This is consistent with the conformation usually observed for proline-rich peptides bound to SH3 domains and also found in the conformation of p41 in complex with Abl-SH3 [29].

The near-UV CD spectrum of SPCp41 has a clearly greater positive ellipticity than that of S19P20s. This difference is most pronounced at around 278 nm, which indicates the immobilisation of the tyrosine side chains in SPCp41 within the binding pocket of the SH3 motif due to contacts with the 13-residue extension, which accords well with the interactions observed between p41 and several aromatic residues of Abl-

SH3 in the crystal structure of the complex [29]. An additional, though smaller, difference in CD occurs at around 294 nm, indicating changes in the environment of a tryptophan side chain, most probably Trp24.

Titration of S19P20s with p41 monitored by tryptophan fluorescence produced a clear increase in intensity and a blue shift at the fluorescence spectrum's maximum, which was undoubtedly caused by the binding of p41 to the SH3 binding site (data not shown), as observed elsewhere for the Abl-SH3 domain [22]. This indicates that S19P20s is capable of binding the decapeptide p41, although with quite low affinity ($K_D = 160 \pm 34 \mu\text{M}$). The fluorescence spectrum of SPCp41, on the other hand, was completely unaffected by the addition of increasing amounts of p41, suggesting that the binding site was occupied by the extended sequence in the chimera. The fluorescence spectrum of SPCp41 was very similar to that of S19P20s saturated with p41 (Fig. 2C), indicating that the environments of Trp24 in the chimera SPCp41 and in the S19P20s–p41 complex are very similar.

We made a more detailed characterisation of the thermal unfolding of SPCp41 by DSC. Fig. 3A shows the heat-capacity profiles, C_p , obtained at different pH values between 2.0 and 7.0. The curves were fitted using the two-state unfolding model with the same method employed previously for the analysis of the DSC curves of the circular permutant S19P20s [19]. The quality of the fitting indicates that SPCp41 unfolds in a two-state process under all the conditions investigated. This result also reveals that the interactions between the SH3 motif and the p41 sequence are highly cooperative with the rest of the domain's structure.

For comparison, we made additional DSC experiments with S19P20s under the same conditions used for SPCp41. Table 1 shows the thermodynamic parameters of the unfolding of these two proteins at different pH values. The unfolding parameters for S19P20s are similar to those found in a previous work [19]. Fig. 3B shows the correlation between the unfolding enthalpy, ΔH_m , and the transition temperature, T_m , for both proteins. It is clear from the plot that the unfolding processes of the two proteins have quite different enthalpy functions, the unfolding enthalpy of the chimeric protein being significantly higher than that of S19P20s. The difference in enthalpy between the two correlations was about 32 kJ mol^{-1} at 50°C and hardly changed with temperature (Fig. 3B). This difference with opposite sign would account for the net enthalpy balance of the interactions between the p41

sequence and the SH3 binding site. This enthalpy is of similar magnitude and sign to the enthalpy of binding of the isolated p41 decapeptide to the wild-type α -spectrin SH3 domain measured directly by ITC (results not shown) and is also similar to the enthalpies found for the binding of other proline-rich peptides to different SH3 domains [6,7], which implies that the contribution of new interactions involving the cross-link sequence to the enthalpy of unfolding of SPCp41 is negligible.

The thermal unfolding of SPCp41 and S19P20s was also followed by CD both in the far-UV and near-UV wavelength ranges at pH 3.0 (results not shown). All the unfolding profiles are very well described by the two-state model with equal enthalpy and T_m values in the two wavelength ranges and also very similar to those found by DSC, which confirms the two-state character of the unfolding of both proteins.

Table 1 includes the calculated Gibbs energy of unfolding for SPCp41 and S19P20s at 25°C . The difference in stability at 25°C between both species varies between 6.1 and 8.3 kJ mol^{-1} , depending upon the pH. This would in principle be related to the Gibbs energy balance upon the severing of the interactions between the extended sequence and the binding motif in SPCp41.

From the Gibbs energy of folding of S19P20s ($\Delta G_{\text{fold,SH3}}^0 = -6.4 \pm 0.7 \text{ kJ mol}^{-1}$) and the standard Gibbs energy of binding of p41 to S19P20s ($\Delta G_{\text{bind}}^0 = -21.7 \pm 0.5 \text{ kJ mol}^{-1}$), both at pH 3.0 and 25°C , the Gibbs energy of folding and docking of the S19P20s–p41 complex at 1 M standard state can be estimated as: $\Delta G_{\text{chimera}}^0 = \Delta G_{\text{fold,SH3}}^0 + \Delta G_{\text{bind}}^0$, which is equal to $-28.1 \pm 1.2 \text{ kJ mol}^{-1}$. Compared to this, and assuming the absence of enthalpy effects produced by cross-linking, as referred to above, the Gibbs energy of folding of SPCp41 ($-14.6 \pm 1.2 \text{ kJ mol}^{-1}$) can be expressed as: $\Delta G_{\text{chimera}}^0 = \Delta G_{\text{fold,SH3}}^0 + \Delta G_{\text{bind}}^0 - T\Delta S_{\text{link}}^0$, where the last quantity is the contribution to the stability of the system arising from the change in configurational entropy upon the insertion of the three-residue link between the two polypeptide chains. The subtraction of these two Gibbs energy values allows us to estimate $-T\Delta S_{\text{link}}^0$, which is equal to $13.5 \pm 2.4 \text{ kJ mol}^{-1}$. According to this calculation, the introduction of the cross-link would add a large unfavourable entropic term to the stability of the chimera SPCp41 compared to that of the complex between S19P20s and p41. The entropic contribution of introducing a six-residue cross-link in a dimeric system of two associating helices has recently been calculated theoretically [34]. The authors also report a similar entropic destabilisation (13.4 kJ mol^{-1}) upon cross-linking the helices at 1 M standard state and 300 K. This paradoxical result appears to arise from the fact that these magnitudes correspond to the 1 M standard state concentration, which is unrealistically high for macromolecules. For a more realistic reference state (1 μM concentration of reagents) the entropic contribution deriving from cross-linking the helices becomes strongly stabilising (-21 kJ mol^{-1}), as might be expected.

In our system, using as reference state a concentration of reactants of $50 \mu\text{M}$, similar to that of the titration experiments, the Gibbs energy of binding between p41 and S19P20s, ΔG_{bind} , is $2.8 \pm 1.2 \text{ kJ mol}^{-1}$, indicating that at this reference concentration the complex is intrinsically unstable. Under these conditions, $-T\Delta S_{\text{link}}$ is calculated as $-11.0 \pm 2.4 \text{ kJ mol}^{-1}$, which is comparable to the values reported theoretically and experimentally for other systems [34]. Therefore, the cross-link adds extra stability to the interactions at the

Table 1
Thermodynamic parameters for the unfolding of SPCp41 and S19P20s deriving from an analysis of the DSC curves according to the two-state unfolding model

pH	Protein	T_m ($^\circ\text{C}$)	ΔH_m (kJ mol^{-1})	ΔG_U (25°C) (kJ mol^{-1})
2.0	S19P20s	–	–	–
	SPCp41	45.5 ± 0.1	160 ± 1	8.5 ± 0.6
2.5	S19P20s	32.7 ± 0.6	88 ± 3	1.9 ± 0.3
	SPCp41	47.7 ± 0.1	175 ± 1	10.2 ± 0.7
3.0	S19P20s	45.7 ± 0.3	131 ± 2	6.3 ± 0.7
	SPCp41	55.8 ± 0.1	197 ± 1	14.6 ± 1.2
3.5	S19P20s	55.1 ± 0.3	154 ± 2	9.6 ± 1.2
	SPCp41	62.7 ± 0.1	207 ± 1	17.5 ± 1.7
7.0	S19P20s	59.2 ± 0.2	178 ± 2	12.5 ± 1.4
	SPCp41	64.4 ± 0.1	213 ± 1	18.6 ± 1.8

binding interface within the chimera SPCp41, which in turn behaves as a two-state unfolding protein.

In spite of this two-state behaviour, the thermodynamic magnitudes of unfolding of SPCp41 are strongly biased by the thermodynamics of the interactions between the SH3 motif and the p41 sequence, as can be seen in Table 1. Therefore, with all other factors constant, changes in these interactions will be reflected in changes in the thermodynamic parameters of unfolding of the chimera, quantifiable by DSC.

We conclude that the chimeric protein described in this work mimics the structural and energetic features of the binding process between the SH3 domain and the ligand p41, with a PPII helix conformation for the p41 sequence, a similar immobilisation of aromatic residues at the binding interface and an interaction enthalpy at the interface analogous to that reported for the binding of proline-rich peptides to other SH3 domains.

There are two comparable examples in the literature of the intra-molecular binding of a proline-rich sequence to an adjacent SH3 domain, one being part of the natural regulatory mechanism of the T-cell-specific kinase ITK [35] and the other an artificial fusion between the Hck-SH3 domain and a proline-rich sequence of human GAP [36]. In these proteins, however, the proline-rich sequences were connected by a long polypeptide segment to the N- and C-terminus of the SH3 motif, respectively. The SPCp41 chimera, on the other hand, has a novel topology and a very short link of three residues, as a result of using a circular permutant of the SH3 domain, thus bringing its C-terminus very close to the binding site. In all cases the ligand sequence occupies intra-molecularly the binding site favoured by the removal of configurational entropy for the bi-molecular binding reaction. One of the reasons for the high unfolding cooperativity of SPCp41 may be the shortness of the link, which would result in a relatively small conformational entropy loss upon the immobilisation of the extended sequence at the binding site. Consequently, the thermal unfolding of this chimera agrees with a simple two-state process, which is easy to characterise energetically by typical unfolding experiments such as DSC.

In addition to a general interest in the design of these chimeric constructions in protein folding, the possibility also exists of their being used as tools to optimise interactions at the binding site of SH3 by site-directed mutagenesis and simple stability measurements. This may short-cut the traditionally inefficient and expensive procedures used until now to design peptides with improved affinity and specificity for SH3 domains.

Acknowledgements: This work has been financed by Grants HPRN-CT-2002-00241 from the European Union and BIO2000-1459 from the Spanish Ministry of Science and Technology. We also thank Dr. J. Trout for revising the English text.

References

- [1] Mayer, B.J. (2001) *J. Cell Sci.* 114, 1253–1263.
- [2] Kay, B.K., Williamson, M.P. and Sudol, M. (2000) *FASEB J.* 14, 231–241.
- [3] Leavitt, S. and Freire, E. (2001) *Curr. Opin. Struct. Biol.* 11, 560–566.
- [4] Arold, S.T., Ulmer, T.S., Mulhern, T.D., Werner, J.M., Ladbury, J.E., Campbell, I.D. and Noble, M.E. (2001) *J. Biol. Chem.* 276, 17199–17205.
- [5] Ferreon, J.C. and Hilser, V.J. (2003) *Protein Sci.* 12, 447–457.
- [6] Lemmon, M.A., Ladbury, J.E., Mandiyan, V., Zhou, M. and Schlessinger, J. (1994) *J. Biol. Chem.* 269, 31653–31658.
- [7] Wang, C., Pawley, N.H. and Nicholson, L.K. (2001) *J. Mol. Biol.* 313, 873–887.
- [8] Viguera, A.R., Martinez, J.C., Filimonov, V.V., Mateo, P.L. and Serrano, L. (1994) *Biochemistry* 33, 2142–2150.
- [9] Grantcharova, V.P. and Baker, D. (1997) *Biochemistry* 36, 15685–15692.
- [10] Grantcharova, V.P., Riddle, D.S., Santiago, J.V. and Baker, D. (1998) *Nat. Struct. Biol.* 5, 714–720.
- [11] Martinez, J.C., Pisabarro, M.T. and Serrano, L. (1998) *Nat. Struct. Biol.* 5, 721–729.
- [12] Plaxco, K.W., Gujjarro, J.I., Morton, C.J., Pitkeathly, M., Campbell, I.D. and Dobson, C.M. (1998) *Biochemistry* 37, 2529–2537.
- [13] Gujjarro, J.I., Morton, C.J., Plaxco, K.W., Campbell, I.D. and Dobson, C.M. (1998) *J. Mol. Biol.* 276, 657–667.
- [14] Martinez, J.C. and Serrano, L. (1999) *Nat. Struct. Biol.* 6, 1010–1016.
- [15] Riddle, D.S., Grantcharova, V.P., Santiago, J.V., Alm, E., Ruczinski, I. and Baker, D. (1999) *Nat. Struct. Biol.* 6, 1016–1024.
- [16] Ding, F., Dokholyan, N.V., Buldyrev, S.V., Stanley, H.E. and Shakhnovich, E.I. (2002) *Biophys. J.* 83, 3525–3532.
- [17] Shea, J.E., Onuchic, J.N. and Brooks, C.L. (2002) *Proc. Natl. Acad. Sci. USA* 99, 16064–16068.
- [18] Borreguero, J.M., Dokholyan, N.V., Buldyrev, S.V., Shakhnovich, E.I. and Stanley, H.E. (2002) *J. Mol. Biol.* 318, 863–876.
- [19] Martinez, J.C., Viguera, A.R., Berisio, R., Wilmanns, M., Mateo, P.L., Filimonov, V.V. and Serrano, L. (1999) *Biochemistry* 38, 549–559.
- [20] Gsponer, J. and Caflisch, A. (2001) *J. Mol. Biol.* 309, 285–298.
- [21] Viguera, A.R., Serrano, L. and Wilmanns, M. (1996) *Nat. Struct. Biol.* 3, 874–880.
- [22] Pisabarro, M.T. and Serrano, L. (1996) *Biochemistry* 35, 10634–10640.
- [23] Kunkel, T.A. (1985) *Proc. Natl. Acad. Sci. USA* 82, 488–492.
- [24] Peranen, J., Rikkonen, M., Hyvonen, M. and Kaariainen, L. (1996) *Anal. Biochem.* 236, 371–373.
- [25] Sadqi, M., Casares, S., Abril, M.A., Lopez-Mayorga, O., Conejero-Lara, F. and Freire, E. (1999) *Biochemistry* 38, 8899–8906.
- [26] Gill, S.C. and Vonhippel, P.H. (1989) *Anal. Biochem.* 182, 319–326.
- [27] Cobos, E.S., Filimonov, V.V., Galvez, A., Valdivia, E., Maqueda, M., Martinez, J.C. and Mateo, P.L. (2002) *Biochim. Biophys. Acta* 1598, 98–107.
- [28] Guex, N. and Peitsch, M.C. (1997) *Electrophoresis* 18, 2714–2723.
- [29] Pisabarro, M.T., Serrano, L. and Wilmanns, M. (1998) *J. Mol. Biol.* 281, 513–521.
- [30] Musacchio, A., Noble, M., Pauptit, R., Wierenga, R. and Saraste, M. (1992) *Nature* 359, 851–855.
- [31] Viguera, A.R., Arrondo, J.L., Musacchio, A., Saraste, M. and Serrano, L. (1994) *Biochemistry* 33, 10925–10933.
- [32] Manning, M.C. and Woody, R.W. (1991) *Biopolymers* 31, 569–586.
- [33] Rabanal, F., Ludevid, M.D., Pons, M. and Giralt, E. (1993) *Biopolymers* 33, 1019–1028.
- [34] Zaman, M.H., Berry, R.S. and Sosnick, T.R. (2002) *Proteins* 48, 341–351.
- [35] Andreotti, A.H., Bunnell, S.C., Feng, S., Berg, L.J. and Schreiber, S.L. (1997) *Nature* 385, 93–97.
- [36] Gmeiner, W.H., Xu, J.Z., Horita, D.A., Smithgall, T.E., Engen, J.R., Smith, D.L. and Byrd, R.A. (2001) *Cell. Biochem. Biophys.* 35, 115–126.

The high-resolution NMR structure of a single-chain chimeric protein mimicking a SH3–peptide complex

Adela M. Candel^a, Francisco Conejero-Lara^a, Jose C. Martinez^a, Nico A.J. van Nuland^{a,*},
Marta Bruix^b

^a *Departamento de Química Física e Instituto de Biotecnología, Facultad de Ciencias, Universidad de Granada, Campus Fuentenueva s/n, 18071 Granada, Spain*

^b *Instituto de Química-Física “Rocasolano”, Consejo Superior de Investigaciones Científicas, Serrano 119, 28006 Madrid, Spain*

Received 30 November 2006; revised 23 December 2006; accepted 11 January 2007

Available online 22 January 2007

Edited by Miguel De la Rosa

Abstract Here we present the high-resolution NMR structure of a chimera (SPCp41) between α -spectrin SH3 domain and the decapeptide p41. The tertiary structure mimics perfectly the interactions typically found in SH3–peptide complexes and is remarkably similar to that of the complex between the separate Spc-SH3 domain and ligand p41. Relaxation data confirm the tight binding between the ligand and SH3 part of the chimera. This chimera will serve as a tool for a deeper understanding of the relationship between structure and thermodynamics of binding using a combination of NMR, stability and site-directed mutagenesis studies, which can lead to an effective strategy for ligand design.

© 2007 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

Keywords: NMR; Protein structure; Chimeric protein; SH3; Proline-rich peptide

1. Introduction

Adapter domains that recognize *proline-rich* sequences are considered to be of key importance as they mediate transient protein–protein interactions that are involved in signal transduction. The Src-homology region-3 (SH3) domains are an example of these classes of small protein modules. Their interactions with *polyproline* peptides have been found to be highly promiscuous and relatively weak, with binding affinities ranging between 5 μ M and 100 μ M [1,2]. Several residues of the peptide are in a polyproline II (PPII) helical conformation within the complex that is positioned on top the hydrophobic surface of the SH3 domain defined by the side chains of preserved aromatic residues, flanked by the n-Src and RT loops,

which play an important role in both the affinity and the specificity of binding [3–5].

SH3 domains are known to be involved in a variety of diseases [3,4,6,7] and thus have become very important targets for drug design [7,8]. The success in the design of high-affinity ligands for these small domains is, however, limited as the structural-thermodynamic determinants of binding affinity and specificity in these interactions are not well understood. One of the reasons for the relatively slow progress in understanding the determinants of SH3–peptide binding affinity and specificity is the difficulty in linking the structural information with a quantitative accounting of the energetics of these interactions [5,9–14]. This is due, in part, to the difficulty in determining high-resolution structures of low-affinity SH3–ligand complexes, which are often difficult to crystallize for X-ray diffraction and show few NOEs between the binding partners in nuclear magnetic resonance (NMR) studies. Additionally, most efforts to elucidate the determinants of binding have generally relied upon laborious titration studies, mainly by isothermal titration calorimetry, involving the use of vast amounts of protein and costly systematically-modified synthetic peptides, which explains the relatively small amount of this kind of studies in the literature. One of the possible approaches to circumvent these problems is to link covalently the two binding partners so that the intramolecular SH3–peptide complex is stabilised entropically.

With these aims we have previously designed a chimeric protein, named SPCp41, which includes the decapeptide APSYSPPPPP (p41) covalently attached to the α -spectrin SH3 domain (Spc-SH3) [15]. In this design we used a three-residue linker to attach p41 to a circular permutant of Spc-SH3 (Spc-S19P20s), which has its new C-terminus at the RT-loop, quite close to the binding site for proline-rich peptides and adjacent to the specificity pocket [16]. On the basis of the higher stability observed for SPCp41 relative to Spc-S19P20s and the structural differences observed by CD and fluorescence, we proposed that the p41 sequence occupies the putative SH3 binding site in a similar fashion to that usually found in the SH3–peptide complexes. The p41 ligand was initially designed as a high-affinity ligand for the Abl-SH3 domain [2] and binds the Spc-SH3 domain and some of its mutants with moderate affinity (50–100 μ M) [17,18].

Here we present the high resolution NMR structure of SPCp41 and show that its three-dimensional structure contains the interactions typically found in the complexes between SH3

*Corresponding author. Fax: +34 958 272879.

E-mail address: najvan@ugr.es (N.A.J. van Nuland).

Abbreviations: SH3, Src-homology domain 3; Spc-SH3, SH3 domain of chicken brain α -spectrin; R21A Spc-SH3, the R21A mutant of Spc-SH3; p41, APSYSPPPPP decapeptide; SPCp41, single-chain chimeric protein made by fusion between Spc-SH3 and p41; NMR, nuclear magnetic resonance; NOESY, nuclear Overhauser enhancement spectroscopy; TOCSY, total correlation spectroscopy; RMSD, root mean square deviation

Table 1
Structural statistics for the 20 conformers of SPCp41^a

Restrains ^b	
Short-range ($ i-j \leq 1$)	725
Medium range ($2 \leq i-j \leq 4$)	144
Long range ($ i-j \geq 5$)	712
Torsion angle restraints	
PHI	40
PSI	38
All	1659
Restraint statistics ^c	
Distance violations >0.0 Å	0
Torsion angle violations >0°	0
Target function value (Å ²)	
Average/best	0.066/0.046
Pairwise RMSD ^d of residues 7–77 in Å	
Backbone N, CA, C'	0.34 ± 0.06
Heavy atoms	0.78 ± 0.09
Ramachandran plot ^e	
Most favoured regions (%)	73.5
Additional allowed regions (%)	24.4
Generously allowed regions (%)	0.5
Disallowed regions (%)	1.6

^aThe statistics is obtained from an ensemble of twenty lowest-energy structures for SPCp41 chimera.

^bRestraint statistics reported for unique, unambiguous assigned NOEs.

^cViolations are only reported when present in six or more structures.

^dCoordinate precision is given as the average pair-wise Cartesian coordinate root mean square deviations over the ensemble.

^eValues obtained from the PROCHECK-NMR analysis [27] excluding the first seven highly flexible residues.

domains and *proline-rich* peptides. Furthermore, the SH3:p41 interactions are closely similar to those in the complex between the separate ligand p41 and the R21A mutant of Spc-SH3 [17]. This chimera will now facilitate extensive structure-based thermodynamic studies of the interactions at the SH3:peptide interface by the use of site-directed mutagenesis in combination with stability measurements and high-resolution NMR.

2. Materials and methods

2.1. Protein expression and purification

The plasmid pETM-11 (Protein Expression and Purification Core Facility, EMBL Heidelberg, Germany) containing the chimera SPCp41 gene covalently linked to a N-terminal 6×His tag and a TEV protease cleavage site was expressed in *Escherichia coli* BL21(DE3) strain (Novagen). Protein expression was induced (OD₆₀₀ = 0.6–0.8) with 1 mM IPTG at 37 °C overnight. Cells were resuspended in 50 mM phosphate, 0.3 M NaCl, pH 8.0 buffer (Column Buffer: CB) and broken in a French press. The supernatant was loaded on 5 mL of Ni-NTA resin (Qiagen) equilibrated with CB and washed with CB, CB + 20 mM imidazole and CB + 50 mM imidazole. The protein was eluted with CB + 500 mM imidazole and dialyzed in TEV buffer (50 mM Tris-HCl pH 8.0, 0.5 mM EDTA) to eliminate imidazole. The His-tag was cleaved by incubation at room temperature overnight with TEV protease in presence of 1 mM dithiothreitol. The cleaved SPCp41 was recovered by a second Ni-NTA chromatography step. ¹⁵N- and ¹³C/¹⁵N-labelled proteins were purified from cells grown in M9 minimal medium containing either ¹⁵NH₄Cl (Tracer Tecnologias Analiticas S.L.) or ¹⁵NH₄Cl and (¹³C₆)_D-glucose (Spectra Stable Isotopes) as the sole sources of nitrogen and carbon, respectively. Sample concentration was determined spectrophotometrically using an extinction coefficient of 16 502 M⁻¹ cm⁻¹, determined using the Gill and von Hippel method [19].

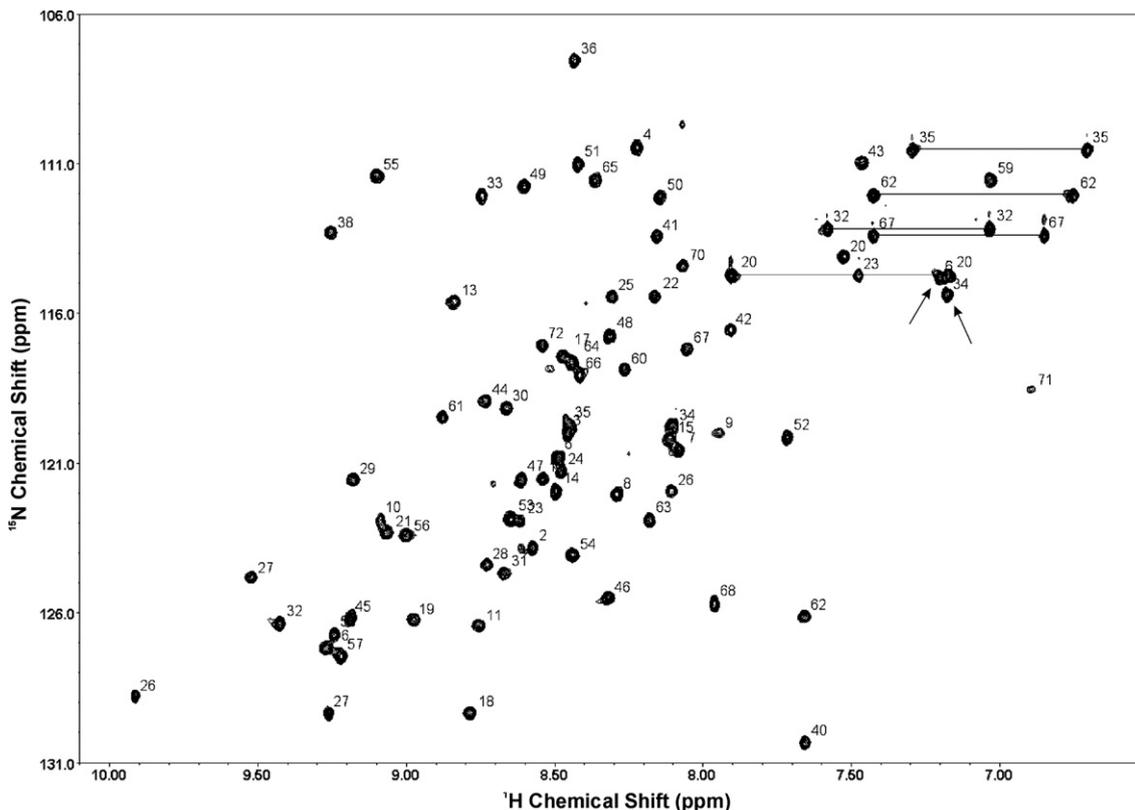


Fig. 1. 800 MHz ¹H–¹⁵N HSQC spectrum of the SPCp41 chimera recorded at 30 °C. Assignments are indicated by the corresponding residue number in the amino acid sequence. The two side chain NH₂ protons of Asn and Gln residues are connected by a solid line in the figure. The ¹Hε–¹⁵Nε cross-peaks of the two arginine residues are indicated by an arrow.

2.2. NMR spectroscopy

All NMR spectra were recorded on 1 mM SPCp41 samples in 93% H₂O/7% D₂O or in D₂O in 20 mM D5-glycine buffer pH 3.5 at 30 °C on a Bruker US2 800 MHz spectrometer. Two-dimensional homonuclear ¹H total correlation spectroscopy (2D TOCSY, 60 ms mixing time) [20] and nuclear Overhauser enhancement spectroscopy (2D NOESY, 100 ms mixing time) [21] spectra were acquired on an unlabelled sample in H₂O and D₂O. 2D heteronuclear single quantum correlation (¹⁵N-HSQC), ¹⁵N-TOCSY-HSQC and ¹⁵N-NOESY-HSQC [22] were acquired on a ¹⁵N-labelled sample. Triple resonance spectra CBCA(CO)NH, CBCANH, HNCA, HN(CO)CA, HNCO and HCCH-TOCSY were recorded on a ¹³C/¹⁵N-labelled sample [23].

NMR data were processed using XWINNMR/TOPSPIN (Bruker Biospin, Karlsruhe, Germany) or NMRPipe [24], and analyzed by SPARKY (T.D. Goddard and D.G. Kneller, University of California) or NMRVIEW [25].

2.3. Spectral assignment

HNCA, HN(CO)CA, CBCANH, CBCA(CO)NH and HNCO experiments were used for the backbone assignment and HCCH-TOCSY and ¹⁵N-TOCSY-HSQC for the side-chain assignment. The assigned resonances have been deposited in the BioMagResBank database with accession code 15013.

2.4. Structure calculation

NOE cross peaks, obtained from the high-resolution 800 MHz 2D NOESY spectra, were assigned using the automated NOE assignment of CYANA [26]. Phi and psi torsion angle restraints were included based on analysis of HN, ¹⁵N, ¹H_α, ¹³C_α, ¹³CO and ¹³C_β chemical shifts using the program PREDITOR [27]. The standard protocol [26] was used with seven cycles of combined automated NOE assignment and structure calculation of 100 conformers in each cycle, of which the 20 structures with lowest target function value were selected for further minimization and analysis. Input data and structure calculation statistics are summarized in Table 1. PROCHECK-NMR [28] was used to analyze the quality of the structures and MOLMOL [29] and PYMOL (<http://www.pymol.org>) for visualization. The structural coordinates and experimentally derived restraints have been deposited in the Protein Data Bank with accession number 2JMC.

Solvent-accessible surface areas (ASA) were calculated using the program NACCESS [30] using the default parameters for the water probe radius (1.4 Å), z-slice (0.05 Å) and van der Waals radii. The buried surface area within the chimera was calculated by assuming residues 1–67 as the SH3 domain and residues 68–77 as the p41 part of the chimeric protein.

2.5. Heteronuclear NOE experiments

¹H-¹⁵N steady-state NOEs of SPCp41 were obtained from two series of 2D experiments with coherence selection achieved by pulse field gradients [31] at 30 °C at 800 MHz and on a Varian NMR Systems

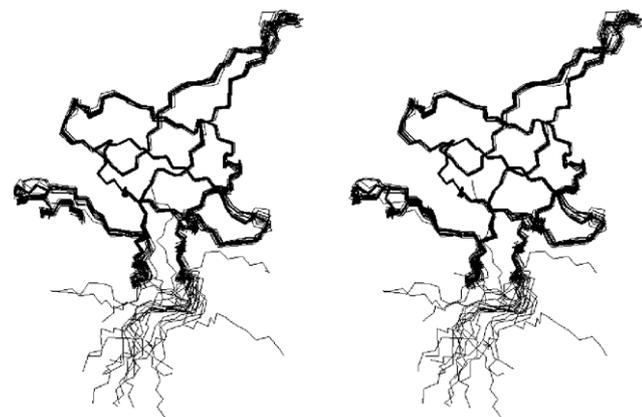


Fig. 2. Stereo view of the ensemble of 20 lowest-energy NMR structures of the SPCp41 chimera superimposed on the backbone atoms (N, Ca, C') of residues 7–77.

600 MHz spectrometer at 25 °C. The ¹H-¹⁵N heteronuclear NOEs were determined from the ratio of peak intensities with and without the saturation of the amide protons for 3 s.

3. Results and discussion

3.1. The structure of the chimera SPCp41

¹H, ¹⁵N and ¹³C assignments were obtained by the use of the combination of selected triple-resonance spectra. Fig. 1 shows a ¹N-¹⁵N HSQC spectrum of SPCp41 with the ¹H and ¹⁵N assignment indicated. Fig. 2 shows a stereo representation of the superposition of the ensemble of 20 SPCp41 structures derived from the automatic NOE assignment protocol in CYANA (see Section 2). The ensemble fulfils very well the experimental data (Table 1). Moreover, the structures satisfy

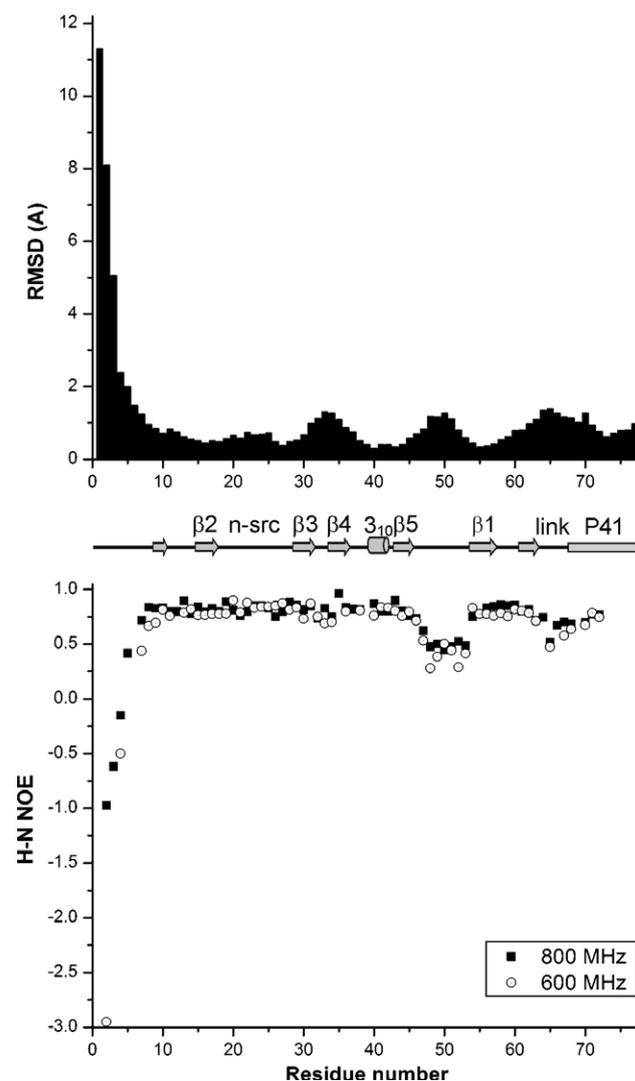


Fig. 3. Comparison of the backbone dynamics and backbone RMSD calculated over the ensemble of 20 structures presented in Fig. 2. (A) Bar diagram of the backbone positional RMSD calculated using MOLMOL. (B) The heteronuclear ¹H-¹⁵N NOEs (H-N NOE) measured at 600 and 800 MHz are presented. Errors are estimated to be in the order of 2%. The different structural elements are shown in the centre of the two panels with β -strands numbered as in WT α -spectrin SH3.

all convergence criteria for successful structure calculations using CYANA [32]. The success of the approach can be attributed to the use of a single high-quality 800 MHz 2D NOESY spectrum and the complete set of assignments of almost all proton frequencies (>95%) used for the automatic NOE assignment stage and the addition of torsion angle restraints derived from the combination of selected ^1H , ^{15}N and ^{13}C chemical shifts. The ^1H – ^{15}N NOE data reveal a good correlation with the backbone root mean square deviations (RMSD) of the calculated ensemble (Fig. 3). The structural disorder observed for the N-terminus, loop 48–52 and the linker region (residues 65–67) connecting the SH3 domain to the p41 ligand (residues 68–77) can be ascribed to the higher flexibility in these regions of the protein as indicated by the smaller heteronuclear NOE values. Moreover, the NOE values following the linker region, which corresponds to the first five residues of the p41 part, increase and are closely similar to those of the more rigid regions of the protein, confirming the tight intramolecular binding between the ligand and SH3 domain part in the chimera.

Fig. 4 shows a superposition of the structure of SPCp41 on that of the circular permutant Spc-S19P20s (pdb code: 1TUC [16]) that was used as template for the chimeric protein. The SH3 part of the structure is very similar to the Spc-S19P20s structure with a backbone pair wise RMSD of 0.87 Å. The structure is also closely similar to that of the R21A mutant of Spc-SH3 (pdb code: 2JMA) that was recently determined by NMR [17] with a backbone pair wise RMSD of 0.83 Å. Moreover, the unfavorable torsion angles for residue 47 found in the WT X-ray structure of Spc-SH3 [33] and in the NMR structure of R21A Spc-SH3 [17] as well as for the corresponding residue in Spc-S19P20s [16] is also present in all structures of the SPCp41 NMR ensemble (residue Asn32). The positive ϕ angle sign for Asn32 was experimentally confirmed by the analysis of a ^{15}N -HSQC derived spectrum [34].

3.2. Comparison between the SPCp41 chimera and the R21A Spc-SH3:p41 and complex

Fig. 5 shows the binding interface of SPCp41 and that of the R21A Spc-SH3:p41 complex [17]. The p41 ligand adopts a

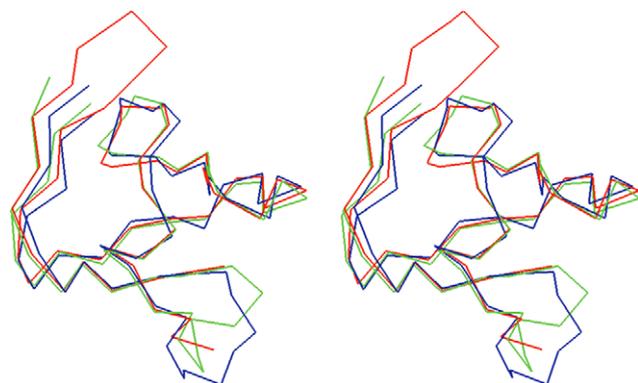


Fig. 4. Stereoview of the lowest-energy SPCp41 structure (blue) superimposed on the WT X-ray structure of the permutant Spc-S19P20s (1TUC, green) and a representative NMR structure of the R21A Spc-SH3:p41 complex (2JMA, only the SH3 domain is shown in red) for comparison. Superposition was done on the backbone N, Ca, C' atoms of 8–46 and 53–62 of SPCp41 on the corresponding residues 23–61 and 7–16 in Spc-S19P20s and in R21A Spc-SH3, respectively.

very similar conformation in both structures (Fig. 5A). Small structural differences are mainly around Tyr4, likely due to adapting to the slightly different conformation of the chain at the RT loop (Lys18 in R21A Spc-SH3 forms a H-bond with Tyr4 of p41). The aromatic residues and Pro39 in SPCp41 (Pro54 in R21A Spc-SH3), highly preserved among SH3 domains and critical for binding *polyproline* motifs, are positioned in the same way in both structures (Fig. 5B). We used LIGPLOT [35] to analyze hydrophobic contacts and hydrogen bonds within the chimera and the R21A Spc-SH3:p41 complex; the results are summarized in Table 2. Generally the same contacts are found in the two protein–ligand complexes, although a few differences are observed. The absence or presence of certain contacts are due to the structural differences around Tyr4 of p41 in relation to the original RT loop in R21A Spc-SH3, partly due to the absence of the H-bond partner in SPCp41 (Lys18 in R21A Spc-SH3). If we assume the p41

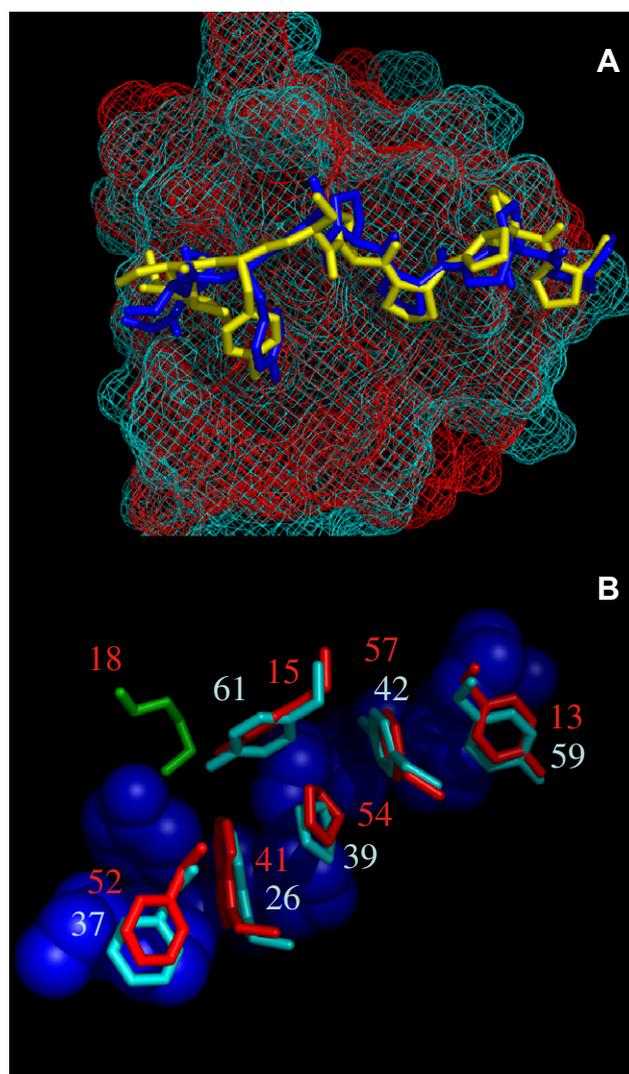


Fig. 5. The SPCp41 and R21A Spc-SH3:p41 binding interface. (A) Stick representation of SPCp41 (p41 blue, SH3 cyan mesh) and p41 in complex with R21A Spc-SH3 (p41 yellow, SH3 red mesh). Structural differences are mainly around Tyr4 likely due to the absence of the H-bond partner in SPCp41 (Lys18 in R21A Spc-SH3). (B) Residues of SH3 involved in binding (SPCp41 red, R21A Spc-SH3 cyan, p41 blue). Residue 18 in R21A Spc-SH3 is colored in green.

Table 2
Intermolecular contacts^a statistics calculated over the ensemble of 20 chimera structures

Interacting residues		Hydrogen bonds			Non-bonded contacts	Occurrence total (%)	Interacting residue in R21A Spc-SH3
p41	Chimera	M–M	S–S	M–S			
Ala1	Phe37	0	0	0	1	35	Phe52 (3, 95%)
Pro2	Trp26	0	0	0	5	100	Trp41 (7, 100%)
Pro2	Phe37	0	0	0	4	100	Phe52 (2, 50%)
Tyr4	Trp26	0	0	0	3	100	Trp41 (1, 95%)
Tyr4	Trp26	0	0	1	0	100	Trp41 (0, 100%)
Pro6	Pro39	0	0	0	1	100	Pro54 (1, 70%)
Pro6	Asp25	0	0	0	2	25	Asp40 (2, 100%)
Pro6	Trp26	0	0	0	2	100	Trp41 (3, 100%)
Pro7	Tyr61	0	0	0	1	95	Tyr15 (1, 80%)
Pro9	Tyr59	0	0	0	3	100	Tyr13 (2, 100%)
Pro9	Tyr42	0	0	0	2	100	Tyr57 (3, 100%)
Pro10	Tyr59	0	0	0	8	100	Tyr13 (2, 65%)
Pro2	Asn23	0	0	0	1	90	–
Ser5	Trp26	0	0	0	1	85	–
Pro7	Trp26	0	0	0	2	100	–
Pro7	Pro39	0	0	0	3	100	–
Tyr4	–	0	1	0	–	–	Lys18 (0, 100%)
Pro7	–	0	0	1	–	–	Tyr57 (0, 95%)

^aIntermolecular contacts were analyzed with LIGPLOT [35] using the default settings. Contacts are reported if present in at least 10 of the 20 structures or in at least 1 of the 20 structures when the corresponding contact in R21A Spc-SH3:P41 is present in 10 or more of the 20 models in pdb entry 2JMA [16]. Hydrogen bonds are reported for main-chain to main-chain (M–M), side-chain to side-chain (S–S) and main-chain to side-chain (M–S). The average number of non-bonded contacts is reported, as well as the percentage of occurrences of a contact. The corresponding interacting residue in R21A Spc-SH3 for a certain contact and the number of contacts and occurrence in the 20 structures (between brackets, respectively) are given in the last column. The last two rows correspond to contacts present in R21A Spc-SH3:P41 that are not present in any of the 20 structures reported here. For comparison, residue numbering of 1–10 is used of the p41 part of the chimera (corresponding to residues 68–77 in SPCp41).

part of the chimera (residues 68–77) separate and as ligand for the SH3 part of the chimera (residues 1–67), the interaction buries $648 \pm 8 \text{ \AA}^2$ of non-polar area and $227 \pm 28 \text{ \AA}^2$ of polar area, slightly less than in the R21A Spc-SH3:P41 complex ($686 \pm 11 \text{ \AA}^2$ and $241 \pm 11 \text{ \AA}^2$, respectively [17]), in good agreement with the structural details of both complexes. These small differences can at least partly be ascribed to the simple model used to calculate the accessible area for the chimera.

3.3. Contribution of cross-linking the protein and ligand on the binding entropy

In a previous article we reported the differences in thermodynamic magnitudes of unfolding of Spc-S19P20s and SPCp41 [15] in comparison with the Gibbs energy of binding of p41 to Spc-S19P20s. It was concluded that the crosslink has a considerable contribution to the configurational entropy of binding ($-T\Delta S$ about -11 kJ/mol) that is crucial to stabilize of the interactions at the binding interface. In a recent paper [36], Zhou has analyzed quantitatively the stabilizing effect of cross-linking between SH3 domains and proline-rich peptides using flexible linker segments. This analysis predicts for SPCp41 a net stabilization of 14.9 kJ/mol relative to Spc-S19P20s, consistent with our experimental results.

The stabilization produced by the crosslink is interpreted in terms of the entropic benefit of increasing the effective concentration of p41 up to about 65 mM . Accordingly, although the isolated p41 ligand and the Spc-S19P20s domain have a low intrinsic affinity ($K_d = 160 \text{ \mu M}$), it is largely compensated by the entropy gain of tethering, so that the chimera behaves very much as a two-state unfolding protein. The results presented in this work confirm the tight intra-molecular binding between the p41 sequence segment and the SH3 binding site, with a well-defined conformation and relatively high heteronuclear NOE values for the residues at the binding interface.

4. Concluding remarks

The success in the design of high-affinity ligands for SH3 domains is up-to-date limited as we do not know how changing a particular interaction at the binding interface modifies the binding Gibbs energy. The work described here shows that we fully succeeded in obtaining a single-chain globular, chimeric protein that mimics the interactions found in the complex between the α -spectrin SH3 domain and the *polyproline* ligand p41. We have now a tool in hand enabling a straightforward approach to increase our understanding on how structure is linked to the thermodynamics of binding by the use of a combination of NMR, stability measurements and site-directed mutagenesis. Such an increased understanding will facilitate an effective strategy of rational design of ligands.

Acknowledgements: This research is funded by Grants BIO2003-04274 and BIO2005-04650 from the Spanish Ministry of Education and Science (MEC), Grant INTAS-03-51-5569 from the European Union and Grant FQM-123 from the Andalusia Regional Government. A.M.C. was supported by a short-term grant from the MEC. N.A.J.v.N. is recipient of a Ramón y Cajal Research Contract from the MEC. The 600 MHz spectra were recorded in the Centre for Scientific Instrumentation (CIC) of the University of Granada.

References

- [1] Viguera, A.R., Arrondo, J.L., Musacchio, A., Saraste, M. and Serrano, L. (1994) Characterization of the interaction of natural proline-rich peptides with five different SH3 domains. *Biochemistry* 33, 10925–10933.
- [2] Pisabarro, M.T. and Serrano, L. (1996) Rational design of specific high-affinity peptide ligands for the Abl-SH3 domain. *Biochemistry* 35, 10634–10640.
- [3] Kay, B.K., Williamson, M.P. and 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.
- [4] Mayer, B.J. (2001) SH3 domains: complexity in moderation. *J. Cell. Sci.* 114, 1253–1263.
- [5] Arold, S.T., Ulmer, T.S., Mulhern, T.D., Werner, J.M., Ladbury, J.E., Campbell, I.D. and 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, 17199–17205.
- [6] Gmeiner, W.H. and Horita, D.A. (2001) Implications of SH3 domain structure and dynamics for protein regulation and drug design. *Cell Biochem. Biophys.* 35, 127–140.
- [7] Vidal, M., Gigoux, V. and Garbay, C. (2001) SH2 and SH3 domains as targets for anti-proliferative agents. *Crit. Rev. Oncol. Hematol.* 40, 175–186.
- [8] Dalgarno, D.C., Botfield, M.C. and Rickles, R.J. (1997) SH3 domains and drug design: ligands, structure, and biological function. *Biopolymers* 43, 383–400.
- [9] Renzoni, D.A., Pugh, D.J., Siligardi, G., Das, P., Morton, C.J., Rossi, C., Waterfield, M.D., Campbell, I.D. and 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, 15646–15653.
- [10] Pisabarro, M.T., Serrano, L. and 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, 513–521.
- [11] Ladbury, J.E. and Arold, S. (2000) Searching for specificity in SH domains. *Chem. Biol.* 7, R3–R8.
- [12] Wang, C., Pawley, N.H. and Nicholson, L.K. (2001) The role of backbone motions in ligand binding to the c-Src SH3 domain. *J. Mol. Biol.* 313, 873–887.
- [13] Ferreon, J.C. and Hilser, V.J. (2004) Thermodynamics of binding to SH3 domains: the energetic impact of polyproline II (PII) helix formation. *Biochemistry* 43, 7787–7797.
- [14] Palencia, A., Cobos, E.S., Mateo, P.L., Martinez, J.C. and Luque, I. (2004) Thermodynamic dissection of the binding energetics of proline-rich peptides to the Abl-SH3 domain: implications for rational ligand design. *J. Mol. Biol.* 336, 527–537.
- [15] Martin-Sierra, F.M., Candel, A.M., Casares, S., Filimonov, V.V., Martinez, J.C. and Conejero-Lara, F. (2003) A binding event converted into a folding event. *FEBS Lett.* 553, 328–332.
- [16] Viguera, A.R., Serrano, L. and Wilmanns, M. (1996) Different folding transition states may result in the same native structure. *Nat. Struct. Biol.* 3, 874–880.
- [17] Casares, S., AB, E., Eshuis, H., Lopez-Mayorga, O., van Nuland, N.A.J. and Conejero-Lara, F. (2006) 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.* (submitted).
- [18] Casares, S., López-Mayorga, O., Vega, M.C., Cámara-Artigas, A. and Conejero-Lara, F. (2006) Cooperative propagation of local stability changes from low-stability and high-stability regions in a SH3 domain. *Proteins* (in press).
- [19] Gill, S.C. and Von Hippel, P.H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* 182, 319–326.
- [20] Bax, A. and Davis, D.G. (1985) MLEV-17 based two-dimensional homonuclear magnetisation transfer spectroscopy. *J. Magn. Res.* 65, 355–360.
- [21] Kumar, A., Ernst, R.R. and Wüthrich, K. (1980) A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton–proton cross-relaxation networks in biological macromolecules. *Biochem. Biophys. Res. Commun.* 95, 1–6.
- [22] Palmer, A.G., Cavanagh, J., Byrd, R.A. and Rance, M. (1992) Sensitivity improvement in three-dimensional heteronuclear correlation NMR spectroscopy. *J. Magn. Res.* 96, 416–424.
- [23] Sattler, M., Schleucher, J. and Griesinger, C. (1999) Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. Nucl. Magn. Reson. Spectrosc.* 34, 93–158.
- [24] Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J. and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* 6, 277–293.
- [25] Johnson, B.A. and Blevins, R.A. (1994) NMRView: a computer program for the visualization and analysis of NMR data. *J. Biomol. NMR* 4, 603–614.
- [26] Herrmann, T., Güntert, P. and Wüthrich, K. (2002) Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J. Mol. Biol.* 319, 209–227.
- [27] Berjanskii, M.V., Neal, S. and Wishart, D.S. (2006) PREDITOR: a web server for predicting protein torsion angle restraints. *Nucleic Acids Res.* 34, W63–W69.
- [28] Laskowski, R.A., Rullmann, J.A., MacArthur, M.W., Kaptein, R. and Thornton, J.M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* 8, 477–486.
- [29] Koradi, R., Billeter, M. and Wüthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* 14, 51–55, 29–32.
- [30] Hubbard, S.J. and Thornton, J.M. (1993) NACCESS, Computer Program, Department of Biochemistry and Molecular Biology, University College, London.
- [31] Farrow, N.A., Muhandiram, R., Singer, A.U., Pascal, S.M., Kay, C.M., Gish, G., Shoelson, S.E., Pawson, T., Forman-Kay, J.D. and Kay, L.E. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ¹⁵N NMR relaxation. *Biochemistry* 33, 5984–6003.
- [32] Jee, J.G. and Güntert, P. (2003) Influence of the completeness of chemical shift assignments on NMR structures obtained with automated NOE assignment. *J. Struct. Funct. Genom.* 4, 179–189.
- [33] Musacchio, A., Noble, M., Pauptit, R., Wierenga, R. and Saraste, M. (1992) Crystal structure of a Src-homology 3 (SH3) domain. *Nature* 359, 851–855.
- [34] Crowley, P., Ubbink, M. and Otting, G. (2000) Phi Angle restraints in protein backbones from dipole–dipole cross-correlation between ¹HN-¹⁵N and ¹HN-¹H vectors. *J. Am. Chem. Soc.* 122, 2968–2969.
- [35] Wallace, A.C., Laskowski, R.A. and Thornton, J.M. (1995) Ligplot – a program to generate schematic diagrams of protein ligand interactions. *Protein Eng.* 8, 127–134.
- [36] Zhou, H-X. (2006) Quantitative relation between intermolecular and intramolecular binding of pro-rich peptides to SH3 domains. *Biophys. J.* 91, 3170–3181.

Analysis of the Thermodynamics of Binding of an SH3 Domain to Proline-rich Peptides using a Chimeric Fusion Protein

Adela M. Candel, Nico A. J. van Nuland, Francisco M. Martin-Sierra
Jose C. Martinez and Francisco Conejero-Lara*

Departamento de Química
Física e Instituto de
Biotecnología, Facultad de
Ciencias, Universidad de
Granada, 18071 Granada
Spain

Received 11 October 2007;
received in revised form
5 November 2007;
accepted 17 November 2007

A complete understanding of the thermodynamic determinants of binding between SH3 domains and proline-rich peptides is crucial to the development of rational strategies for designing ligands for these important domains. Recently we engineered a single-chain chimeric protein by fusing the α -spectrin Src homology region 3 (SH3) domain to the decapeptide APSYSPPPPP (p41). This chimera mimics the structural and energetic features of the interaction between SH3 domains and proline-rich peptides. Here we show that analysing the unfolding thermodynamics of single-point mutants of this chimeric fusion protein constitutes a very useful approach to deciphering the thermodynamics of SH3–ligand interactions. To this end, we investigated the contribution of each proline residue of the ligand sequence to the SH3–peptide interaction by producing six single Pro-Ala mutants of the chimeric protein and analysing their unfolding thermodynamics by differential scanning calorimetry (DSC). Structural analyses of the mutant chimeras by circular dichroism, fluorescence and NMR together with NMR-relaxation measurements indicate conformational flexibility at the binding interface, which is strongly affected by the different Pro-Ala mutations. An analysis of the DSC thermograms on the basis of a three-state unfolding model has allowed us to distinguish and separate the thermodynamic magnitudes of the interaction at the binding interface. The model assumes equilibrium between the “unbound” and “bound” states at the SH3–peptide binding interface. The resulting thermodynamic magnitudes classify the different proline residues according to their importance in the interaction as P2~P7~P10>P9~P6>P8, which agrees well with Lim’s model for the interaction between SH3 domains and proline-rich peptides. In addition, the thermodynamic signature of the interaction is the same as that usually found for this type of binding, with a strong enthalpy–entropy compensation for all the mutants. This compensation appears to derive from an increase in conformational flexibility concomitant to the weakening of the interactions at the binding interface. We conclude that our approach, based on DSC and site-directed mutagenesis analysis of chimeric fusion proteins, may serve as

*Corresponding author. E-mail address: conejero@ugr.es.

Abbreviations used: SH3, Src homology region 3; p41, decapeptide with the sequence APSYSPPPPP; P7Ap, decapeptide with the sequence APSYSPAPPP; S19P20s, circular permutant of the α -spectrin SH3 domain cleaved between positions S19 and P20 with its N and C-terminal ends linked; SPCp41, chimeric S19P20s elongated from its C terminus with a three-residue link with the sequence DGN plus the p41 sequence; P2A, P6A, P7A, P8A, P9A and P10A, mutants of SPCp41 with a proline replaced by alanine at each position of the p41 sequence; PPII, poly-proline II helix; DSC, differential scanning calorimetry; ITC, isothermal titration calorimetry; CD, circular dichroism spectroscopy; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; HSQC, heteronuclear single-quantum coherence; EEC, enthalpy–entropy compensation.

a suitable tool to analyse the energetics of weak biomolecular interactions such as those involving SH3 domains.

© 2007 Published by Elsevier Ltd.

Edited by J. E. Ladbury

Keywords: Src-homology region 3 domain; proline-rich peptide; binding thermodynamics; protein stability; differential scanning calorimetry

Introduction

Src-homology region 3 (SH3) domains have become a very important focus of interest for drug design due to their involvement in transient protein–protein interactions related to a large number of cell processes, many of them connected to pathological conditions.^{1–4} Molecular recognition mediated by SH3 domains is very often an intramolecular event that plays an essential role in the regulation of many signalling proteins, such as, for example, in several families of tyrosine kinases.^{5–7} An understanding of the thermodynamic determinants underlying these interactions is therefore essential if we are to unveil the regulatory mechanisms of these signalling proteins.

SH3 domains recognize proline-rich sequences with low-to-moderate affinity through a highly preserved mode of interaction. Their binding site consists of a hydrophobic surface with three shallow pockets defined by the side-chains of preserved aromatic residues. These pockets are flanked by the n-src and RT loops, which play an important role in both the affinity and the specificity of the interaction.^{2,3,8} One of the most prominent characteristics of this interaction is the polyproline II (PPII) helical conformation of the ligand chain in the complex. This favours the presence of proline residues in the ligand sequences due to a reduction in the entropic cost of achieving the PPII conformation upon binding. Nevertheless, there are certain positions in the ligand where proline is most frequently found, giving rise to consensus sequences for the SH3 ligands around a core sequence $\Phi P x \Phi P$ where Φ is a hydrophobic residue and x can be any residue.² The PPII helix has three residues per turn and a roughly triangular cross-section, with one of its faces lying against the SH3 binding surface. In fact two of the three hydrophobic pockets are occupied by the side-chains of each of the ΦP dipeptides of the core sequence. This has been attributed to the special geometry of the PPII conformation, which favours the hydrophobic packing of the core side-chains in the binding pockets.⁹ Additionally, each of the side-chains of a tryptophan and a tyrosine in the SH3 domain donates a hydrogen bond to the carbonyl groups of the peptide ligand. The third pocket establishes additional interactions with less preserved residues of the ligand, conferring additional specificity to the binding. This pocket, known as the “compass” pocket, also dictates the orientation of the peptide chain on the binding site, giving rise to class-1 or class-2 binding modes.¹⁰

The success of rational-design approaches to high-affinity ligands for SH3 domains has to a great extent been impaired by a lack of understanding of the structural and thermodynamic determinants that govern the binding energy. Indeed the experimental thermodynamic magnitudes of this interaction are often confusing. Isothermal titration calorimetry (ITC) studies into several SH3 domains and their ligands have always revealed that binding produces strongly negative enthalpy changes, which are partially balanced out by negative entropy changes.^{11–15} But these results are not to be expected because the binding interface is primarily hydrophobic. Several recent studies have proposed various possible contributions to the unusual thermodynamic features of this interaction: the reduction in the conformational motions of the protein;¹⁶ the folding of the peptide into the binding-competent PPII conformation;¹² or the establishment of hydrogen-bond networks mediated by water molecules buried at the binding interface.¹⁴ Nevertheless, the relative importance of these possible effects upon the global thermodynamic magnitudes of binding remains unclear.

Many efforts to elucidate the determinants of binding of SH3 domains to their putative ligands have generally relied upon laborious titration studies, mainly by ITC. Due to the low binding affinities of these interactions, ITC studies usually involve large amounts of costly, systematically modified, synthetic peptides and proteins, which might explain why relatively few studies of this kind appear in the literature. Looking for a simpler and more straightforward approach, we designed a single-chain chimeric protein that mimics in its tertiary fold the structural and energetic features of the binding between SH3 domains and proline-rich ligands.¹⁷ This chimera, named SPCp41, was made by fusing a previously engineered, circular permutant of the SH3 domain of α -spectrin (S19P20s)¹⁸ with the proline-rich decapeptide APSYSPPPPP (p41), which was originally designed to bind the Abl-SH3 domain with high-affinity and specificity.¹⁹ The RT-loop of the S19P20s permutant is cleaved and the new C terminus is close to the binding site for proline-rich peptides and adjacent to the specificity pocket, so that the p41 decapeptide could be attached using a three-residue link¹⁷ (Figure 1(a)). Using differential scanning calorimetry (DSC) we showed that the thermodynamic magnitudes of the thermal unfolding of the chimeric protein are strongly biased by interactions at the SH3–peptide binding interface. We recently reported the high-resolution structure of SPCp41 revealed by NMR²⁰

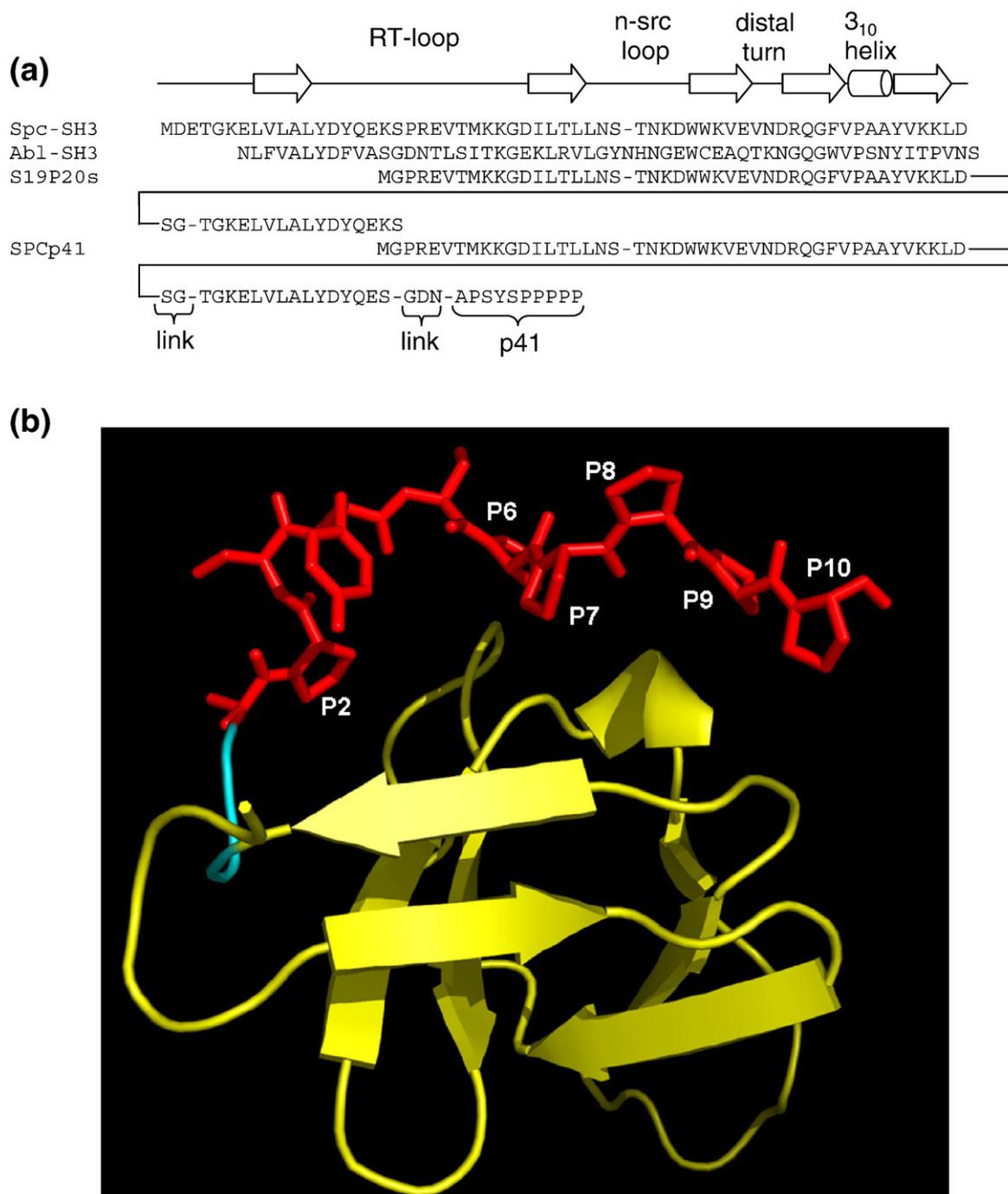


Figure 1. (a) Sequence alignment between the Spc-SH3 and Abl-SH3 domains, the S19P20s circular permutant and the SPCp41 chimera. The secondary structure elements are depicted schematically at the top (b) The high-resolution structure of SPCp41 determined by NMR.²⁰ The SH3 domain is drawn in the form of a fluid sketch, whilst the p41 ligand segment is represented in sticks. The SH3 domain is yellow, the linker in light blue and the p41 ligand segment red. The different mutated proline residues are labelled according to their numbering in the p41 sequence.

165 (Figure 1(b)), confirming the tight binding between
 166 the p41 peptide segment and the SH3 part of the
 167 chimera. The SH3–p41 interactions found in SPCp41
 168 are practically identical to those found in
 169 the complexes of the p41 peptide with the Spc-SH3
 170 and the Abl-SH3 domains.^{15,21} It is thus possible to
 171 use this chimeric protein as a tool to analyse the
 172 energetic determinants of binding between SH3
 173 domains and proline-rich peptides by site-directed

mutagenesis and thermal-unfolding measurements.¹⁷⁴
 We tested this approach by evaluating the relative
 175 importance of each proline residue of p41 in its
 176 interaction with the SH3 binding domain. We
 177 produced six single mutants of SPCp41, within
 178 which each proline in the p41 sequence has been
 179 replaced by alanine. We analysed the thermody-
 180 namic stabilities of the mutant proteins by DSC and
 181 compared them with those of SPCp41 and S19P20s,
 182

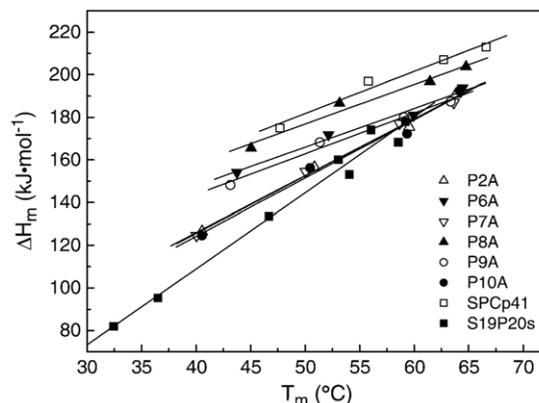
183 in conjunction with structural information from
 184 circular dichroism, fluorescence and NMR, as well
 185 as ^1H - ^{15}N NMR relaxation measurements. We inter-
 186 preted the data assuming a three-state equilibrium
 187 model, in which an intermediate state characterized
 188 by a loss of interactions at the SH3–p41 binding
 189 interface is significantly populated. This analysis has
 190 allowed us to separate the thermodynamic mag-
 191 nitudes of the SH3–p41 interaction from those of
 192 the SH3 unfolding. The results provide a quantita-
 193 tive evaluation of the energetic impact of each pro-
 194 line substitution in the p41 ligand and give insight
 195 into the thermodynamic determinants of binding
 196 between SH3 domains and proline-rich ligands,
 197 thus demonstrating the effectiveness of chimeric
 198 fusion proteins as tools to analysing protein-ligand
 199 interactions.

200 Results

201 Thermal unfolding of SPCp41 Pro-Ala mutants 202 by differential scanning calorimetry

203 We studied the thermal unfolding of each Pro-Ala
 204 mutant of SPCp41 by DSC at several pH values
 205 (pH 2.5, 3.0, 3.5 and 4.0) under the same buffer
 206 conditions as those used in previous experiments
 207 with SPCp41 and the circular permutant SH3
 208 domain S19P20s.¹⁷ Under these conditions the
 209 unfolding of all the mutant chimeras was comple-
 210 tely reversible. In contrast, at pH 7.0 in sodium
 211 phosphate buffer, the chimeric mutants were prone
 212 to aggregate and their thermal unfolding was
 213 irreversible, thus preventing our comparative ana-
 214 lysis at a near-physiological pH. Nevertheless, the
 215 binding of p41 to the Abl–SH3 and the Spc–SH3
 216 domains depends only slightly upon pH^{14,15} so the
 217 thermodynamic magnitudes measured at acid pH
 218 are fully representative of those typically observed
 219 for the binding between SH3 domains and their
 220 ligands.

221 In an initial analysis we fitted the heat-capacity
 222 curves, C_p , of all the proteins using the two-state
 223 unfolding model. All the fittings were very good
 224 (results not shown) as we had already found for the
 225 SPCp41 chimera.¹⁷ Figure 2 shows a plot of the
 226 apparent enthalpy changes of unfolding, ΔH_m ,
 227 versus the melting temperatures, T_m , derived from
 228 the two-state fittings for all mutants in comparison
 229 with the values obtained under the same conditions
 230 for SPCp41 and S19P20s.^{17,18} It is immediately
 231 evident from the T_m values that the Pro-Ala muta-
 232 tions in the p41 sequence decrease the stability of
 233 the chimeric proteins significantly, depending to a
 234 great extent upon the position of mutation. Muta-
 235 tion P8A, located on the solvent-exposed side of the
 236 PPII helix, results in the least decrease in stability,
 237 followed by mutants P6A and P9A, which show
 238 very similar values (Figure 2). Mutants P2A, P7A
 239 and P10A were much less stable to a similar degree
 240 at every pH value analysed. These results demon-



241 **Figure 2.** Plots of the enthalpy change of unfolding,
 242 ΔH_m , versus the unfolding temperature, T_m , arrived at by
 243 two-state analysis of the DSC data for S19P20s, SPCp41
 244 and all its Pro-Ala mutants studied at different pH values,
 245 as indicated in Results. The lines correspond to the linear
 246 regression of the values obtained for each protein.

247 strate that the stability of these chimeric proteins is
 248 very sensitive to changes in the interactions at the
 249 binding interface.

250 A closer look at Figure 2 reveals an interesting fact:
 251 assuming a two-state unfolding model and accord-
 252 ing to the usual interpretation of this kind of plot,
 253 the slope of the linear dependence of ΔH_m with T_m
 254 yields the heat-capacity change of unfolding, ΔC_p ,
 255 which is approximately constant within the tem-
 256 perature range of the T_m values. Surprisingly, the
 257 greatest ΔC_p value ($3.6(\pm 0.2)$ kJ mol⁻¹) is found for
 258 S19P20s in the absence of interactions at the binding
 259 site. All the chimeric proteins have a considerably
 260 lower ΔC_p of unfolding in spite of the presence of
 261 the SH3–p41 interaction in their structure, burying
 262 875 \AA^2 of mostly hydrophobic surface in SPCp41.²⁰
 263 This is apparently at odds with the well established
 264 correlation between the unfolding ΔC_p and the
 265 amount of surface area exposed to the solvent upon
 266 unfolding^{22,23} and also inconsistent with the nega-
 267 tive value of the change in the heat capacity of
 268 binding between p41 and Spc–SH3¹⁵ or Abl–SH3.¹⁴
 269 Therefore there must be additional contributions
 270 to the temperature dependence of the apparent
 271 unfolding enthalpies shown in Figure 2 that cannot
 272 be justified by the simple assumption of two-state
 273 unfolding.

268 Structural analysis of the mutant chimeras by 269 circular dichroism and fluorescence

270 We analysed the structural effects of each Pro-Ala
 271 mutation in the p41 sequence segment of the chi-
 272 meric proteins using both fluorescence and circular
 273 dichroism (CD) spectroscopy. The buffer conditions
 274 were pH 2.5, 3.0, 3.5 and 4.0, in 20 mM glycine or
 275 20 mM sodium acetate, identical to those used in the
 276 thermal stability experiments. Fluorescence spectra
 277 of each mutant protein were recorded at 25 °C
 278 between 305 nm and 400 nm with selective trypto-

phan excitation at 298 nm and compared with the analogous spectra of SPCp41 and S19P20s (Figure 3). All the chimeric proteins showed a considerable enhancement in fluorescence intensity and a blue shift of the maximum emission wavelength as compared to S19P20s. These effects are due to the interaction between the C-terminal peptide sequence and the SH3 binding site in the chimeras, which alters the chemical environment and solvent exposure of the Trp26 indole group at the binding site surface.^{15,20,24} Similar fluorescence emission enhancements have been observed when the p41 peptide binds to S19P20s,¹⁷ the wild-type (WT) and R21A mutant Spc-SH3 domain¹⁵ and the Abl-SH3 domain.¹⁹ The intensity of the emission is considerably higher for some mutant chimeric proteins, indicating a different degree of interaction between the C-terminal peptide sequence and the SH3 binding site, which runs in the approximate order: SPCp41 > P8A ≥ P9A ≈ P6A > P2A ≈ P7A ≈ P10A.

Far and near-UV CD spectra were also recorded for all the chimeric proteins at 25 °C under the same buffer conditions. When compared with S19P20s,

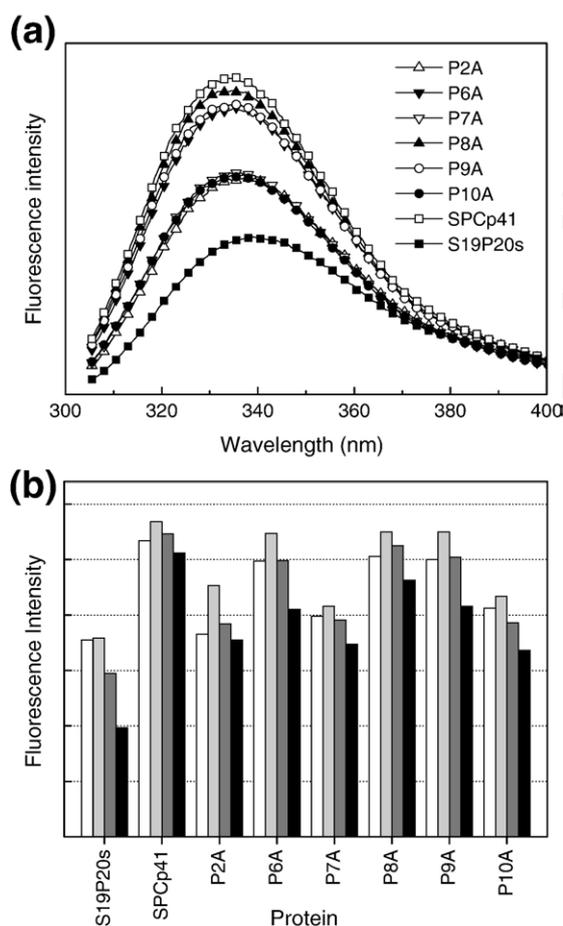


Figure 3. (a) Tryptophan fluorescence emission spectra of S19P20s, SPCp41 and the various Pro-Ala mutants at pH 3.0 and 25 °C. (b) Relative fluorescence intensity for each chimeric protein and S19P20s at 25 °C and different pH values: white, pH 4.0; light grey, pH 3.5; grey, pH 3.0; black, pH 2.5.

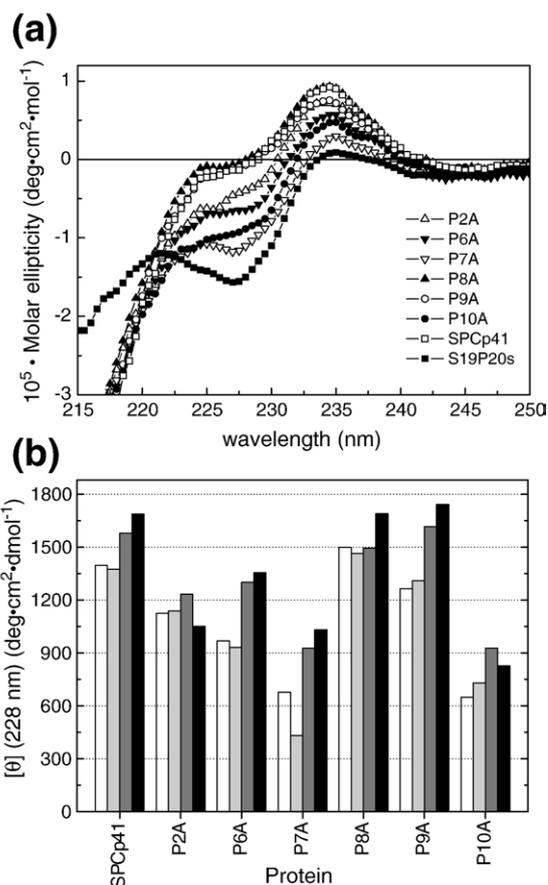


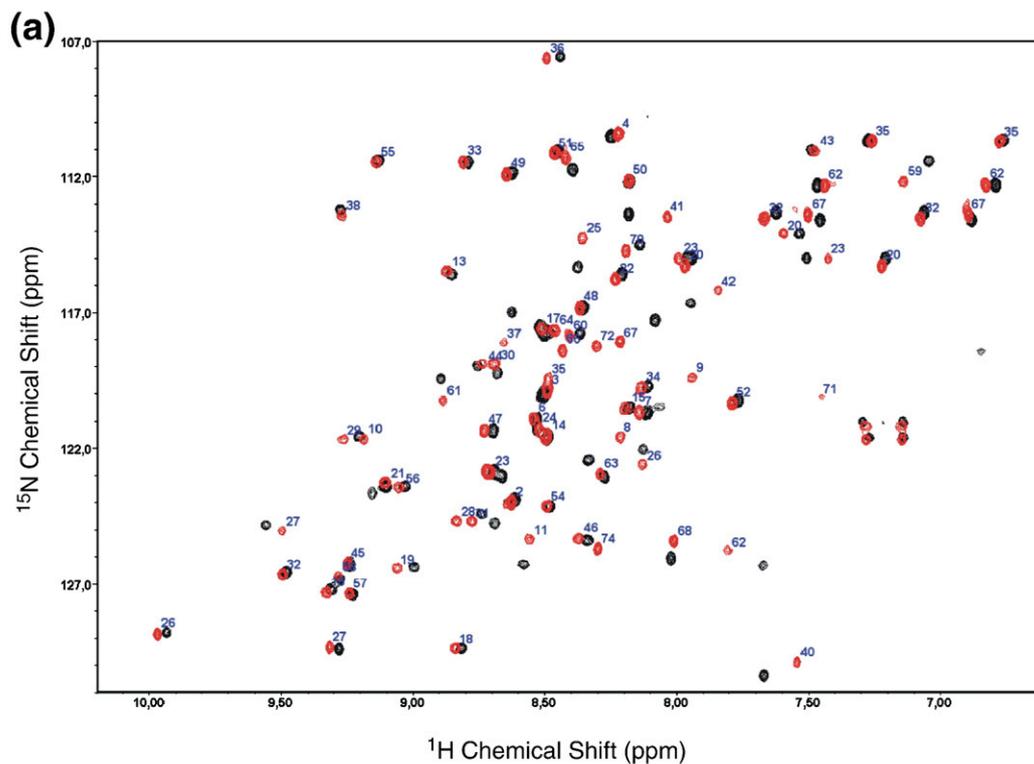
Figure 4. (a) Far-UV CD spectra of S19P20s, SPCp41 and all the Pro-Ala mutant chimeras at pH 3.5 and 25 °C. (b) Differences in CD ellipticity at 228 nm between each chimeric protein and S19P20s at 25 °C and different pH values. Colours of bars are as for Figure 2.

the far-UV CD spectra of SPCp41 and the mutant chimeras show a considerable increase in a positive ellipticity band centred at around 228 nm together with an enhancement in a negative ellipticity band at ≈ 205 nm (Figure 4; see also Figure 2 of Martin-Sierra *et al.*¹⁷). The difference spectra obtained by subtracting the S19P20s spectrum from those of the chimeric proteins shows a positive band at ≈ 228 nm, characteristic of PPII helix conformation,²⁵ indicating that the ligand sequence adopts a similar conformation in the chimeric proteins to that typically observed for proline-rich peptides bound to SH3 domains. Alterations in the environment of aromatic side-chains have also been reported as contributing to the positive band at about 230 nm observed for some proteins,²⁶ as is the case with the SH3 domains. The differences in ellipticity at 228 nm between each chimeric protein and S19P20s (Figure 4(b)) agree well with the changes in tryptophan fluorescence for each mutant chimera.

The near-UV CD spectra of SPCp41 and all Pro-Ala mutants show an increased positive ellipticity band centred at ≈ 278 nm in comparison with that of S19P20s under the same conditions (results not

326 shown). The near-UV CD changes are consistent
 327 with the immobilisation of aromatic side-chains
 328 in the SH3 binding site by interactions with the
 329 C-terminal peptide chain. Mutating the different
 330 proline residues of the p41 sequence produced con-

331 siderable variations both in the intensity and shape
 332 of the ellipticity band, thus suggesting significant
 333 differences in the degree of interaction at the binding
 334 interface, which agrees with the far-UV CD and
 335 fluorescence results.



(b)

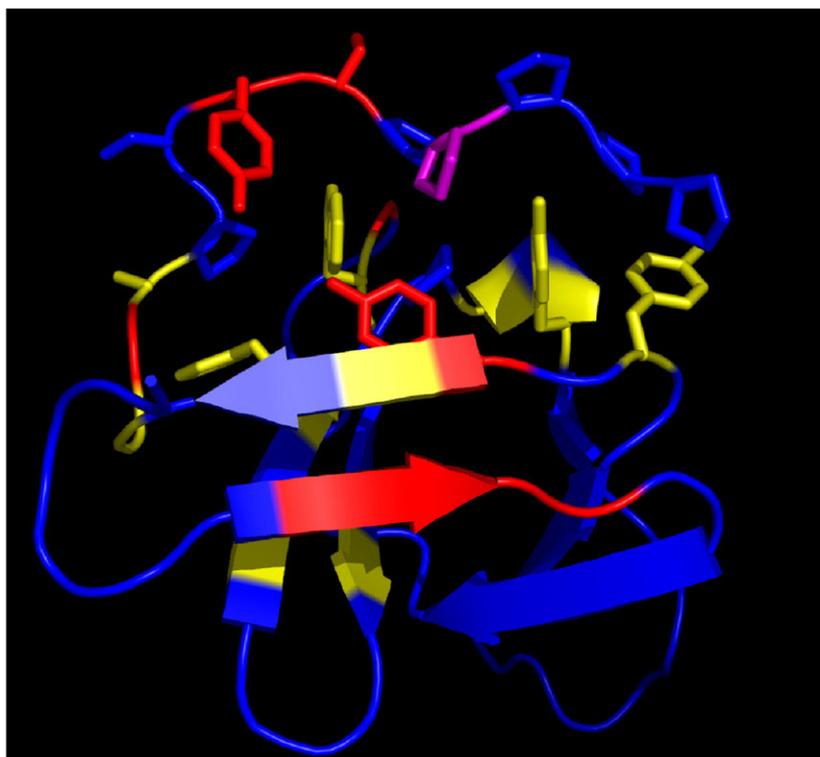


Figure 5 (legend on next page)

336 Comparative NMR analysis of SPCp41 and the 337 P7A mutant

338 The fluorescence and CD data indicate consider- 387
339 able structural changes at the binding interface 388
340 depending upon the proline mutation site. To obtain 389
341 further details about the structural changes at the 390
342 binding interface produced by one of the most 391
343 destabilizing mutations we analysed ¹⁵N-enriched 392
344 samples of SPCp41 and P7A by NMR. The backbone 393
345 amide ¹H and ¹⁵N frequencies of the P7A mutant
346 were assigned by direct comparison of the 3D ¹⁵N-
347 nuclear Overhauser effect spectroscopy (NOESY)-
348 heteronuclear single-quantum coherence (HSQC)
349 spectrum of P7A with that of SPCp41, which had
350 already been assigned.²⁰ Figure 5 shows an overlay
351 of the assigned ¹H-¹⁵N-HSQC spectra and a
352 representation of the differences in the chemical
353 shifts in the three-dimensional structure of SPCp41.
354 Many of the backbone ¹H-¹⁵N chemical shifts were
355 affected by the P7A mutation, especially those of
356 residues in the vicinity of the SH3–p41 interface, but
357 also in regions further away along the SH3 domain.
358 These changes probably reflect a lower degree of
359 interaction between the p41 segment and the SH3
360 binding site in the P7A mutant.

361 To search for more direct structural evidence in
362 support of a weakening of the interaction between
363 the SH3 and p41 parts of the protein produced by
364 the P7A mutation we compared the NOEs involving
365 the H^{ε1} (indole) protons of the two tryptophan
366 residues in SPCp41 with those in the P7A mutant.
367 These two tryptophan residues are highly preserved
368 among SH3 domains. Trp26 interacts directly with
369 the p41 peptide, not only in SPCp41 but also in the
370 complex between the separate partners.^{15,20} Trp27 is
371 known to be important for the stability of the SH3
372 domain and has many contacts with neighbouring
373 residues in the domain's core. The NOEs detected in
374 the 3D ¹⁵N-NOESY-HSQC spectra for the H^{ε1}
375 protons of both Trp residues in SPCp41 and P7A
376 are shown in Figure 6. All the NOEs, except one with
377 very low intensity, can be observed for Trp27 H^{ε1} in
378 both proteins. This indicates that the P7A mutation
379 in the p41 sequence does not cause any significant
380 change in the structure of the SH3 domain. Most of
381 the NOEs involving Trp26 H^{ε1}, on the other hand,
382 disappeared upon mutation, indicating that the
383 average distance between the Trp26 H^{ε1} proton
384 and the p41 segment increased substantially upon
385 mutation, because NOE intensity has an inversed
386 sixth power relationship with the proton–proton

distance. These observations provide direct evi- 387
dence that interactions at the binding interface are 388
substantially weakened by the P7A mutation whilst 389
the SH3 domain structure remains largely unaltered, 390
which is consistent with the similarity between the 391
structures of the Spc–SH3 domain in its free and 392
p41-complexed forms.¹⁵ 393

394 Backbone dynamics from ¹⁵N relaxation data: 395 increased flexibility at the binding interface 396 upon mutation 397

¹⁵N NMR relaxation data are sensitive probes for 397
both the molecular rotational diffusion and the 398
local backbone dynamics of a protein. Thus, to 399
analyse the effect of the P7A mutation upon the 400
conformational flexibility of the chimeric protein 401
we measured the relaxation parameters (¹⁵N *R*₁, *R*₂ 402
and ¹H-¹⁵N steady-state NOE) at 600 MHz and 403
25 °C for SPCp41 and the P7A mutant. The NOE 404
values in the SH3 domain region and the SH3– 405
peptide linker are almost identical for both pro- 406
teins, high NOE values (about 0.8) being observed 407
for most of the SH3 domain chain (Figure 7(a)). 408
Reduced NOE values, indicating extensive high- 409
frequency motions (ps to ns), are found for residues 410
2 to 7 at the N terminus, for residues 46 to 53, 411
corresponding to the link connecting the original N 412
and C termini in the S19P20s permutant and for 413
residues 65 to 67 in the linker connecting the 414
S19P20s and ligand segments. Conversely, the NOE 415
values observed in the ligand segment (residues 68, 416
70, 71 and 72 in SPCp41 plus residue 74 in P7A) are 417
substantially lower for the P7A mutant than for 418
SPCp41, reflecting a higher disorder due to an 419
enhancement in high-frequency motions produced 420
by the P7A mutation.²⁷ 421

The ¹⁵N *R*₁ and *R*₂ relaxation rates were also 422
measured for SPCp41 and P7A. Unfortunately, the 423
P7A mutant protein slowly aggregated during the 424
time of the measurements, which is reflected in a 425
general increase in *R*₂ values. This complicates our 426
comparative analysis and demands caution when 427
interpreting the results pertaining to relaxation 428
data.²⁸ Nevertheless, local differences in dynamics 429
between SPCp41 and P7A may also be revealed 430
directly and concisely by analysing the *R*₂/*R*₁ ratios 431
(Figure 7(b)). An analysis of the relaxation data of 432
33 residues in the most ordered regions of SPCp41 433
using TENSOR2²⁹ indicates average ¹⁵N *R*₂/*R*₁ ra- 434
tios corresponding to an apparent rotational corre- 435
lation time, τ_{mv}, of 5.15(± 0.04) ns, which is within the 436

Figure 5. Chemical-shift changes produced by the mutation of proline 7 to alanine in the chimeric protein. (a) Overlay of 600 MHz ¹H-¹⁵N HSQC spectra of SPCp41 (black) and P7A (red). Peaks of the mutant protein are labelled by their corresponding residue number. (b) Representative, lowest-energy structure of SPCp41 (PDB code 2JMC) coloured according to the magnitude of the chemical-shift changes (Δδ) upon mutation, calculated using the equation:

$$\Delta\delta_{ppm} = \sqrt{(\Delta\delta_{HN})^2 + \left(\frac{\Delta\delta_N}{6.51}\right)^2}$$

Colour code: blue, Δδ ≤ 0.25 ppm; yellow, 0.25 ppm > Δδ ≤ 0.75 ppm; and red, Δδ > 0.75 ppm. The side-chains of the p41 segment and the aromatic residues of the SH3 binding-site surface are drawn as sticks. The mutated proline (P7) is magenta.

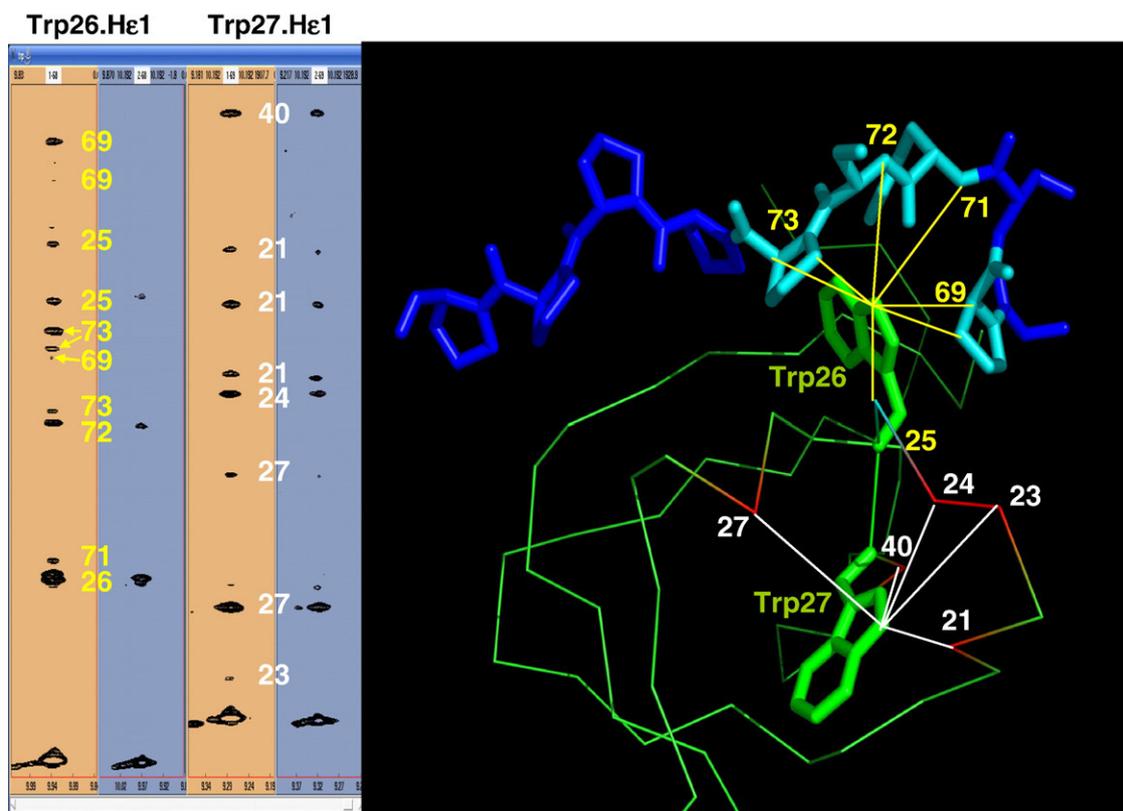


Figure 6. Differences in NOEs involving the indole H ϵ^1 protons of the two tryptophan residues (Trp26 and Trp27) between SPCp41 and P7A. The left-hand panel shows strips taken at the $^1\text{H}^{\epsilon^1}$ and $^{15}\text{N}^{\epsilon^1}$ frequencies of Trp26 and Trp27, as indicated, from the 3D ^{15}N -HSQC-NOESY spectra of the SPCp41 chimera (brownish background) and the P7A mutant (blue background). Each NOE is labelled with the number of the residue to which the NOE was assigned. The right-hand panel shows these NOEs represented as continuous lines in the structure of SPCp41. NOEs involving Trp26–H ϵ^1 are shown in yellow and those involving Trp27–H ϵ^1 in white. The p41 segment and the Trp side-chains are represented as thick sticks and the rest of the chimera is drawn in thin lines.

437 expected range for a monomeric 8 kDa protein. 438 Moreover, using Fast-HYDRONMR³⁰ we calculated 439 a rotational correlation time of 5.1 ns for the lowest- 440 energy SPCp41 structure. Locally decreased R_2/R_1 441 ratios can be seen in SPCp41 at regions largely co- 442 inciding with those showing reduced NOE values, 443 confirming the presence in these regions of fast 444 local motions on the ps to ns timescale. Conversely, 445 increased R_2/R_1 ratios, such as those visible in 446 SPCp41 for residues 68, 71 and 72 in the p41 447 sequence and the neighbouring residues 9, 10, 62 448 and 63 in the RT-loop region of the SH3 domain, 449 indicate slow motions and/or conformational 450 exchange in the μs – ms timescale, which cause 451 higher R_2 values.²⁷ These effects suggest a certain 452 degree of slow conformational exchange at the 453 binding interface in the SPCp41 chimera. Increased 454 R_2/R_1 ratios are also observed for residue 17 and 455 particularly residue 33 at the distal turn, far away 456 from the binding interface.

457 An overall increase in R_2 produced by aggregation 458 is clearly visible in the sequential plot of the R_2/R_1 459 ratios of P7A for most residues in the SH3 domain, 460 except for those undergoing high-frequency motions 461 that compensate the effect of aggregation. Residue 462 33 still has a high R_2/R_1 ratio in the P7A mutant,

463 indicating a similar conformational exchange phe- 464 nomenon at the distal turn. A considerable reduc- 465 tion in the R_2/R_1 ratio is observed for the linker and 466 the p41 part of the P7A chimera, which implies that 467 relaxation in this part of the mutant protein becomes 468 governed by high-frequency motions. Accordingly, 469 the P7A mutation shifts the time-scale of motions of 470 the p41 ligand chain from slow conformational 471 fluctuations in the SPCp41 chimera to fast motions 472 in the P7A mutant.

473 Global three-state analysis of the heat-capacity 474 curves of unfolding

475 The results of the structural and dynamic char- 476 acterization of the mutant chimeras shown above 477 indicate that the different Pro-Ala mutations along 478 the p41 sequence of SPCp41 weaken the interactions 479 at the binding interface and increase the high- 480 frequency motions of the peptide segment. This is 481 accompanied by increased exposure of the aromatic 482 side-chains to the solvent and a reduction in the 483 fraction of PII helix for the p41 segment, as shown 484 by the fluorescence and CD results. These effects are 485 greatest for P2A, P7A and P10A, moderate for P6A 486 and P9A and lowest for P8A, and correlate quite

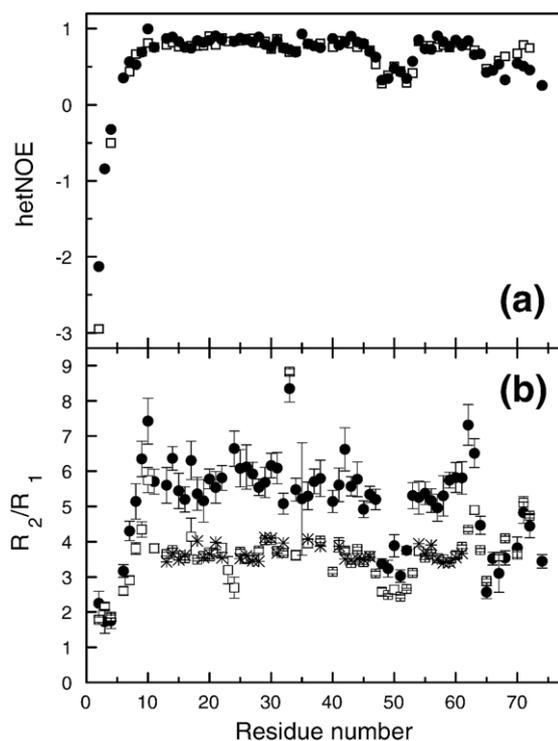


Figure 7. Backbone dynamics of the SPCp41 and P7A chimeras at pH 3.5 and 25 °C measured at 600 MHz. (a) Heteronuclear ^1H - ^{15}N NOEs for SPCp41 (open squares) and the P7A mutant (filled black circles). Errors were estimated to be around 2% for the WT and 4% for the mutant protein. (b) Ratio between the R_2 and R_1 ^{15}N relaxation rates for SPCp41 (open squares) and the P7A mutant (filled black circles). Errors were estimated as described in Materials and Methods. Back-calculated R_2/R_1 values using Fast-HYDRONMR for the 33 residues in SPCp41 used in the TENSOR2 analysis are shown as asterisks.

well with the apparent changes in stability measured by DSC. As mentioned above, the apparent heat capacity changes of unfolding of these proteins obtained *via* a two-state analysis vary in an opposite order and are inconsistent with the structural effects of the mutations.

Although the heat-capacity changes of protein unfolding and binding have been frequently interpreted in terms of protein hydration effects,^{22,23} other sources for changes in heat capacity have also been proposed, even though their importance and sign remain uncertain.³¹ The effects of additional equilibria linked to folding and binding processes can be of great importance in contributing to the heat-capacity changes.^{32,33}

In the case of these chimeric proteins it is conceivable that the weakening of interactions produced by the Pro-Ala mutations may promote local conformational fluctuations at the binding interface, which would undergo a dynamic equilibrium between “bound” and “unbound” states. Mutations alter the position of this equilibrium and consequently change the average degree of disorder

at the binding interface. If these conformational states had different enthalpy and heat capacity (see below) their equilibrium would be displaced toward the “unbound” state by the increase in temperature, thus changing their relative populations. The heat involved in this shift would contribute significantly to the heat-capacity change of the overall unfolding of the chimeric protein, thus resulting in an anomalous value when determined by two-state analysis. This is important because the derivation of thermodynamic magnitudes of unfolding of mutant proteins using simple two-state analyses may disregard the influence of conformational flexibility and lead to incorrect estimates of the thermodynamic magnitudes. This scenario can be simplified in the equilibrium model:



where N and I are states in which the SH3–peptide binding interface is in a “bound” and “unbound” conformation respectively, and U is the fully unfolded state of the protein. K_1 and K_2 are the equilibrium constants of each transition. Using this three-state equilibrium model we carried out the global fitting of the C_p curves measured at different pH values for each of the chimeric proteins. The equations of this model and the details of the fitting procedure are provided as Supplementary Data.

In our analysis we assumed that the 12 residue C-terminal peptide is completely disordered and solvent exposed in the unbound state, I, and does not influence the thermodynamics of unfolding of the SH3 domain. We therefore assumed that the thermodynamic magnitudes of the transition between the I and U states for all the chimeric proteins are the same as those of the two-state unfolding of S19P20s at each pH value, thus allowing us to incorporate them into the global fittings as fixed parameters. The C_p functions of the I and U states were also taken from the S19P20s DSC analysis and corrected for the contribution of the disordered 12 residue C-terminal segment according to Makhadze & Privalov.³⁴ Under these assumptions, the adjustable parameters for the fitting of the DSC curves of each chimeric protein at different pH values are the equilibrium constant, K_1 , the enthalpy change, ΔH_1 , and the heat capacity change, ΔC_{p1} , for the transition between the states N and I at any particular reference temperature, plus the temperature dependence of C_p for the state N, which is taken to be a linear function.

For simplicity’s sake we assumed that the enthalpy change, ΔH_1 , is independent of pH within the range studied. This is a good approximation considering the absence of charged residues at the C-terminal sequence and is supported by the very slight dependence upon pH of the enthalpy of binding of the p41 decapeptide to the Abl–SH3 domain¹⁴ and the Spc–SH3 domain.¹⁵

Figure 8 shows the global fittings for all the chimeric variants and Table 1 summarises the

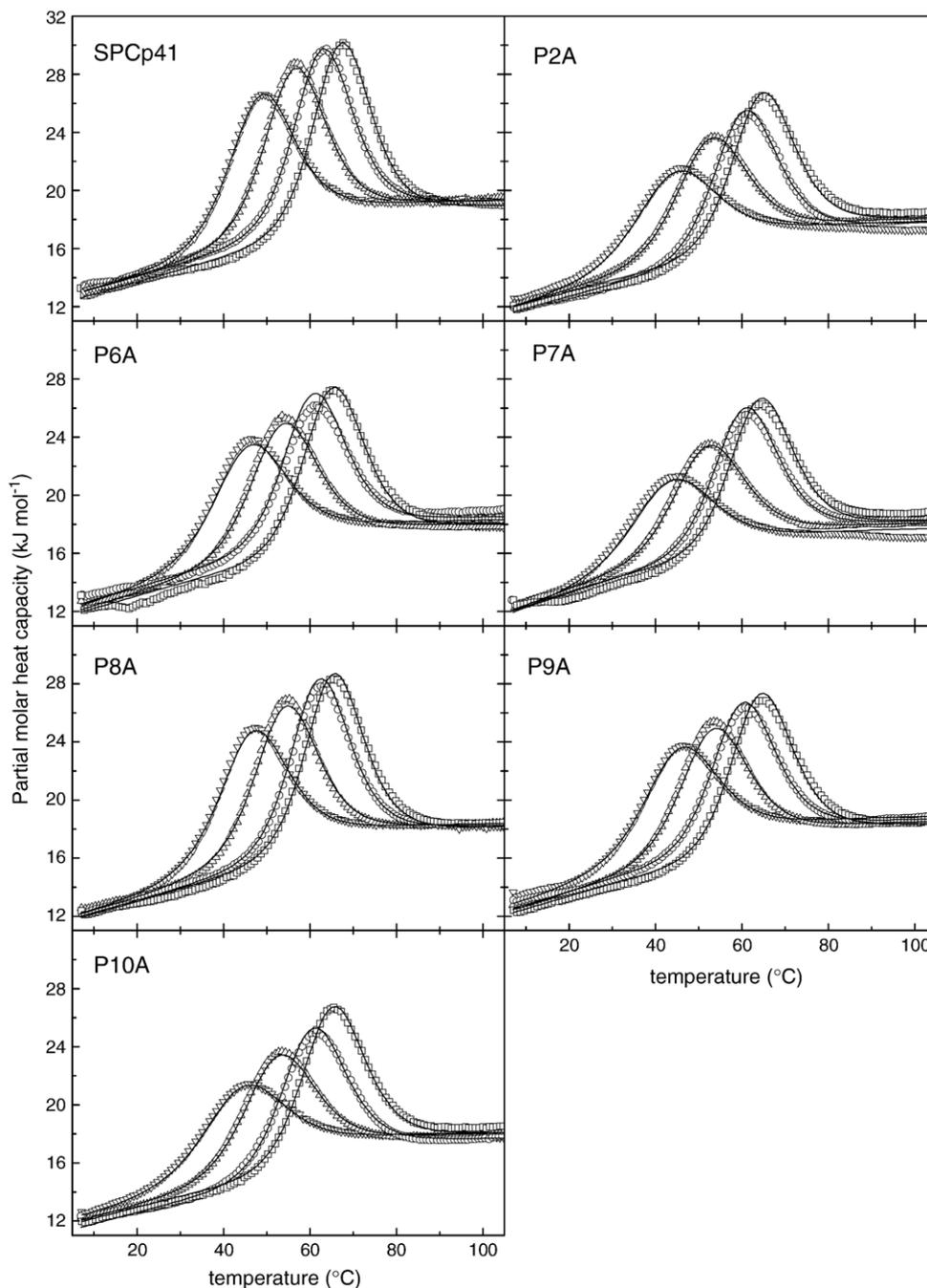


Figure 8. Three-state analysis of the DSC thermograms of the chimeric proteins at various pH values. Symbols represent the experimental DSC thermograms at different pH values: upside-down triangles, pH 2.5; triangles, pH 3.0; circles, pH 3.5; and squares, pH 4.0. Lines correspond to the best fitting using the equations of the three-state model, as described in the text.

570 thermodynamic magnitudes resulting from this
 571 analysis. In spite of the constraints imposed in the
 572 analysis, the C_p curves are very well reproduced
 573 both in shape and pH dependence, which supports
 574 the validity of the assumptions described above.
 575 More importantly, this analysis allows us to distin-
 576 guish between the thermodynamic magnitudes of
 577 the interaction at the binding interface and those
 578 of the unfolding of the SH3 domain.

579 These results indicate that the apparent structur-
 580 al properties of the different Pro-Ala mutants, as

observed by CD, fluorescence and NMR, are
 representative average properties in a dynamic
 conformational equilibrium between the bound
 and unbound conformations, which is character-
 ized by the equilibrium constant, K_1 , for each chimeric
 protein at each temperature. The thermodynamic
 magnitudes given in Table 1 rank each mutated
 proline residue according to its relative importance
 in the interaction. Furthermore, the SH3-peptide
 interaction is always thermodynamically favour-
 able although in general, and especially for some

t1.1 **Table 1.** Thermodynamic magnitudes at 25 °C for the SH3–peptide interaction in the chimeric proteins deriving from a
 t1.2 three-state analysis of the DSC curves

t1.3	Protein	pH	K_1	ΔG_1 (kJ mol ⁻¹)	ΔH_1 (kJ mol ⁻¹)	$T \cdot \Delta S_1$ (kJ mol ⁻¹)	ΔC_{p1} (kJ mol ⁻¹ K ⁻¹)
t1.4	SPCp41	4.0	0.0234±0.0008	9.31±0.08	56.6±0.9	47.3±1.0	0.242±0.011
t1.5		3.5	0.0288±0.0010	8.80±0.09			
t1.6		3.0	0.0389±0.0013	8.05±0.08			
t1.7		2.5	0.0541±0.0017	7.22±0.08			
t1.8	P2A	4.0	0.163±0.011	4.49±0.17	37.8±1.4	33.3±1.6	0.02±0.04
t1.9		3.5	0.171±0.011	4.37±0.16			
t1.10		3.0	0.35±0.03	2.60±0.21			
t1.11		2.5	0.51±0.04	1.69±0.20			
t1.12	P6A	4.0	0.079±0.006	6.28±0.19	49.0±1.8	42.8±2.0	-0.01±0.05
t1.13		3.5	0.098±0.008	5.75±0.20			
t1.14		3.0	0.142±0.011	4.83±0.19			
t1.15		2.5	0.170±0.012	4.39±0.17			
t1.16	P7A	4.0	0.153±0.012	4.65±0.19	47.2±1.6	42.5±1.8	-0.12±0.04
t1.17		3.5	0.129±0.012	5.08±0.21			
t1.18		3.0	0.40±0.04	2.27±0.25			
t1.19		2.5	0.54±0.05	1.52±0.23			
t1.20	P8A	4.0	0.0449±0.0021	7.69±0.12	56.0±1.2	48.3±1.3	0.19±0.03
t1.21		3.5	0.0384±0.0018	8.08±0.12			
t1.22		3.0	0.076±0.003	6.39±0.10			
t1.23		2.5	0.096±0.004	5.80±0.10			
t1.24	P9A	4.0	0.117±0.006	5.31±0.13	48.4±1.1	43.1±1.2	-0.06±0.03
t1.25		3.5	0.131±0.007	5.04±0.13			
t1.26		3.0	0.179±0.009	4.26±0.12			
t1.27		2.5	0.200±0.009	3.99±0.11			
t1.28	P10A	4.0	0.181±0.009	4.24±0.12	28.2±1.2	24.0±1.3	0.014±0.018
t1.29		3.5	0.228±0.012	3.66±0.13			
t1.30		3.0	0.49±0.03	1.78±0.15			
t1.31		2.5	0.73±0.05	0.79±0.17			

t1.32 Errors in the parameters correspond to 95% confidence intervals in the fittings.

592 mutants (P2A, P7A and P10A), and at low pH
 593 values, the Gibbs energy of the interaction is low.

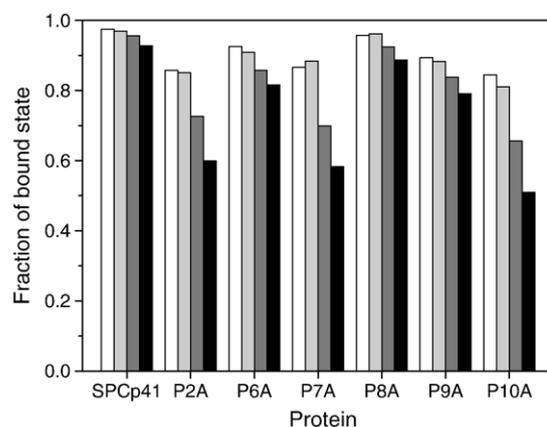
594 The values of ΔH_1 and $T \cdot \Delta S_1$ are positive and the
 595 interaction is dominated by a large, favourable
 596 enthalpy, partially balanced out by an unfavourable
 597 entropic term. This is usually observed in the bind-
 598 ing of proline-rich peptides to SH3 domains.^{8,11–14}

599 Except for the P8A mutant, the Pro-Ala muta-
 600 tions produce a considerable decrease in the en-
 601 thalpy of the interaction, which is particularly
 602 pronounced in the P2A and P10A mutants. This
 603 is accompanied by a concomitant reduction in
 604 the entropic terms, resulting in strong enthalpy-
 605 entropy compensation. The ΔC_{p1} values for the
 606 SPCp41 and P8A chimeric proteins are relatively
 607 small and close to zero for the rest of the chimeras.
 608 The magnitudes of the ΔC_{p1} values are smaller
 609 than those reported for the heat-capacity change of
 610 binding of p41 to either Abl-SH3 (-0.73 kJ K⁻¹
 611 mol⁻¹)¹⁴ or Spc-SH3 (-1.3 kJ K⁻¹ mol⁻¹),¹⁵ suggest-
 612 ing a higher degree of hydration of the binding
 613 interface in the bound state of all the chimeras than
 614 in either of these complexes.

615 The thermodynamic magnitudes given in Table 1
 616 allowed us to estimate the populations of each
 617 state of the chimeric proteins as functions of
 618 temperature. The populations of the bound state
 619 of each chimeric protein at 25 °C are set out in
 620 Figure 9. These populations compare quite well
 621 with the relative variations in the intrinsic Trp
 622 fluorescence and far-UV CD signal shown in
 623 Figures 3 and 4.

Binding thermodynamics by isothermal titration calorimetry 624

625
 626 To validate our approach towards determining the
 627 energetics of the p41-based peptides binding to the
 628 SH3 domain we measured directly the thermody-
 629 namic magnitudes of binding of the isolated p41
 630 peptide to the S19P20s domain by ITC at pH 3.5 and
 631 25 °C. Previous titration experiments by tryptophan
 632 fluorescence had estimated a binding constant, K_b ,
 633 equal to $6.3(\pm 1.3) \times 10^3$ M⁻¹, which is equivalent to a



624 **Figure 9.** Fraction of intra-molecularly “bound” state
 625 for each chimeric protein at 25 °C and different pH values
 626 calculated from the thermodynamic magnitudes of Table
 627 1. Bar colours are as for Figure 2.

634 K_d of $160(\pm 34) \mu\text{M}$.¹⁷ Since this affinity is very low, for an accurate determination of the binding thermodynamics we needed to design the ITC experiment to have a c value (the product of the binding constant and the concentration of protein in the ITC cell) within the appropriate range.³⁵ Thus, we used 0.44 mM of protein in the cell and 10.8 mM of p41 peptide in the syringe, giving a c value of around 3. Figure 10 shows the result of the ITC experiment and Table 2 summarises the thermodynamic magnitudes of the interaction. The binding affinity measured by ITC ($K_d=262(\pm 12) \mu\text{M}$) is slightly lower than that determined by fluorescence and confirms the low affinity of the p41 sequence for this permuted SH3 domain. The large and negative binding enthalpy is partially balanced out by an unfavourable binding entropy, as is usually found in

Table 2. Thermodynamic parameters of binding between isolated decapeptides and S19P20s determined by ITC at pH 3.5 and 25 °C

Peptide	K_b (M^{-1})	ΔH_b (kJ mol^{-1})	n^b	ΔG_b (kJ mol^{-1})
APSYSPPPPP (p41)	3800 ± 180	-58.1 ± 1.6	1.09 ± 0.02	-20.44 ± 0.12
APSYSPAPPP (P7Ap) ^a	480 ± 100	-46 ± 13	–	-15.3 ± 0.5

Errors correspond to 95% confidence intervals in the determination of the parameters.

^a Analysed by displacement ITC.

^b The number of independent sites was fixed in the analysis of the displacement experiment.

this type of interaction. The standard Gibbs energy of binding at pH 3.5 and 25 °C between S19P20s and the p41 peptide is $-20.44(\pm 0.12) \text{ kJ mol}^{-1}$. These thermodynamic magnitudes are similar to those measured previously for the binding of p41 to the WT Spc-SH3 domain and its R21A mutant under the same conditions,¹⁵ indicating that the sequence permutation in the SH3 domain does not have a significant effect upon this interaction. By comparing the experimental Gibbs energies of intra- and intermolecular binding (Tables 1 and 2) we estimated the effect of tethering the ligand and the SH3 domain with a short linker upon the binding energy, which amounts to 11.1 kJ mol^{-1} . As discussed previously,¹⁷ this apparently unfavourable energy contribution resulting from the tethering of the binding partners is just an artefact arising from the unrealistic standard conditions (1 M of reactants) used to quantify the binding Gibbs energy. The SH3–peptide linker is common to all chimeric mutant proteins and appears largely disordered in the solution structure of SPCp41, without contacts with the rest of the structure.²⁰ The flexibility of the linker is also reflected in the NMR relaxation measurements (see above). If we assume, therefore, that the contribution of the tethering to the energy of the interaction is the same for all chimeric proteins, the dissociation constants could be estimated for the corresponding isolated Pro-Ala mutant peptides using the Gibbs energy values given in Table 1. For example, K_d would be $430 \mu\text{M}$ for P8A, 1.5 mM for P7A and 2.6 mM for P10A at pH 3.5 and 25 °C. These affinities are very low and make it extremely difficult to investigate these weak interactions by direct titration using the separate binding partners. Nevertheless, we carried out displacement ITC experiments³⁶ to estimate the thermodynamic magnitudes of the binding of the P7A mutant of the isolated p41 peptide (P7Ap) to the S19P20s domain at pH 3.5 and 25 °C. In these experiments a solution of S19P20s in the presence of P7Ap was titrated with peptide p41. The details and conditions of these experiments are described in Materials and Methods. The thermodynamic magnitudes resulting from this analysis appear in Table 2. With this experimental design the dissociation constant for the P7Ap peptide was determined as $2.1(\pm 0.4) \text{ mM}$, which

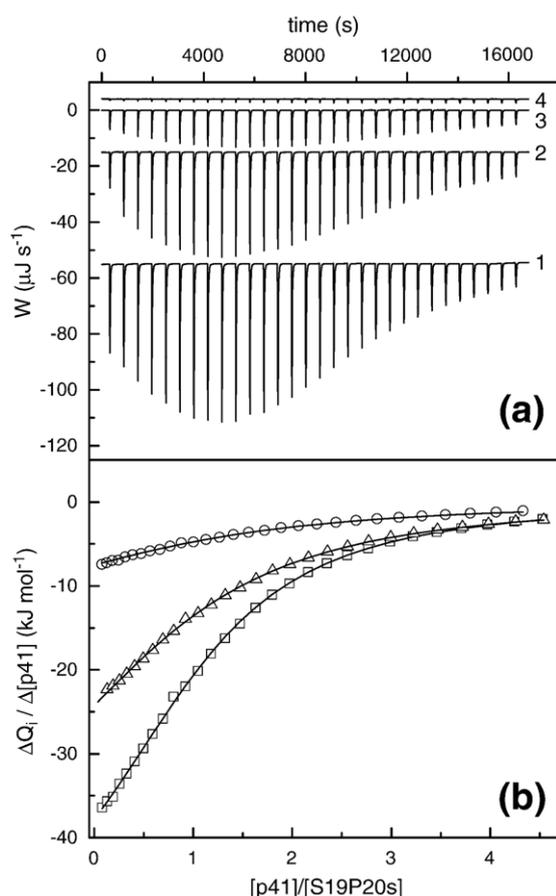


Figure 10. Calorimetric titrations of the S19P20s domain with the p41 decapeptide at 25 °C in 20 mM glycine (pH 3.5), both in the absence and presence of the P7Ap decapeptide. (a) Heat effects associated with the injection of variable volumes of a 10.83 mM p41 solution into the calorimeter cell containing: (1) 0.44 mM S19P20s; (2) 0.44 mM S19P20s + 1.32 mM P7Ap; (3) 0.46 mM S19P20s + 5.17 mM P7Ap; (4) buffer. (b) Normalised and corrected binding isotherms obtained from the thermograms shown in (a). Symbols represent the experimental data: squares, no P7Ap; triangles, 1.32 mM P7Ap; circles, 5.17 mM P7Ap. The continuous lines correspond to the best fitting, as described in Materials and Methods.

698 compares well with the results of the DSC analysis
699 of the mutant chimeric protein P7A (see above). The
700 determination of the binding enthalpy is somewhat
701 less accurate but also agrees satisfactorily with the
702 DSC results.

703 Discussion

704 Chimeric fusion proteins as suitable tools to 705 investigate the thermodynamics of weak 706 interactions

707 One of the biggest problems involved in analysing
708 the thermodynamics of weak SH3-mediated inter-
709 actions is that their affinities usually fall at the lower
710 edge of the range accessible to direct study by ITC.³⁵
711 Accurate determination of the thermodynamic mag-
712 nitudes requires large quantities of highly concen-
713 trated proteins and/or synthetic peptides, which
714 renders these studies impractical due both to the cost
715 and to possible collateral problems such as limited
716 solubility and/or aggregation. DSC transitions of
717 proteins have proved to be very sensitive to ligand
718 binding and so DSC has often been used in the
719 past to analyse the thermodynamics of protein-
720 ligand interaction in preference to ITC under some
721 circumstances.^{37–40} We have shown here and
722 elsewhere¹⁷ that intramolecular binding of low-
723 affinity sequences has a considerable impact upon
724 the stability and dynamic properties of an SH3
725 domain. Thus, since the thermodynamics of the
726 SH3–peptide interaction are affected considerably
727 by simple mutations, these effects (not only the
728 changes in affinity but also in the enthalpy–entropy
729 balance) can be separated from the unfolding
730 thermodynamics of the SH3 domain by a three-
731 state analysis of the DSC data of the chimeric
732 proteins. Our approach is applied here to an artificial
733 chimeric protein SPCp41 that mimics perfectly the
734 interactions typically found between SH3 domains
735 and proline-rich peptides. In fact, the high similarity
736 between the interactions at the SH3–p41 binding
737 interface within SPCp41 and those found in inter-
738 molecular complexes²⁰ is clearly reflected in parallel
739 thermodynamic magnitudes of binding, as de-
740 scribed in Results. Furthermore, we have demon-
741 strated that the three-residue link tethering the SH3
742 domain and the ligand is quite flexible and does not
743 influence the thermodynamics of binding, except
744 for the introduction of an energy offset imposed by
745 the reduction in configurational entropy, common to
746 all chimeric variants. This allowed us to evaluate
747 directly the changes in the thermodynamic mag-
748 nitudes of binding elicited by single-point mutations.

749 To our knowledge this thermodynamic approach
750 is completely novel although similar artificial fusion
751 proteins have been used by other researchers to
752 analyse structurally the interactions between SH3
753 domains and proline-rich peptides.^{41,42} Examples of
754 similar intramolecular interactions between SH3
755 and proline-rich peptides tethered by short links

are also found in nature, as observed for the Itk–SH3
756 domain.⁵ Although the particular topology of these
757 protein constructs is different to that of our system,
758 our thermodynamic approach is conceptually inde-
759 pendent of the nature of the interacting molecules
760 and, therefore, it may be fully applicable to these
761 intramolecular SH3–ligand interactions. Moreover,
762 this method may also be extended to other bio-
763 molecular interactions when the binding partners
764 can be conveniently tethered by protein engineering
765 or chemical methods. 766

The importance of proline residues in SH3–p41 767 binding 768

An analysis of the site-directed mutagenesis of
769 the chimeric proteins described here establishes very
770 clearly the importance of each proline residue of p41
771 in the SH3–peptide interaction. Mutation P8A has
772 the lowest impact on the interaction, which con-
773 forms with the solvent-exposed orientation of this
774 residue in the SPCp41 structure and the lack of
775 contact between the proline side-chain and the SH3
776 domain.²⁰ Substitutions of prolines 2, 7 and 10 have,
777 on the other hand, the highest impact on binding.
778 These three proline residues were found to be
779 essential for the binding of proline-rich peptides
780 to the Abl–SH3 domain⁴³ and proline residues 7 and
781 10 occupy preserved positions in the consensus
782 sequence of SH3 ligands $\Phi P x \Phi P$.² Proline residues 6
783 and 9 are intermediate in the order of importance,
784 occupying a less-preserved position in the consen-
785 sus sequence. These results agree very well with
786 Lim’s model for SH3–peptide interactions^{9,10,44} with
787 residues 6 and 9 in an “internal packing” position,
788 according to the terminology of the model, and
789 residues 7 and 10 in an “external packing” situation,
790 which is optimum for a proline residue to establish
791 strong hydrophobic interactions with the SH3 sur-
792 face. The proline 2 side-chain is also optimally
793 oriented to interact hydrophobically with the nearby
794 compass pocket, which explains the strong destabi-
795 lising effect of the P2A mutation. 796

The thermodynamic impact of Pro-Ala 797 substitutions upon the SH3–peptide interaction 798

The thermodynamic parameters shown in Table 1
799 for the SH3–p41 binding reveal that in all the
800 chimeras the binding is driven by a favourable
801 negative enthalpy change, partially balanced out by
802 an unfavourable negative binding entropy, as is
803 usual in SH3–peptide interactions.^{8,11–15} Interest-
804 ingly, as the binding affinity is progressively reduced
805 for the different mutants, the magnitudes of both
806 enthalpy and entropy terms diminish considerably.
807 These differences are relatively large between
808 mutants. For instance, the difference in enthalpy
809 between SPCp41 and mutant P10A is about 28 kJ
810 mol⁻¹ at 25 °C. These effects would be difficult to
811 explain merely by the loss of a few van der Waals
812 contacts in the bound state after each Pro-Ala
813 mutation. Moreover, the enthalpic effect from the
814

loss of van der Waals contacts would be compensated by an opposite change in the enthalpy of hydration.⁴⁵ Neither do changes in the hydrophobicity of the interaction at the SH3–peptide interface produced by the Pro-Ala mutations provide us by themselves with an appropriate explanation for the changes in the binding entropy. The hydrophobic effect makes a strong favourable contribution to the binding entropy due to the release of solvating water molecules produced by the interaction. Alanine substitution of any proline side-chain interacting with the SH3 surface does, however, produce a small reduction in the surface area that is buried upon binding, which would, in contrast to our observations, imply a small increase in the binding entropy with the Pro-Ala mutation.

The folding of the ligand into the PPII conformation should also be influential in the thermodynamics of this type of interaction.¹² In the absence of other effects, replacement of a proline by alanine in the ligand would be expected to result in a less favourable binding entropy due to the increase in conformational freedom of the free ligand. This contribution can in principle be evaluated with the mutation of the solvent-exposed Pro8 side-chain, which is not directly involved in the interfacial interaction. Our results for the P8A mutant indicate, however, only slightly more negative binding entropy accompanied by a slightly less negative enthalpy (both changes $<1 \text{ kJ mol}^{-1}$) (see Table 1). This is a somewhat unexpected result, given that the Pro-Ala mutation should increase the conformational degeneracy of the mutated residue and of its preceding one by several times.⁴⁶ A previous ITC study has shown that replacement of the similar Pro8 by Leu in the isolated p41 ligand produces a significant unfavourable change in enthalpy, partially balanced out by a more favourable entropy of binding to the Abl–SH3 domain, which is contrary to the initially expected result.¹⁴ Other effects, such as the special hydrogen-bond acceptor capability of the proline carbonyl group or the existence of cooperative effects propagated *via* specific water-mediated interactions through the rest of the ligand and the SH3 domain itself, have been put forward to explain such unexpected behaviour (see below). In contrast to these results, Hilser and co-workers, studying the binding of the Sos peptide to the C-terminal Sem-5 SH3 domain, reported that the mutation of solvent-exposed proline residues (at positions equivalent to Pro8 in this study) to either glycine or alanine produces a more negative binding enthalpy (of about -4 kJ mol^{-1}), partially offset by a more negative binding entropy (of about -2 kJ mol^{-1}).^{13,47} One possible explanation for these discrepancies is that the free p41 ligand and the mutant sequences may be more biased to PPII than are the Sos peptides. Since any residue followed by a proline residue is restricted by steric interactions primarily to the β -sheet region of the ϕ and ψ angles,^{48–50} the increase in configurational freedom of the unbound ligand segment would be relatively

small for most Pro-Ala mutants, as observed for P8A. For similar reasons, the differences in binding enthalpy and entropy between the different Pro-Ala mutants cannot be put down to differences in the energetic cost of the adoption of the PPII conformation of the ligand chain. These results therefore cannot be rationalised in terms of simple changes in the conformational freedom of the ligand chain alone and other contributions must therefore play a part in this behaviour.

Additional contributions to the special thermodynamics of the SH3–peptide interaction are to be found in the literature, including the presence of water-mediated hydrogen-bond networks at the binding interface and a reduction in conformational dynamics upon complexation.^{14–16,51} The significance of these effects in determining the binding energetics, though very important, is still poorly understood. It is quite likely that their relative contribution depends upon the specific details of the interaction between each SH3 domain and its binding partner. For instance, the high similarity in the structures of the complexes between the p41 peptide and either Abl–SH3 or Spc–SH3 has led to the suggestion that the large differences in their thermodynamic magnitudes of binding may be closely related to the absence of several water-mediated interactions in the latter.¹⁵ Nevertheless, if the mutational effects observed in our chimeras were primarily due to the destabilization of specific water-bridged interactions we would have found the greatest mutational effects at specific positions, instead of the apparent periodicity of the energy effects observed throughout the PPII helix. Accordingly, although our results do not rule out the influence of water-mediated interactions, a significant contribution to the thermodynamic changes produced by the Pro-Ala mutations may come from their effect upon the conformational dynamics of the SH3–peptide complex, as discussed below.

The intramolecular SH3–peptide binding shows enthalpy–entropy compensation

Strong enthalpy–entropy compensation in the thermodynamics of the SH3–peptide interaction is observed here for the various chimeric proteins (Figure 11). The existence of enthalpy–entropy compensation (EEC) is a ubiquitous property of biomolecular processes⁵² and has been interpreted as a natural property of multiple, weak, non-covalent interactions, such as those existing in macromolecular systems in water.^{53,54} An EEC behaviour, similar to that found here, has also been reported recently for the binding between bovine carbonic anhydrase II and para-substituted benzenesulfonamides with chains of different lengths.⁵⁵ These authors observed an increase in mobility at the binding interface of the complex concomitant with the increase in the ligand length, which results in a less unfavourable $T\Delta S$ and a less favourable ΔH of binding, similar to our observations. Other examples of EEC behaviour in binding processes

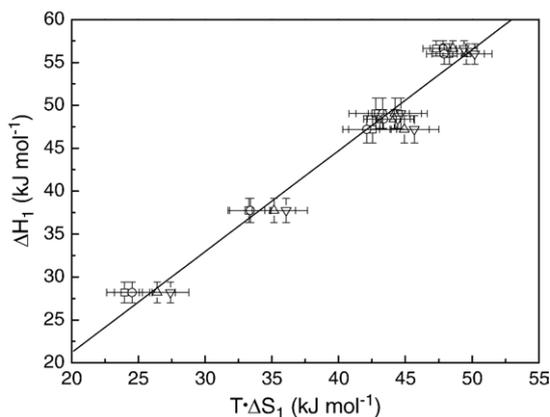


Figure 11. Enthalpy–entropy compensation for the intramolecular binding in the SH3–peptide chimeric proteins. The values of ΔH_1 and $T \cdot \Delta S_1$ (Table 1) were obtained from the three-state analysis of the DSC data shown in Figure 8. Symbols are as for Figure 8.

939 reviewed recently⁵⁶ illustrate an inverse relationship
 940 that frequently exists between structural tightening
 941 and residual motions, intensified by cooperative
 942 effects during the binding event. The strong EEC
 943 behaviour in our results also indicates that mobility
 944 in the bound conformation of the chimeric proteins
 945 diminishes concomitantly with an increase in affinity
 946 caused by cooperative improvement in the inter-
 947 facial interactions. This is supported by the NMR
 948 relaxation parameters, which show a significant
 949 increase in fast motions over the whole binding
 950 interface for the P7A mutant compared to SPCp41.

951 The decrease in conformational mobility may not
 952 be restricted just to the binding interface, as shown
 953 for several examples of binding between proline-
 954 rich peptides and SH3 domains and other protein
 955 ligand systems.^{15,16,57–59} Long-range propagation of
 956 conformational tightening would therefore produce
 957 an amplification of the local energy changes of
 958 single-point mutations. Unfortunately, however, the
 959 comparative NMR relaxation study shown here for
 960 the SPCp41 chimera and the P7A mutant does not
 961 provide us with information about long-range
 962 effects in the SH3 domain brought about by the
 963 slow aggregation of the mutant chimera, which
 964 hindered a more thorough analysis of the data.

965 Intramolecular versus intermolecular binding of 966 low-affinity sequences

967 We have shown here that despite the intrinsically
 968 low binding affinities of the isolated proline-rich
 969 peptides and the permutant SH3 domain, when the
 970 partners are tethered by a short linker the intramo-
 971 lecular binding exerts a surprisingly strong thermo-
 972 dynamic effect upon the stability of the chimeras.
 973 SH3-mediated intramolecular binding plays an
 974 essential role in the regulation of many signalling
 975 proteins such as, for example, in several families of
 976 tyrosine kinases, and it is well known that the in-
 977 tramolecular ligands in these proteins differ sig-

nificantly from high-affinity sequences for SH3
 978 domains.^{5–7} This has been described as a require-
 979 ment for effective regulation because they must be
 980 out-competed by external high-affinity ligands
 981 overcoming the entropic advantage imposed by
 982 the intramolecular tether.⁶ 983

Zhou discovered a simple, quantitative relation-
 984 ship between intramolecular and intermolecular
 985 binding constants between SH3 domains and their
 986 proline-rich ligands⁶⁰ with the simple form $K_i = K_b p$,
 987 where K_i and K_b are the intramolecular and in-
 988 termolecular binding constants, respectively, and p
 989 has the role of an effective concentration, generally
 990 in the mM range. This effective concentration is
 991 determined by the length and flexibility of the linker
 992 as a result of the distance restriction imposed be-
 993 tween the binding partners. Here we used DSC and
 994 ITC to measure both binding constants for the p41
 995 ligand and its P7A mutant (see Tables 1 and 2),
 996 arriving at a value for the effective concentration, p ,
 997 of between 9 and 16 mM, which agrees well with the
 998 values predicted for the linker used in these chimeric
 999 proteins.⁶⁰ A finely tuned balance between the
 1000 effective concentration and the intrinsic affinity
 1001 seems to be crucial to keeping the right equilibrium
 1002 between active and inactive forms of tyrosine
 1003 kinases, allowing some basal activity and efficient
 1004 kinase activation by a variety of signals.⁶¹ The
 1005 length of the linker between SH3 domains and
 1006 proline-rich sequences has also shown itself to be
 1007 essential to determining the competition between
 1008 intra- and intermolecular binding in the Tec-family
 1009 kinases.⁶² Understanding quantitatively the ener-
 1010 getic effects of tethering SH3 domains and their
 1011 ligands by flexible linkers is therefore of great
 1012 importance and our approach combining mutagen-
 1013 esis and DSC analyses of chimeric proteins may
 1014 serve as a powerful tool to unveil the thermody-
 1015 namic effects of varying both the length and com-
 1016 position of the linker, thus providing a convenient
 1017 way of validating the theoretical work. 1018

1019 Conclusions

We conclude that a mutagenesis analysis of the
 1020 thermodynamic stability of chimeric protein con-
 1021 structs mimicking the interaction between an SH3
 1022 domain and a proline-rich peptide may provide a
 1023 useful approach to analysing the thermodynamic
 1024 determinants of binding for this type of weak
 1025 interaction. The thermodynamic magnitudes of the
 1026 unfolding of these chimeras have shown themselves
 1027 to be very sensitive to perturbations in the binding
 1028 interactions produced by Pro-Ala substitutions in
 1029 the ligand sequence. A global analysis of the
 1030 unfolding of the different Pro-Ala mutant chimeras
 1031 on the basis of a three-state model has allowed us to
 1032 extract selectively the energetics of the intramo-
 1033 lecular binding and also to establish the energetic
 1034 importance of each proline residue in the interac-
 1035 tion, which agrees well with the model by Lim for
 1036 the binding between SH3 and proline-rich peptides. 1037

1038 The thermodynamic effects produced by the Pro-Ala
1039 mutations in the intramolecular binding can be put
1040 down mainly to the cooperative tightening of
1041 interactions and a considerable reduction in con-
1042 formational mobility, which extend over the whole
1043 binding interface and possibly throughout the
1044 remainder of the SH3 domain.

1045 Materials and Methods

1046 Protein samples

1047 The genes encoding the mutant chimeric proteins and
1048 the S19P20s domain were obtained by the polymerase
1049 chain reaction (PCR) method.⁶³ Mutant fragments were
1050 cloned into the pETM-11 plasmids (Protein Expression
1051 and Purification Core Facility, EMBL, Heidelberg, Ger-
1052 many) and over-expressed in *Escherichia coli* BL21/DE3
1053 cells. Proteins were purified as described.²⁰ Synthetic
1054 peptides for ITC were bought from Pepscan Presto B.V.
1055 (Lelystad, The Netherlands). Samples were prepared for
1056 experimental work by extensive dialysis against a large
1057 volume of the appropriate buffer, except for the NMR
1058 samples, which were prepared by dissolving the lyophi-
1059 lized protein (previously dialysed against water) directly
1060 in the buffer and subsequently readjusting their pH.
1061 Protein concentration was determined by absorption
1062 measurements at 280 nm using an extinction coefficient
1063 of 16,502 cm⁻¹ M⁻¹, determined as described.⁶⁴

1064 Circular dichroism measurements

1065 Circular dichroism (CD) measurements were made in a
1066 Jasco J-715 spectropolarimeter equipped with a tempera-
1067 ture-controlled cell holder. Far-UV spectra were recorded
1068 from 250 nm to 190 nm in a 1 mm path-length cuvette. For
1069 high-accuracy measurements of the ellipticity signal at
1070 228 nm, spectra were recorded from 250 nm to 215 nm
1071 using sample concentrations of ≈0.25 mg ml⁻¹–0.50 mg
1072 ml⁻¹. Near-UV spectra were measured from 320 nm to
1073 250 nm in a 5 mm path-length cuvette, with sample
1074 concentrations of ≈1.5–2.5 mg ml⁻¹. Each CD spectrum
1075 was obtained by averaging 20 accumulations collected at a
1076 scan rate of 50 nm min⁻¹. Baseline spectra obtained with
1077 pure buffer were subtracted for all spectra. Experimental
1078 data were normalized as mean-residue molar ellipticity.

1079 Intrinsic fluorescence measurements

1080 Fluorescence measurements were made in a Perkin-
1081 Elmer LS-55 spectrofluorimeter equipped with a tempera-
1082 ture-controlled cell holder. Spectra were recorded at 25 °C
1083 between 305 nm and 400 nm with an excitation wave-
1084 length of 298 nm. Each spectrum was the result of five
1085 accumulations collected at a scan rate of 200 nm min⁻¹. All
1086 the spectra were normalized by protein concentration.

1087 Differential scanning calorimetry

1088 Differential scanning calorimetry (DSC) experiments
1089 were carried out in a VP-DSC microcalorimeter (MicroCal
1090 Inc., Northampton, MA) at a scan rate of 1.5 K min⁻¹ using
1091 protein concentrations within the range of 1 mg ml⁻¹ to
1092 1.5 mg ml⁻¹. The samples were routinely heated from 5 °C

to 110 °C, then cooled down inside the calorimeter and
1093 reheated once again to check the reversibility of the
1094 unfolding process. Each experimental DSC thermogram
1095 was routinely corrected from the time-response of the
1096 calorimeter and from the instrumental baseline obtained
1097 with the calorimeter cells filled with the corresponding
1098 buffer. After normalization with the protein concentration,
1099 the partial molar heat-capacity curves (C_p) were calculated
1100 from the resulting thermograms, assuming 0.73 ml g⁻¹ for
1101 the partial specific volume and 8453.5 Da and 8427.5 Da
1102 for the molecular mass of the SPCp41 chimera and the Pro-
1103 Ala mutants, respectively. The C_p curves were subse-
1104 quently analysed using ORIGIN software. 1105

1106 Isothermal titration calorimetry

1107 Calorimetric titrations of the S19P20s domain with p41
1108 peptide were undertaken at 25 °C in a MCS titration
1109 microcalorimeter (Microcal Inc, Northampton, MA). The
1110 S19P20s solution was prepared by extensive dialysis
1111 overnight before being filtered, degassed and equili-
1112 brated at 25 °C prior to the experiment. The protein
1113 concentration in the cell was 0.44 mM. The p41 ligand
1114 solution was prepared at 10.8 mM by dissolving the
1115 lyophilised peptide in the dialysis buffer and readjusting
1116 the pH. Due to the low binding affinities the titrations
1117 were made by using a profile of injection volumes
1118 (between 2 μl and 13 μl) of ligand solution in order to
1119 define better the titration curve. After baseline subtraction,
1120 the area under each peak of the thermogram was
1121 integrated to determine the heat produced by the binding
1122 between the ligand and the protein after each injection. An
1123 independent identical titration with buffer in the cell was
1124 made with the same ligand solution to determine the
1125 corresponding heats of dilution, which were subtracted
1126 from the heats obtained in the titration of the protein. The
1127 apparent binding constant, K_b , and the enthalpy change of
1128 binding, ΔH_b , were determined by fitting the resulting
1129 titration isotherm to a model of binding to a single set of
1130 identical sites, as described.¹⁵

1131 Displacement ITC experiments were made by titration
1132 with a 10.8 mM solution of p41 over a 0.44 mM solution of
1133 S19P20s in the presence of the decapeptide APSYSPAPPP
1134 (P7Ap). Two independent experiments were made with
1135 P7Ap concentrations of 1.32 mM and 5.17 mM and a fixed
1136 0.44 mM concentration of S19P20s in the cell. The resulting
1137 binding isotherms were analysed as described by
1138 Sigurskjold,³⁶ fixing the thermodynamic parameters of
1139 the S19P20s:p41 interaction. 1139

1140 NMR spectroscopy

1141 ¹⁵N-labelled samples of SPCp41 and the P7A mutant
1142 were prepared for NMR experiments at 1 mM and
1143 0.5 mM, respectively, in 93% H₂O/7% ²H₂O, 50 mM per-
1144 deuterated glycine (pH 3.5). All NMR spectra were
1145 recorded at 25 °C on a Varian NMR Direct-Drive Systems
1146 600 MHz spectrometer (¹H frequency of 600.25 MHz).
1147 ¹⁵N-NOESY-HSQC spectra⁶⁵ were acquired for both
1148 samples for comparison and to assign the NMR spectrum
1149 of the P7A mutant. Due to the lower concentration, twice
1150 as many scans were accumulated to obtain the NMR
1151 spectra for the P7A mutant protein. 1151

1152 Heteronuclear ¹H-¹⁵N steady-state NOEs were obtained
1153 for both SpCp41 and mutant P7A from a series of 2D
1154 experiments with coherence selection achieved by pulse-
1155 field gradients at 600 MHz.⁶⁶ The ¹H-¹⁵N heteronuclear
1156 NOEs were determined from the ratio of peak intensities

1157 ($I_{\text{on}}/I_{\text{off}}$) with and without the saturation of the amide
 1158 protons for 3 s. To correct for the slight loss in intensity due
 1159 to the slow aggregation of P7A, the I_{on} peak intensities
 1160 were determined from the average intensities of the
 1161 spectra with saturation recorded just before and after the
 1162 non-saturation experiment.
 1163 ^{15}N R_1 and ^{15}N R_2 relaxation rates were measured for
 1164 SPCp41 and the P7A mutant. R_1 values were measured
 1165 from ten spectra with different relaxation delays: 10, 100,
 1166 200, 300, 400, 500, 700, 900, 1200 and 1500 ms. R_2 values
 1167 were extracted from $R_{1\rho}$ experiments^{67,68} recorded with
 1168 spin-lock pulses 10, 20, 40, 60, 80, 100, 130, 180 ms. The
 1169 P7A mutant chimera aggregated slowly during the
 1170 experimental time needed for these measurements, comp-
 1171 licating the analysis of such data. For this reason ^1H - ^{15}N
 1172 HSQC spectra were recorded after each R_1 and $R_{1\rho}$
 1173 experiment and used to correct the loss in intensity caused
 1174 by this slow aggregation. All NMR data were processed
 1175 using NMRPipe⁶⁹ and analysed by SPARKY (T.D. God-
 1176 dard and D.G. Kneller, University of California) or
 1177 NMRview.⁷⁰ Relaxation parameters were extracted with
 1178 the SPARKY program, using a two-parameter fitting and
 1179 Monte Carlo simulation to estimate the errors. The
 1180 rotational correlation time and R_2/R_1 ratios of SPCp41
 1181 were calculated with the Fast-HYDRONMR program³⁰
 1182 using a temperature of 5 °C, a viscosity of 1.1 cP and a
 1183 radius of 2.6 Å for the spherical elements used to build the
 1184 initial shell.

1185 Acknowledgements

1186 This research was funded by grants BIO2003-
 1187 04274, BIO2005-04650 and BIO2006-15517.C02.01
 1188 from the Spanish Ministry of Education and Science;
 1189 INTAS-03-51-5569 from the European Union; and
 1190 FQM-123 from the Andalusian Regional Govern-
 1191 ment. A.M.C. and N.A.J.v.N. benefited from an F.P.I.
 1192 fellowship and a “Ramón y Cajal” research grant,
 1193 respectively, from the Spanish Ministry of Education
 1194 and Science. We thank A. Palencia for his help with
 1195 the ITC experiments and the support of the C.I.C. of
 1196 the University of Granada and Varian Inc. with the
 1197 NMR measurements. We also thank our colleague
 1198 Dr J. Trout for revising our English text.

1199 Supplementary Data

1200 Supplementary data associated with this article
 1201 can be found, in the online version, at doi:10.1016/
 1202 j.jmb.2007.11.060

1203 References

- 1204 1. Gmeiner, W. H. & Horita, D. A. (2001). Implications of
 1205 SH3 domain structure and dynamics for protein
 1206 regulation and drug design. *Cell. Biochem. Biophys.*
 1207 **35**, 127–140.
- 1208 2. Kay, B. K., Williamson, M. P. & Sudol, M. (2000). The
 1209 importance of being proline: the interaction of proline-
 1210 rich motifs in signaling proteins with their cognate
 1211 domains. *Faseb J.* **14**, 231–241.

3. Mayer, B. J. (2001). SH3 domains: complexity in
 1212 moderation. *J. Cell Sci.* **114**, 1253–1263. 1213
4. Vidal, M., Gigoux, V. & Garbay, C. (2001). SH2 and
 1214 SH3 domains as targets for anti-proliferative agents. 1215
Crit. Rev. Oncol. Hematol. **40**, 175–186. 1216
5. Andreotti, A. H., Bunnell, S. C., Feng, S., Berg, L. J. &
 1217 Schreiber, S. L. (1997). Regulatory intramolecular
 1218 association in a tyrosine kinase of the Tec family. 1219
Nature, **385**, 93–97. 1220
6. Nguyen, J. T. & Lim, W. A. (1997). How Src exercises
 1221 self-restraint. *Nature Struct. Biol.* **4**, 256–260. 1222
7. Boggon, T. J. & Eck, M. J. (2004). Structure and regu-
 1223 lation of Src family kinases. *Oncogene*, **23**, 7918–7927. 1224
8. Arold, S., O'Brien, R., Franken, P., Strub, M. P., Hoh, F.,
 1225 Dumas, C. & Ladbury, J. E. (1998). RT loop flexibility
 1226 enhances the specificity of Src family SH3 domains for
 1227 HIV-1 Nef. *Biochemistry*, **37**, 14683–14691. 1228
9. Nguyen, J. T., Turck, C. W., Cohen, F. E., Zuckermann,
 1229 R. N. & Lim, W. A. (1998). Exploiting the basis
 1230 of proline recognition by SH3 and WW domains:
 1231 design of N-substituted inhibitors. *Science*, **282**,
 1232 2088–2092. 1233
10. Musacchio, A. (2002). How SH3 domains recognize
 1234 proline. *Adv. Protein Chem.* **61**, 211–268. 1235
11. Renzoni, D. A., Pugh, D. J., Siligardi, G., Das, P.,
 1236 Morton, C. J., Rossi, C. *et al.* (1996). Structural and
 1237 thermodynamic characterization of the interaction
 1238 of the SH3 domain from Fyn with the proline-rich
 1239 binding site on the p85 subunit of PI3-kinase. 1240
Biochemistry, **35**, 15646–15653. 1241
12. Ferreon, J. C. & Hilser, V. J. (2004). Thermodynamics of
 1242 binding to SH3 domains: the energetic impact of
 1243 polyproline II (PII) helix formation. *Biochemistry*, **43**,
 1244 7787–7797. 1245
13. Ferreon, J. C. & Hilser, V. J. (2003). The effect of the
 1246 polyproline II (PPII) conformation on the denatured
 1247 state entropy. *Protein Sci.* **12**, 447–457. 1248
14. Palencia, A., Cobos, E. S., Mateo, P. L., Martinez, J. C.
 1249 & Luque, I. (2004). Thermodynamic dissection of the
 1250 binding energetics of proline-rich peptides to the Abl-
 1251 SH3 domain: implications for rational ligand design. 1252
J. Mol. Biol. **336**, 527–537. 1253
15. Casares, S., AB, E., Eshuis, H., Lopez-Mayorga, O.,
 1254 van Nul, N. A. J. & Conejero-Lara, F. (2007). The high-
 1255 resolution NMR structure of the R21A Spc-SH3:P41
 1256 complex: understanding the determinants of binding
 1257 affinity by comparison with Abl-SH3. *BMC Struct.* 1258
Biol. **7**. 1259
16. Wang, C., Pawley, N. H. & Nicholson, L. K. (2001).
 1260 The role of backbone motions in ligand binding to the
 1261 c-Src SH3 domain. *J. Mol. Biol.* **313**, 873–887. 1262
17. Martin-Sierra, F. M., Candel, A. M., Casares, S.,
 1263 Filimonov, V. V., Martinez, J. C. & Conejero-Lara, F. 1264
 (2003). A binding event converted into a folding
 1265 event. *FEBS Letters*, **553**, 328–332. 1266
18. Martinez, J. C., Viguera, A. R., Berisio, R., Wilmanns,
 1267 M., Mateo, P. L., Filimonov, V. V. & Serrano, L. (1999).
 1268 Thermodynamic analysis of alpha-spectrin SH3 and
 1269 two of its circular permutants with different loop
 1270 lengths: discerning the reasons for rapid folding in
 1271 proteins. *Biochemistry*, **38**, 549–559. 1272
19. Pisabarro, M. T. & Serrano, L. (1996). Rational design
 1273 of specific high-affinity peptide ligands for the Abl-
 1274 SH3 domain. *Biochemistry*, **35**, 10634–10640. 1275
20. Candel, A. M., Conejero-Lara, F., Martinez, J. C. & van
 1276 Nul, N. A. (2007). and M. Bruix. The high-resolution
 1277 NMR structure of a single-chain chimeric protein
 1278 mimicking a SH3-peptide complex. *FEBS Letters*, **581**,
 1279 687–692. 1280

- 1281 21. Pisabarro, M. T., Serrano, L. & Wilmanns, M. (1998).
1282 Crystal structure of the abl-SH3 domain complexed
1283 with a designed high-affinity peptide ligand: implica-
1284 tions for SH3-ligand interactions. *J. Mol. Biol.* **281**,
1285 513–521.
- 1286 22. Makhatazde, G. I. & Privalov, P. L. (1995). Energetics
1287 of protein structure. *Adv. Protein Chem.* **47**, 307–425.
- 1288 23. Spolar, R. S. & Record, M. T., Jr (1994). Coupling of
1289 local folding to site-specific binding of proteins to
1290 DNA. *Science*, **263**, 777–784.
- 1291 24. Viguera, A. R., Arrondo, J. L., Musacchio, A.,
1292 Saraste, M. & Serrano, L. (1994). Characterization
1293 of the interaction of natural proline-rich peptides
1294 with five different SH3 domains. *Biochemistry*, **33**,
1295 10925–10933.
- 1296 25. Kelly, M. A., Chellgren, B. W., Rucker, A. L., Trout-
1297 man, J. M., Fried, M. G., Miller, A. F. & Creamer, T. P.
1298 (2001). Host-guest study of left-handed polyproline II
1299 helix formation. *Biochemistry*, **40**, 14376–14383.
- 1300 26. Woody, R. W. (1994). Contributions of tryptophan side
1301 chains to the far-ultraviolet circular dichroism of
1302 proteins. *Eur. Biophys. J.* **23**, 253–262.
- 1303 27. Jarymowycz, V. A. & Stone, M. J. (2006). Fast time
1304 scale dynamics of protein backbones: NMR relaxation
1305 methods, applications, and functional consequences.
1306 *Chem. Rev.* **106**, 1624–1671.
- 1307 28. Korzhnev, D. M., Karlsson, B. G., Orekhov, V. Y. &
1308 Billeter, M. (2003). NMR detection of multiple transi-
1309 tions to low-populated states in azurin. *Protein Sci.* **12**,
1310 56–65.
- 1311 29. Dosset, P., Hus, J. C., Blackledge, M. & Marion, D.
1312 (2000). Efficient analysis of macromolecular rotational
1313 diffusion from heteronuclear relaxation data. *J. Biomol.*
1314 *NMR*, **16**, 23–28.
- 1315 30. Ortega, A. & de la Torre, J. G. (2005). Efficient, accurate
1316 calculation of rotational diffusion and NMR relaxation
1317 of globular proteins from atomic-level structures and
1318 approximate hydrodynamic calculations. *J. Am. Chem.*
1319 *Soc.* **127**, 12764–12765.
- 1320 31. Prabhu, N. V. & Sharp, K. A. (2005). Heat capacity in
1321 proteins. *Annu. Rev. Phys. Chem.* **56**, 521–548.
- 1322 32. Eftink, M. R., Anusiem, A. C. & Biltonen, R. L. (1983).
1323 Enthalpy entropy compensation and heat-capacity
1324 changes for protein ligand interactions - general ther-
1325 modynamic models and data for the binding of nu-
1326 cleotides to ribonuclease-A. *Biochemistry*, **22**, 3884–3896.
- 1327 33. Fisher, H. F., Colen, A. H. & Medary, R. T. (1981).
1328 Temperature-dependent ΔC_P generated by a
1329 shift in equilibrium between macrostates of an en-
1330 zyme. *Nature*, **292**, 271–272.
- 1331 34. Makhatazde, G. I. & Privalov, P. L. (1990). Heat
1332 capacity of proteins. I. Partial molar heat capacity of
1333 individual amino acid residues in aqueous solution:
1334 hydration effect. *J. Mol. Biol.* **213**, 375–384.
- 1335 35. Wiseman, T., Williston, S., Brandts, J. F. & Lin, L. N.
1336 (1989). Rapid measurement of binding constants and
1337 heats of binding using a new titration calorimeter.
1338 *Anal. Biochem.* **179**, 131–137.
- 1339 36. Sigurskjold, B. W. (2000). Exact analysis of competi-
1340 tion ligand binding by displacement isothermal
1341 titration calorimetry. *Anal. Biochem.* **277**, 260–266.
- 1342 37. Brandts, J. F. & Lin, L. N. (1990). Study of strong to
1343 ultratight protein interactions using differential scan-
1344 ning calorimetry. *Biochemistry*, **29**, 6927–6940.
- 1345 38. Martinez, J. C., Filimonov, V. V., Mateo, P. L.,
1346 Schreiber, G. & Fersht, A. R. (1995). A calorimetric
1347 study of the thermal-stability of barstar and its inter-
1348 action with barnaset. *Biochemistry*, **34**, 5224–5233.
- 1349 39. Martinez, J. C., Elharrous, M., Filimonov, V. V., Mateo,
P. L. & Fersht, A. R. (1994). A calorimetric study of
the thermal-stability of barnase and its interaction
with 3'gmp. *Biochemistry*, **33**, 3919–3926.
40. Jelesarov, I. & Bosshard, H. R. (1999). Isothermal titra-
tion calorimetry and differential scanning calorimetry
as complementary tools to investigate the energetics
of biomolecular recognition. *J. Mol. Recognit.* **12**, 3–18.
41. Gmeiner, W. H., Xu, J. Z., Horita, D. A., Smithgall,
T. E., Engen, J. R., Smith, D. L. & Byrd, R. A. (2001).
Intramolecular binding of a proximal PPII helix to an
SH3 domain in the fusion protein SH3Hck : PPIIh-
GAP. *Cell. Biochem. Biophys.* **35**, 115–126.
42. Freund, C., Kuhne, R., Park, S., Thiemke, K., Reinherz,
E. L. & Wagner, G. (2003). Structural investigations
of a GYF domain covalently linked to a proline-rich
peptide. *J. Biomol. NMR*, **27**, 143–149.
43. Ren, R., Mayer, B. J., Cicchetti, P. & Baltimore, D.
(1993). Identification of a ten-amino acid proline-rich
SH3 binding site. *Science*, **259**, 1157–1161.
44. Lim, W. A., Richards, F. M. & Fox, R. O. (1994).
Structural determinants of peptide-binding orienta-
tion and of sequence specificity in SH3 domains.
Nature, **372**, 375–379.
45. Hilser, V. J., Gomez, J. & Freire, E. (1996). The enthalpy
change in protein folding and binding: refinement of
parameters for structure-based calculations. *Proteins:*
Struct. Funct. Genet. **26**, 123–133.
46. Daquino, J. A., Gomez, J., Hilser, V. J., Lee, K. H.,
Amzel, L. M. & Freire, E. (1996). The magnitude of the
backbone conformational entropy change in protein
folding. *Proteins: Struct. Funct. Genet.* **25**, 143–156.
47. Hamburger, J. B., Ferreon, J. C., Whitten, S. T. & Hilser,
V. J. (2004). Thermodynamic mechanism and con-
sequences of the polyproline II (P-II) structural bias
in the denatured states of proteins. *Biochemistry*, **43**,
9790–9799.
48. MacArthur, M. W. & Thornton, J. M. (1991). Influence
of proline residues on protein conformation. *J. Mol.*
Biol. **218**, 397–412.
49. Creamer, T. P. (1998). Left-handed polyproline II helix
formation is (very) locally driven. *Proteins: Struct.*
Funct. Genet. **33**, 218–226.
50. Rucker, A. L., Pager, C. T., Campbell, M. N., Qualls,
J. E. & Creamer, T. P. (2003). Host-guest scale of left-
handed polyproline II helix formation. *Proteins: Struct.*
Funct. Genet. **53**, 68–75.
51. Ferreon, J. C. & Hilser, V. J. (2003). Ligand-induced
changes in dynamics in the RT loop of the C-terminal
SH3 domain of Sem-5 indicate cooperative conforma-
tional coupling. *Protein Sci.* **12**, 982–996.
52. Lumry, R. & Rajender, S. (1970). Enthalpy–entropy
compensation phenomena in water solutions of pro-
teins and small molecules: a ubiquitous property of
water. *Biopolymers*, **9**, 1125–1227.
53. Dunitz, J. D. (1995). Win some, lose some: enthalpy-
entropy compensation in weak intermolecular inter-
actions. *Chem. Biol.* **2**, 709–712.
54. Cooper, A., Johnson, C. M., Lakey, J. H. & Nollmann,
M. (2001). Heat does not come in different colours:
entropy-enthalpy compensation, free energy win-
dows, quantum confinement, pressure perturbation
calorimetry, solvation and the multiple causes of heat
capacity effects in biomolecular interactions. *Biophys.*
Chem. **93**, 215–230.
55. Krishnamurthy, V. M., Bohall, B. R., Semetey, V. &
Whitesides, G. M. (2006). The paradoxical thermo-
dynamic basis for the interaction of ethylene glycol,
glycine, and sarcosine chains with bovine carbonic an-
hydrase II: an unexpected manifestation of enthalpy/

- 1419 entropy compensation. *J. Am. Chem. Soc.* **128**,
1420 5802–5812.
- 1421 56. Williams, D. H., Stephens, E., O'Brien, D. P. & Zhou,
1422 M. (2004). Understanding non-covalent interactions:
1423 Ligand binding energy and catalytic efficiency from
1424 ligand-induced reductions in motion within receptors
1425 and enzymes. *Angew. Chem. Int. Edit.* **43**, 6596–6616.
- 1426 57. Mayne, L., Paterson, Y., Cerasoli, D. & Englander,
1427 S. W. (1992). Effect of antibody-binding on protein
1428 motions studied by hydrogen-exchange labeling and
1429 2-dimensional NMR. *Biochemistry*, **31**, 10678–10685.
- 1430 58. Wildes, D. & Marqusee, S. (2005). Hydrogen exchange
1431 and ligand binding: ligand-dependent and ligand-
1432 independent protection in the Src SH3 domain. *Protein*
1433 *Sci.* **14**, 81–88.
- 1434 59. Polshakov, V. I., Birdsall, B. & Feeney, J. (2006). Effects
1435 of co-operative ligand binding on protein amide NH
1436 hydrogen exchange. *J. Mol. Biol.* **356**, 886–903.
- 1437 60. Zhou, H. X. (2006). Quantitative relation between
1438 intermolecular and intramolecular binding of Pro-rich
1439 peptides to SH3 domains. *Biophys. J.* **91**, 3170–3181.
- 1440 61. Ayrapetov, M. K., Wang, Y. H., Lin, X. F., Gu, X. F.,
1441 Parang, K. & Sun, G. Q. (2006). Conformational basis
1442 for SH2-Tyr(P)(527) binding in Src inactivation. *J. Biol.*
1443 *Chem.* **281**, 23776–23784.
- 1444 62. Laederach, A., Cradic, K. W., Fulton, D. B. &
1445 Andreotti, A. H. (2003). Determinants of intra versus
1446 intermolecular self-association within the regulatory
1447 domains of Rlk and Itk. *J. Mol. Biol.* **329**, 1011–1020.
- 1448 63. Higuchi, R., Krummel, B. & Saiki, R. K. (1988). A
1449 general method of *in vitro* preparation and specific
mutagenesis of DNA fragments: study of protein and
DNA interactions. *Nucl. Acids Res.* **16**, 7351–7367. 1450
1451
64. Gill, S. C. & Von Hippel, P. H. (1989). Calculation of
1452 protein extinction coefficients from amino-acid
1453 sequence data. *Anal. Biochem.* **182**, 319–326. 1454
65. Palmer, A. G., Cavanagh, J., Byrd, R. A. & Rance, M.
1455 (1992). Sensitivity improvement in 3-dimensional
1456 heteronuclear correlation NMR-spectroscopy. *J. Magn.*
1457 *Reson.* **96**, 416–424. 1458
66. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal,
1459 S. M., Kay, C. M., Gish, G. *et al.* (1994). Backbone
1460 dynamics of a free and a phosphopeptide-complexed
1461 Src homology-2 domain studied by N-15 NMR
1462 relaxation. *Biochemistry*, **33**, 5984–6003. 1463
67. Peng, J. W., Thanabal, V. & Wagner, G. (1991).
1464 Improved accuracy of heteronuclear transverse
1465 relaxation-time measurements in macromolecules -
1466 elimination of antiphase contributions. *J. Magn. Reson.*
1467 **95**, 421–427. 1468
68. Sattler, M., Schleucher, J. & Griesinger, C. (1999).
1469 Heteronuclear multidimensional NMR experiments
1470 for the structure determination of proteins in solution
1471 employing pulsed field gradients. *Prog. Nucl. Mag.*
1472 *Res. Sp.* **34**, 93–158. 1473
69. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G.,
1474 Pfeifer, J. & Bax, A. (1995). Nmrpipe - a multidimen-
1475 sional spectral processing system based on Unix
1476 Pipes. *J. Biomol. NMR*, **6**, 277–293. 1477
70. Johnson, B. A. & Blevins, R. A. (1994). Nmr View - a
1478 computer-program for the visualization and analysis
1479 of NMR data. *J. Biomol. NMR*, **4**, 603–614. 1480
1481

1482

1483

1484