

Universidad Miguel Hernández
Instituto de Biología Molecular y Celular

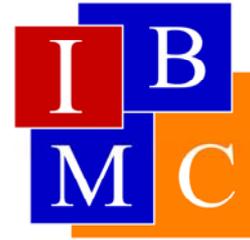


*Estabilidad, plegamiento y asociación
del dominio de dimerización de la
proteína de la cápsida del virus de la
inmunodeficiencia humana*

TESIS DOCTORAL

María del Carmen Lidón Moya

Elche, 2005



D. José Manuel González Ros, Director del Instituto de Biología Molecular y Celular de la Universidad Miguel Hernández de Elche,

DA SU CONFORMIDAD a la lectura de la tesis doctoral titulada: “Estabilidad, plegamiento y asociación del dominio de dimerización de la proteína de la cápsida del virus de la inmunodeficiencia humana” presentada por María del Carmen Lidón Moya.

Para que conste y surta los efectos oportunos, firma el presente certificado en Elche, a 22 de Julio de 2005.

Fdo. José Manuel González Ros



D. José Luis Neira Faleiro, Profesor Titular de Química-Física,
(Departamento de Agroquímica y Medio Ambiente) de la Universidad
Miguel Hernández de Elche,

CERTIFICA:

Que el trabajo de investigación que conduce a la obtención del grado de doctor titulado: “Estabilidad, plegamiento y asociación del dominio de dimerización de la proteína de la cápsida del virus de la inmunodeficiencia humana”, del que es autora Dña. María del Carmen Lidón Moya, ha sido realizado bajo su dirección en el Instituto de Biología Molecular y Celular de la Universidad Miguel Hernández de Elche.

Para que conste y surta los efectos oportunos, firma el presente certificado en Elche, a 22 de Julio de 2005.

Fdo. José Luis Neira Faleiro

AGRADECIMIENTOS

En primer lugar, quiero agradecer al Dr. José Luis Neira, por haberme dado la oportunidad de realizar este trabajo bajo su dirección, por orientarme, y haber compartido conmigo sus amplios conocimientos.

Al Dr. Mauricio G. Mateu, por su estrecha colaboración durante todo el proyecto, y por acogerme en su laboratorio. A Marta del Álamo, por toda su ayuda durante mi estancia.

Al Dr. Javier Sancho, por su colaboración, por permitirme trabajar en su laboratorio, y darme la oportunidad de conocer a una gente estupenda, entre ellos a Marta Bueno, que me hizo sentir como en casa.

Al Dr. Javier Gómez, por su participación en el trabajo realizado, por facilitarme su ayuda, y por su gran amabilidad.

A Fran, por su dedicación, por sus consejos e ideas, y por estar siempre dispuesto a colaborar.

A Maite, por su participación, y su generosidad.

A Marco, por todo lo que me ha enseñado sobre biología molecular, y por sus ánimos. ¡Eres un cielo!

A Estefanía, muy buena compañera y mejor amiga. Gracias por toda tu ayuda, por escucharme, aconsejarme y animarme siempre, y por los buenos momentos que hemos pasado juntas; porque sin ti, mi estancia en el laboratorio no habría sido lo mismo.

A Lourdes, por la confianza que me da, por su simpatía y su bondad, por su paciencia conmigo, y por toda su ayuda. Nunca podré agradecerte todo lo que has hecho por mi. ¡Millones de gracias!

A Marcela y José Antonio, por el interés que han mostrado, por todo su apoyo, y porque sé que puedo contar ellos.

A Maria José, mi amiga de la Vega, por la alegría que me transmite. ¡Gracias por brindarme tu amistad!

A Antonio, por los buenos momentos que me ha hecho pasar, y por enseñarme a ver el lado positivo de las cosas. ¡Gracias primo!

A los compis del laboratorio 2.2, por su disposición, y por ser fuente de simpatía. ¡Os deseo lo mejor!

A Eva Martínez, María, Estefanía, Isabel, Daianne, Nuria, Eva Vera, Ana y Ángeles, por haber tenido siempre una palabra de aliento, y por todos esos pequeños detalles que, en suma, hacen un mundo.

A Javier, por su gran amabilidad, y por estar siempre dispuesto a solucionar esos “problemillas” que van surgiendo día a día. A May, por su implicación personal y cariño demostrado, por su apoyo constante y toda su ayuda. A Carmen, por todas las facilidades prestadas. ¡Qué sería del IBMC sin vosotros!

En general, quiero expresar mi gratitud a toda la gente del IBMC que de una u otra forma me ha ayudado.

A mis amigos, M^a del Mar, Encarni, Eloisa, Coque, Bea, Paqui, Virginia, y M^a Dolores, por todo lo vivido, por las risas y energía positiva.

A Caty, la mejor amiga que se puede tener, por ser como eres, única e imprescindible. Tengo infinita suerte de contar contigo y saber que siempre estarás ahí. ¡No cambies nunca!

A toda mi familia, mis abuelitos, mis tios, en especial a mis padrinos, mis primos, y a mi ahijado, por quererme tanto.

A mis niños mimados: Daniel, por su sinceridad; Isabel M^a, por ser mi fan n^o 1; Lucía, por su espontaneidad; y Paola, por las caritas que me pone. A los cuatro, por todo lo que me habéis enseñado a valorar. ¡Sois mis tesoritos!

A ti, Beatriz, la estrella que me ilumina desde el cielo. No sabes cuánto te echo de menos, pero sé que desde donde quiera que estés me sigues sonriendo. Por todo lo que me diste. ¡Siempre te llevaré en mi corazón!

A mis hermanos, Miguel Ángel y José Manuel, por estar siempre a mi lado, por aguantar mis malos humos, y por ayudarme a superar los peores momentos. ¡Sois los mejores!

A mis padres, por quererme sin límites, por confiar en mi más que yo misma, y por apoyarme incondicionalmente. En definitiva, porque todo os lo debo a vosotros. ¡Os quiero muchísimo!

A mis padres y hermanos



ABREVIATURAS

ADN	Ácido desoxirribonucleico
ANS	Ácido 1-anilino-8-naftalenosulfónico
ARN	Ácido ribonucleico
ASA	Área accesible al disolvente
BrCN	Bromuro de cianógeno
CA	Proteína de la cápsida del VIH-1
CA-C	Dominio C-terminal de la proteína de la cápsida del VIH-1
CA-N	Dominio N-terminal de la proteína de la cápsida del VIH-1
CAC1	Péptido que comprende la α -hélice 2 de CA-C
CAC1Aib	Péptido CAC1 marcado con un grupo antranilo en el grupo ϵ -amino de la cadena lateral del residuo Lys182
ΔC_p	Incremento de capacidad calorífica de desnaturalización
ΔG	Incremento de energía libre de desnaturalización
ΔH_m	Incremento de entalpía a la temperatura media de desnaturalización
ΔS	Incremento de entropía
DC	Dicroísmo circular
DTGS	Sulfato de triglicina deuterado
DTT	Ditiotreitol
EDTA	Ácido etilendiaminotetracético
FTIR	Infrarrojo por transformada de Fourier
GdmHCl	Cloruro de guanidinio
IN	Integrasa del VIH-1
IPTG	Isopropil-1-tio- β -D-galactopiranosido
ITC	Calorimetría de titulación isotérmica
K_D	Constante de disociación
K_{sv}	Constante de Stern-Volmer
LB	Medio de cultivo Luria-Bertani
MA	Proteína de la matriz del VIH-1
MHR	Región principal de homología de CA del VIH-1
NC	Proteína de la nucleocápsida del VIH-1

Abreviaturas

PCR	Reacción en cadena de la polimerasa
pI(s)	Punto(s) isoeléctrico(s) (en unidades de $-\log$)
PR	Proteasa del VIH-1
RMN	Resonancia magnética nuclear
rpm	Revoluciones por minuto
RT	Transcriptasa inversa del VIH-1
SDS	Dodecilsulfato sódico
SDS-PAGE	Electroforesis desnaturalizante en gel de poliacrilamida
SIDA	Síndrome de la inmunodeficiencia adquirida
T_m	Temperatura media de desnaturalización
Tris	Tris-hidroximetil-aminometano
TSP	Sal sódica del ácido 3-(trimetilsilil)propiónico-2,2,2,2- $^2\text{H}_4$
UV	Ultravioleta
V_e	Volumen de elución
VIH-1	Virus de la inmunodeficiencia humana tipo I
VIH-2	Virus de la inmunodeficiencia humana tipo II
$[\Theta]_{222}$	Elipticidad molar por residuo a 222 nm

ÍNDICE

1. INTRODUCCIÓN.	1
1.1. LAS PROTEÍNAS Y SU ESTRUCTURA.	3
1.1.1. LAS PROTEÍNAS.	3
1.1.2. LA ESTRUCTURA DE LAS PROTEÍNAS.	5
1.1.2.1. La estructura primaria.	5
1.1.2.2. La estructura secundaria.	5
1.1.2.3. La estructura terciaria.	9
1.1.2.4. La estructura cuaternaria.	9
1.2. EL PLEGAMIENTO DE LAS PROTEÍNAS: RECONOCIMIENTO INTRAMOLECULAR.	10
1.2.1. EL PROBLEMA DEL PLEGAMIENTO.	10
1.2.2. ESTADOS CONFORMACIONALES EN LA REACCIÓN DE PLEGAMIENTO.	11
1.2.2.1. El estado desnaturalizado.	11
1.2.2.2. El estado nativo.	12
1.2.2.3. Intermediarios de plegamiento: el glóbulo fundido.	12
1.2.3. PLEGAMIENTO DE PROTEÍNAS CON VARIOS DOMINIOS Y DE PROTEÍNAS OLIGOMÉRICAS.	13
1.2.4. ESTUDIO DEL PLEGAMIENTO: TÉCNICAS EXPERIMENTALES.	14
1.3. UNIÓN A MACROMOLÉCULAS: RECONOCIMIENTO INTERMOLECULAR.	14
1.3.1. LA INTERACCIÓN ENTRE MACROMOLÉCULAS.	14
1.3.2. ESTUDIO DE LA INTERACCIÓN CON MACROMOLÉCULAS.	15
1.4. TÉCNICAS BIOFÍSICAS.	16
1.4.1. TÉCNICAS UTILIZADAS PARA DETERMINAR LA ESTRUCTURA DE LAS PROTEÍNAS Y ESTUDIAR EL PLEGAMIENTO Y LA UNIÓN A MACROMOLÉCULAS.	16
1.4.1.1. Técnicas espectroscópicas.	16
1.4.1.1.1. Absorbancia.	16
1.4.1.1.2. Fluorescencia.	17
1.4.1.1.3. Dicroísmo circular.	19

1.4.1.1.4. Infrarrojo por transformada de Fourier.	22
1.4.1.1.5. Resonancia magnética nuclear.	23
1.4.1.2. Técnicas hidrodinámicas: cromatografía de exclusión por tamaño molecular.	24
1.4.2. TÉCNICAS UTILIZADAS PARA DETERMINAR CONSTANTES DE UNIÓN.	25
1.4.2.1. Cromatografía de afinidad analítica.	26
1.4.2.2. Calorimetría de titulación isotérmica.	26
1.5. EL DOMINIO DE DIMERIZACIÓN DE LA PROTEÍNA DE LA CÁPSIDA DEL VIH-1 (CA-C).	27
1.5.1. EL VIRUS DE LA INMUNODEFICIENCIA HUMANA: ESTRUCTURA.	27
1.5.1.1. El virus de la inmunodeficiencia humana.	27
1.5.1.2. Estructura molecular del VIH-1.	28
1.5.2. LA PROTEÍNA DE LA CÁPSIDA DEL VIH-1.	29
1.5.2.1. Maduración, morfología y composición de la cápsida del VIH-1.	29
1.5.2.2. Estructura de la proteína de la cápsida del VIH-1.	29
1.5.3. JUSTIFICACIÓN DE LA IMPORTANCIA DEL ESTUDIO ESTRUCTURAL Y TERMODINÁMICO DE CA-C.	34
1.6. OBJETIVOS.	35
2. CARACTERIZACIÓN TERMODINÁMICA DE CA-C.	37
2.1. MATERIALES Y MÉTODOS.	39
2.1.1. PRODUCCIÓN DE CA-C.	39
2.1.1.1. Expresión de CA-C.	39
2.1.1.2. Purificación de CA-C.	40
2.1.1.3. Determinación de la pureza y concentración de CA-C.	41
2.1.2. FLUORESCENCIA.	42
2.1.2.1. Medidas en el estado estacionario.	42
2.1.2.1.1. Fluorescencia intrínseca de triptófano y tirosina.	42
2.1.2.1.2. Fluorescencia del ANS.	42
2.1.2.1.3. <i>Quenching</i> de fluorescencia de triptófano y tirosina.	43
2.1.2.2. Medidas de desnaturalización térmica.	44
2.1.2.2.1. Fluorescencia intrínseca de triptófano y tirosina.	44

2.1.2.2.2. Fluorescencia del ANS.	44
2.1.2.2.3. Anisotropía.	45
2.1.3. DICROÍSMO CIRCULAR Y ABSORBANCIA.	45
2.1.3.1. Medidas en el estado estacionario.	45
2.1.3.2. Medidas de desnaturalización térmica.	46
2.1.4. CROMATOGRAFÍA DE EXCLUSIÓN POR TAMAÑO MOLECULAR.	47
2.1.5. INFRARROJO POR TRANSFORMADA DE FOURIER.	48
2.1.6. RESONANCIA MAGNÉTICA NUCLEAR.	48
2.1.7. ANÁLISIS DE LAS CURVAS DE DESNATURALIZACIÓN.	49
2.2. RESULTADOS.	51
2.2.1. DESNATURALIZACIÓN POR pH.	51
2.2.1.1. Fluorescencia intrínseca de triptófano y tirosina.	51
2.2.1.2. Fluorescencia del ANS.	53
2.2.1.3. <i>Quenching</i> de fluorescencia de triptófano y tirosina.	54
2.2.1.4. Dicroísmo circular en el ultravioleta lejano.	56
2.2.1.5. Cromatografía de exclusión por tamaño molecular.	57
2.2.1.6. Infrarrojo por transformada de Fourier.	58
2.2.2. DESNATURALIZACIÓN TÉRMICA.	58
2.2.2.1. Fluorescencia intrínseca de triptófano y tirosina.	59
2.2.2.2. Fluorescencia del ANS.	61
2.2.2.3. Anisotropía.	62
2.2.2.4. Dicroísmo circular en el ultravioleta lejano.	62
2.2.2.5. Dicroísmo circular en el ultravioleta cercano y absorbancia.	64
2.2.2.6. Infrarrojo por transformada de Fourier.	66
2.2.2.7. Resonancia magnética nuclear.	68
2.2.3. ESTABILIDAD CONFORMACIONAL Y PARÁMETROS TERMODINÁMICOS DE LA REACCIÓN DE DESPLEGAMIENTO DE LA FORMA MONOMÉRICA DE CA-C A pH 7.	72
2.3. DISCUSIÓN.	74
2.3.1. CAMBIOS CONFORMACIONALES DE CA-C CON EL pH.	74
2.3.1.1. Cambios conformacionales de CA-C a pH ácido.	74
2.3.1.2. Cambios conformacionales de CA-C a pH básico.	76

2.3.2. DESNATURALIZACIÓN TÉRMICA DE CA-C.	76
2.3.2.1. Disociación de la forma dimérica de CA-C.	76
2.3.2.2. Estados conformacionales de la forma monomérica de CA-C.	77
2.3.2.3. Equilibrio de desnaturalización térmica de CA-C.	78
2.3.3. PARÁMETROS TERMODINÁMICOS DE LA DISOCIACIÓN Y DESNATURALIZACIÓN DE LAS FORMAS DIMÉRICA Y MONOMÉRICA DE CA-C: COMPARACIÓN CON MODELOS TEÓRICOS.	80
2.3.4. POSIBLES IMPLICACIONES BIOLÓGICAS.	82
3. DISEÑO DE UN PÉPTIDO Y ESTUDIO DE SU INTERACCIÓN CON CA-C.	83
3.1. MATERIALES Y MÉTODOS.	85
3.1.1. PRODUCCIÓN DE PROTEÍNA.	85
3.1.2. DISEÑO DEL PÉPTIDO, SÍNTESIS Y PURIFICACIÓN.	85
3.1.3. FLUORESCENCIA.	86
3.1.3.1. Determinación de la constante de autodisociación aparente de CAC1 y CAC1Aib mediante fluorescencia en el estado estacionario.	86
3.1.3.2. Determinación de la constante de autodisociación aparente de CAC1 y CAC1Aib mediante anisotropía en el estado estacionario.	88
3.1.3.3. Determinación de la constante de disociación aparente del complejo CAC1-CA-C mediante fluorescencia en el estado estacionario.	88
3.1.3.4. <i>Quenching</i> de fluorescencia del grupo antranilo.	89
3.1.4. DICROÍSMO CIRCULAR.	90
3.1.4.1. Determinación de la constante de autodisociación aparente de CAC1.	90
3.1.4.2. Medidas de desnaturalización térmica.	90
3.1.5. CROMATOGRAFÍA DE EXCLUSIÓN POR TAMAÑO MOLECULAR.	91
3.1.6. RESONANCIA MAGNÉTICA NUCLEAR.	91
3.1.7. CROMATOGRAFÍA DE AFINIDAD ANALÍTICA.	91
3.1.8. CALORIMETRÍA DE TITULACIÓN ISOTÉRMICA.	92
3.2. RESULTADOS.	93
3.2.1. CARACTERIZACIÓN ESPECTROSCÓPICA DE CAC1 Y CAC1Aib.	93
3.2.1.1. Dicroísmo circular en el ultravioleta lejano.	93

3.2.1.2. Fluorescencia en el estado estacionario.	94
3.2.1.3. Resonancia magnética nuclear.	95
3.2.1.4. Diseño de un péptido CAC1 marcado con un grupo antrnilo y estudio de sus propiedades de autoasociación.	96
3.2.2. UNIÓN DE CAC1 A CA-C.	96
3.2.2.1. Interacción CAC1-CA-C.	97
3.2.2.1.1. Dicroísmo circular.	97
3.2.2.1.2. Cromatografía de exclusión por tamaño molecular.	98
3.2.2.1.3. Resonancia magnética nuclear.	99
3.2.2.2. Determinación de la constante de disociación aparente del complejo CAC1-CA-C.	101
3.3. DISCUSIÓN.	103
3.3.1. ESTRUCTURA DE CAC1.	103
3.3.2. RECONOCIMIENTO MOLECULAR DE CA-C POR CAC1.	104
4. DISEÑO, EXPRESIÓN Y PURIFICACIÓN DE MINIPROTEÍNAS.	107
4.1. MATERIALES Y MÉTODOS.	109
4.1.1. DISEÑO DE PROTEÍNAS MUTANTES.	109
4.1.2. MUTAGÉNESIS DIRIGIDA.	110
4.1.2.1. Proceso de mutagénesis.	110
4.1.2.2. Diseño de oligonucleótidos.	111
4.1.2.3. Método experimental.	112
4.1.3. PRODUCCIÓN DE PROTEÍNAS MUTANTES.	113
4.1.3.1. Expresión y purificación de proteínas mutantes.	113
4.1.3.1.1. Comprobación de la expresión.	114
4.1.3.1.2. Expresión a gran escala y purificación.	115
4.1.3.2. Determinación de la pureza y concentración de proteínas mutantes.	115
4.1.4. DIGESTIÓN DE PROTEÍNAS MUTANTES CON BrCN Y PURIFICACIÓN DE LOS FRAGMENTOS OBTENIDOS.	115
4.2. RESULTADOS.	117
4.2.1. OBTENCIÓN DE PROTEÍNAS MUTANTES DE CA-C QUE SÓLO CONTIENEN LOS ELEMENTOS DE ESTRUCTURA SECUNDARIA HASTA LA α-HÉLICE 2.	117

4.2.1.1. Diseño y análisis de la expresión del mutante $\Delta 196$.	118
4.2.1.2. Diseño y análisis de la expresión del mutante M185A/M214I/M215I/N195M.	118
4.2.1.3. Diseño y análisis de la expresión del mutante M185A/M214I/M215I/C198M.	120
4.2.2. OBTENCIÓN DE PROTEÍNAS MUTANTES DE CA-C QUE SÓLO CONTIENEN LOS ELEMENTOS DE ESTRUCTURA SECUNDARIA HASTA LA α-HÉLICE 3.	121
4.2.2.1. Diseño y análisis de la expresión del mutante $\Delta 214$.	121
4.2.2.2. Obtención y digestión del mutante M185A.	122
4.2.2.2.1. Diseño, análisis de la expresión, expresión a gran escala y purificación.	122
4.2.2.2.2. Digestión con BrCN y purificación de los fragmentos obtenidos.	123
4.2.2.3. Obtención y digestión del mutante M185A/C218S.	124
4.2.2.3.1. Diseño, análisis de la expresión, expresión a gran escala y purificación.	124
4.2.2.3.2. Digestión con BrCN y purificación de los fragmentos obtenidos.	125
4.3. DISCUSIÓN.	126
5. CONCLUSIONES.	129
6. BIBLIOGRAFÍA.	133
7. APÉNDICE.	151

INTRODUCCIÓN



1.1. LAS PROTEÍNAS Y SU ESTRUCTURA.

1.1.1. LAS PROTEÍNAS.

Las proteínas son las macromoléculas biológicas más abundantes de las células vivas, estando presentes en todas las células y en todas las partes de las mismas. Son las moléculas que expresan la información contenida en los genes y están dotadas de gran versatilidad, tanto estructural como funcional. De hecho, prácticamente todos los procesos celulares requieren la intervención de una o más proteínas: proporcionan soporte estructural a las células, catalizan reacciones celulares, transportan sustancias a través de las membranas, regulan la función celular, e intervienen en procesos de reconocimiento molecular específicos como la interacción hormona-receptor o antígeno-anticuerpo, entre otros. Por todo ello, son elementos funcionales clave para la vida.

Desde el punto de vista químico, las proteínas son polímeros lineales de aminoácidos. Un aminoácido es una molécula orgánica que consta de un átomo de carbono, denominado C_{α} , al que se unen un grupo amino ($-NH_2$), de naturaleza básica, y un grupo carboxilo ($-COOH$), de carácter ácido. Además, a este C_{α} se une, como tercer sustituyente, un átomo de hidrógeno y, como cuarto sustituyente, un grupo adicional denominado cadena lateral, que es el que diferencia a cada aminoácido de los demás. La mayoría de las proteínas están construidas a partir del mismo conjunto de 20 aminoácidos. Cada aminoácido presenta una cadena lateral diferente (Fig. 1.1) que determina sus propiedades físico-químicas particulares, de modo que las células pueden producir proteínas con propiedades y funciones totalmente diferentes uniendo los mismos 20 aminoácidos en multitud de combinaciones distintas. Los aminoácidos se unen entre sí formando la proteína mediante enlaces peptídicos, que son enlaces amida entre el grupo carboxilo ($-COOH$) de un aminoácido y el grupo amino ($-NH_2$) del siguiente. Una característica muy importante del enlace peptídico es que presenta resonancia (Fig. 1.2). Al formarse el enlace peptídico, el átomo de nitrógeno conserva un par de electrones que puede compartir con el átomo de carbono al que está unido, a condición de que éste rompa uno de sus enlaces con el oxígeno. De esta manera, aparece una forma resonante iónica, con doble enlace $C = N$, que presenta carga positiva en el nitrógeno y negativa en el oxígeno. En realidad, el enlace peptídico no adopta ninguna de estas dos situaciones extremas, sino que es un híbrido resonante de ambas y sus propiedades son intermedias. Así, su distancia de enlace (1.32 \AA) es más

corta que la de un enlace sencillo $C - N$ (1.49 Å), pero algo mayor que la de un enlace doble $C = N$ (1.27 Å), lo que indica que también su fortaleza es intermedia. La consecuencia más importante es que el enlace peptídico presenta un carácter de doble enlace parcial, que impide la libre rotación del enlace $C - N$, por lo que los seis átomos que participan en el enlace peptídico (C_{α} , C , O , N , H , y $C_{\alpha+1}$) están obligados a permanecer en el mismo plano y esto va a determinar muchas de las propiedades conformacionales de las proteínas. Otra característica notable del enlace peptídico es que los enlaces $C = O$ y $N - H$ están polarizados, y como son paralelos y el sentido de sus dipolos es el mismo, esto determina que el enlace peptídico lleve asociado un momento dipolar.

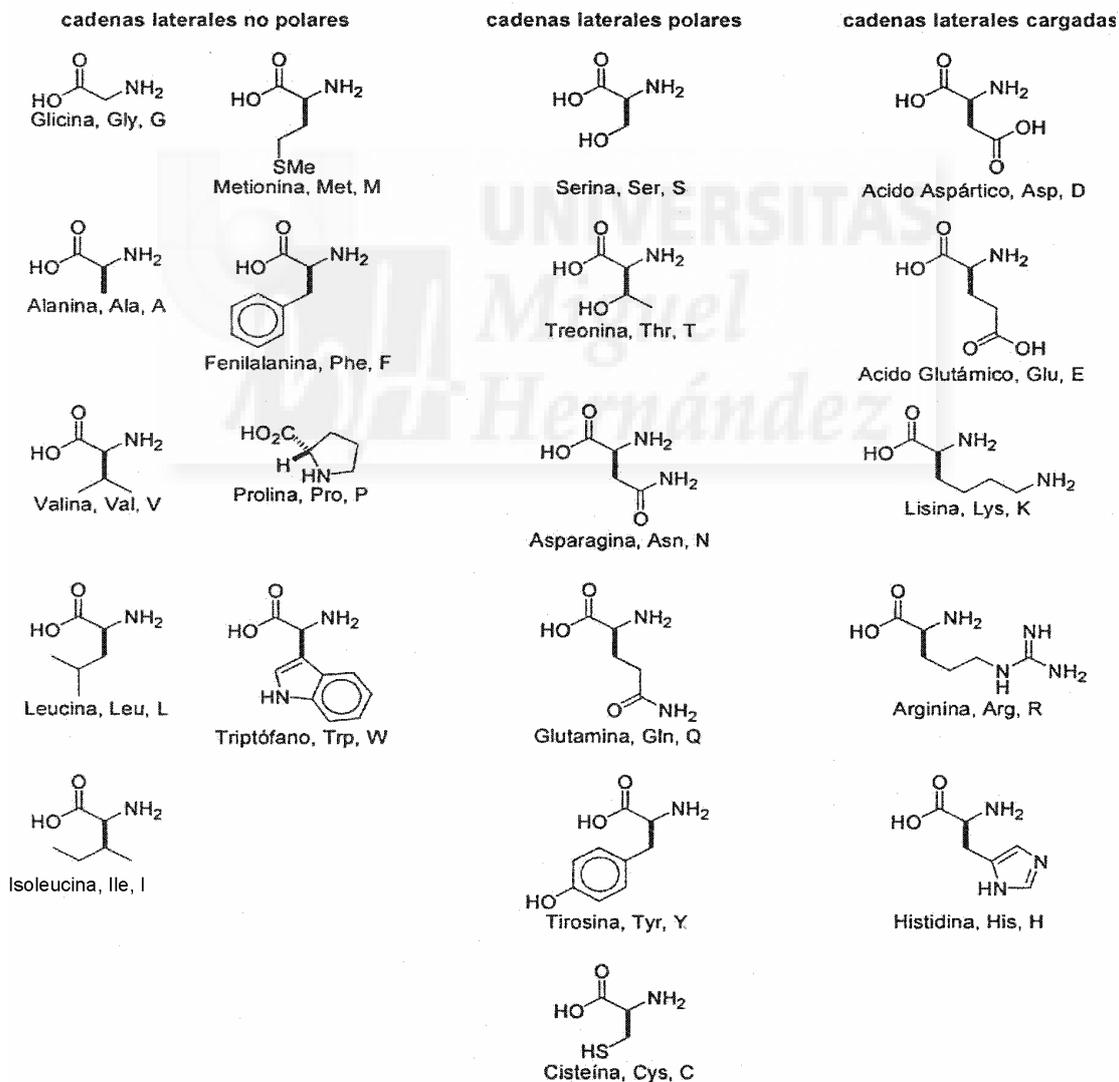


Figura 1.1 Estructura y códigos de una y tres letras para los aminoácidos más comunes, agrupados según la naturaleza de su cadena lateral.

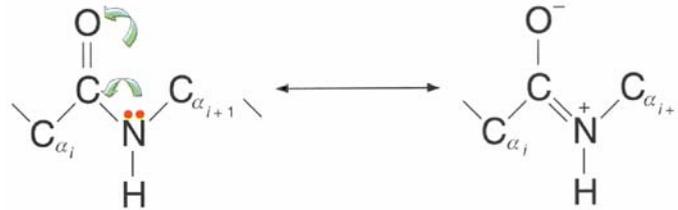


Figura 1.2 Resonancia del enlace peptídico. El enlace peptídico presenta resonancia entre: la forma neutra, con un enlace sencillo uniendo el carbono carbonílico del primer aminoácido (aminoácido i) y el nitrógeno amínico del segundo ($i+1$); y, la forma iónica, con separación de cargas y un enlace doble entre ambos átomos.

La secuencia de aminoácidos de una proteína (definida por la secuencia de nucleótidos del gen que la codifica) contiene toda la información necesaria para que la proteína se pliegue correctamente (Epstein et al., 1963) adoptando su conformación nativa, de modo que entonces puede desempeñar su función biológica. Pero la relación existente entre la secuencia de aminoácidos y la estructura tridimensional de la proteína constituye todavía un problema sin resolver.

1.1.2. LA ESTRUCTURA DE LAS PROTEÍNAS.

A pesar de la enorme complejidad que presenta la estructura de las proteínas, se pueden distinguir cuatro niveles diferentes de organización estructural: la estructura primaria, secundaria, terciaria y cuaternaria, principalmente (Branden y Tooze, 1991).

1.1.2.1. La estructura primaria.

Está definida por la secuencia lineal de los aminoácidos unidos por enlaces peptídicos, así como por la localización de los puentes disulfuro, es decir, comprende todos los enlaces covalentes entre aminoácidos. Es el nivel más sencillo de organización estructural y, en cierto modo, el más importante, puesto que la estructura primaria determina la estructura tridimensional y por tanto la función de la proteína. En este nivel no se especifica la disposición tridimensional de los aminoácidos.

1.1.2.2. La estructura secundaria.

Las cadenas polipeptídicas no son libres de plegarse adoptando cualquier estructura tridimensional al azar, sino que existen impedimentos estéricos y una gran cantidad de interacciones débiles que hacen que unos plegamientos sean más favorables que otros. La estructura secundaria es el conjunto de disposiciones regulares y

repetitivas en el espacio de los residuos que forman la cadena polipeptídica. Los residuos que constituyen estos elementos se enlazan entre sí, fundamentalmente, mediante puentes de hidrógeno entre los átomos de sus enlaces peptídicos. Existen dos tipos principales de estructura secundaria: la hélice α y la lámina β .

La hélice α (Fig. 1.3) es un arrollamiento a derechas de la cadena polipeptídica, en el cual la cadena principal se agrupa de forma muy compacta alrededor del eje longitudinal de la hélice, y las cadenas laterales se proyectan hacia el exterior de la misma, minimizando así cualquier contacto con la propia cadena polipeptídica o con otras cadenas laterales. La estabilidad de la estructura se debe principalmente al óptimo aprovechamiento de los puentes de hidrógeno internos, que se establecen entre el átomo de oxígeno del grupo $-CO$ del residuo i y el átomo de hidrógeno del grupo $-NH$ del residuo $i+4$ (excepto los cuatro aminoácidos de cada extremo de la hélice, que no se enlazan entre sí mediante puentes de hidrógeno intrahelicoidales). Además, la estabilidad de la hélice α también depende de los aminoácidos que la componen y de la posición que éstos ocupan en la hélice, así como de las interacciones que se pueden establecer entre las cadenas laterales de los residuos i e $i+3$ ó i e $i+4$, y de las interacciones de empaquetamiento con el resto de la proteína. La unidad repetitiva es la vuelta de hélice, que ocupa alrededor de 0.54 nm del eje longitudinal y contiene aproximadamente 3.6 aminoácidos (el avance longitudinal por residuo es de 0.15 nm). Una característica a resaltar de las hélices α es su carácter dipolar, ya que los planos de los enlaces peptídicos se disponen de forma casi paralela al eje de la hélice, de modo que sus dipolos quedan alineados, generando así un dipolo global con una carga parcial positiva en el extremo N-terminal y una carga parcial negativa en el extremo C-terminal. Además de las hélices α , un polipéptido puede adoptar otras conformaciones helicoidales, como las denominadas hélices 3_{10} , que se caracterizan porque establecen puentes de hidrógeno entre los aminoácidos i e $i+3$; esta modificación de la α -hélice aparece en ocasiones en el extremo carboxilo de las hélices α . Otro tipo de estructura helicoidal es la hélice π , con puentes de hidrógeno entre los aminoácidos i e $i+5$, que se observa en muy raras ocasiones y sólo dentro de las hélices α más largas.

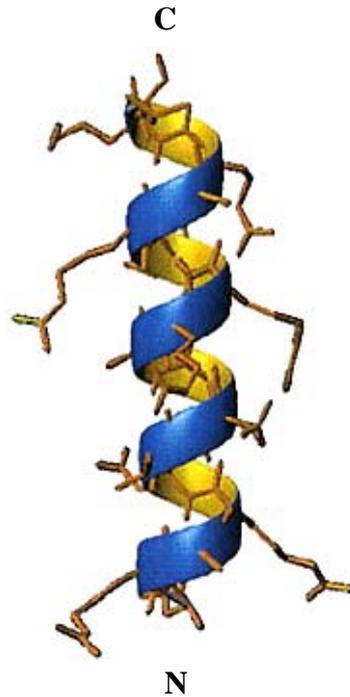


Figura 1.3 Estructura de una hélice α . Diagrama de cintas de la hélice α . Las cadenas laterales se proyectan hacia el exterior de la misma, inclinadas ligeramente hacia el extremo amino de la hélice.

La lámina β (Fig. 1.4) es una conformación extendida en zig-zag de la cadena polipeptídica. Esta estructura se forma a partir de regiones no continuas en la secuencia de la cadena polipeptídica, denominadas hebras β . Cada residuo avanza 0.35 nm a lo largo del eje de la hebra. Estas hebras β se disponen alineadas una al lado de otra, de modo que se establecen puentes de hidrógeno entre los grupos $-\text{CO}$ de una hebra y los grupos $-\text{NH}$ de la otra, formando así las láminas β . Todos los grupos $-\text{CO}$ y $-\text{NH}$ de una hebra β forman puentes de hidrógeno con los de las hebras adyacentes, salvo los que definen los bordes izquierdo y derecho de la lámina. En esta conformación, las cadenas laterales de los residuos consecutivos de las hebras se disponen alternativamente por encima o por debajo del plano en zig-zag de la lámina plegada. La estabilidad de la estructura depende de los aminoácidos que la componen, de las interacciones que se establecen entre las cadenas laterales de hebras adyacentes, y del empaquetamiento de la lámina con el resto de la proteína de la que forma parte. Las láminas β pueden ser de tres tipos: paralelas, cuando todas las hebras que la componen se disponen en el mismo sentido ($\text{N} \rightarrow \text{C}$, por ejemplo); antiparalelas, cuando las hebras contiguas se disponen

alternando los dos sentidos posibles ($N \rightarrow C$ y $C \rightarrow N$); y mixtas, que contienen algunas hebras dispuestas en paralelo y otras antiparalelamente.

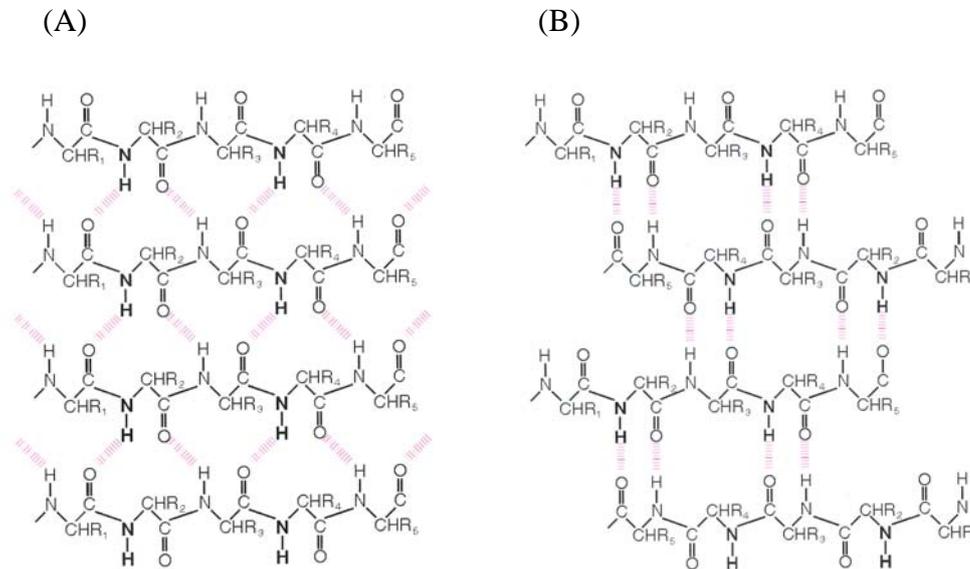


Figura 1.4 Esquema de la estructura de las láminas β . (A) Lámina β paralela. (B) Lámina β antiparalela. En ambas se muestran los puentes de hidrógeno establecidos entre las hebras β , por medio de líneas de puntos.

La conexión entre distintos elementos de estructura secundaria se establece mediante regiones en las que la cadena polipeptídica cambia de dirección; estas regiones reciben el nombre genérico de bucles o lazos, y son necesarios para que la proteína pueda adquirir su particular estructura tridimensional. Los bucles pueden ser de distinta longitud y suelen aparecer en la superficie de las proteínas, garantizando así la buena solvatación de los grupos peptídicos. Los bucles más cortos se denominan giros, y destacan los giros reversos, que están formados por tan sólo cuatro residuos: los residuos i e $i+3$ enlazados por puente de hidrógeno entre sus grupos $-CO$ y $-NH$, respectivamente; y, los residuos $i+1$ e $i+2$ se encuentran sin enlazar. Otro tipo importante son los giros β , que conectan con frecuencia hebras β de láminas antiparalelas. En los giros abundan los residuos de glicina, dado que es un residuo pequeño que puede adoptar multitud de conformaciones estéricamente prohibidas para otros aminoácidos.

Se puede considerar la existencia de un nivel estructural adicional, a medio camino entre la estructura secundaria y la terciaria, que se conoce como estructura

supersecundaria. La estructura supersecundaria es un agrupamiento estable de varios elementos de estructura secundaria, y los más frecuentes son: el motivo α -giro- α , compuesto por dos hélices α unidas por un giro; la horquilla β , formada por dos hebras β unidas por un giro β ; el motivo $\beta\alpha\beta$, formado por dos hebras β paralelas conectadas entre sí a través de una hélice α ; y el meandro β , compuesto por cinco hebras β unidas entre sí por giros β .

1.1.2.3. La estructura terciaria.

Es la descripción de las relaciones espaciales existentes entre todos los aminoácidos de la cadena polipeptídica, y por lo tanto, define la disposición tridimensional de todos los átomos que forman la proteína. Cuando la proteína se pliega adquiriendo su estructura terciaria, los aminoácidos que están alejados en la secuencia pueden interactuar y quedar próximos.

Al plegarse la cadena polipeptídica puede dar lugar a uno o varios dominios independientes. Normalmente, los dominios son unidades compactas, con una estructura tridimensional característica, de plegamiento autónomo, y muchas veces desempeñan funciones específicas. Por ello, son las unidades fundamentales de estructura y función de las proteínas. La estructura terciaria de proteínas con elevada masa molecular suele estar compuesta por varios dominios.

A pesar de que cada proteína presenta una estructura terciaria característica, algunos tipos de plegamiento terciario suelen aparecer de forma repetida. De acuerdo con esto, las proteínas se pueden clasificar según su estructura tridimensional en cuatro grupos principales: proteínas α , cuyo armazón está formado exclusivamente por hélices α (CA-C, la proteína objeto de este trabajo, es una proteína α); proteínas β , constituidas por hebras β en orientación preferentemente antiparalela; proteínas α/β , formadas por la alternancia regular de hélices α y hebras β a lo largo de la secuencia; y proteínas $\alpha+\beta$, caracterizadas por la alternancia de ambos tipos de estructura secundaria de forma no secuencial.

1.1.2.4. La estructura cuaternaria.

Se trata de un nivel adicional de estructura presente sólo en aquellas proteínas formadas por varias cadenas polipeptídicas. La estructura cuaternaria se refiere a la disposición espacial de las distintas cadenas polipeptídicas que forman la proteína y que se unen entre sí mediante interacciones no covalentes. En estas proteínas, cada cadena

polipeptídica constituye una subunidad o monómero, y el conjunto forma un oligómero. CA-C, la proteína objeto de este trabajo, es un dímero.

1.2. EL PLEGAMIENTO DE LAS PROTEÍNAS: RECONOCIMIENTO INTRAMOLECULAR.

1.2.1. EL PROBLEMA DEL PLEGAMIENTO.

Una vez que la proteína ha sido sintetizada en el ribosoma, su cadena polipeptídica tiene que plegarse correctamente para adquirir su estructura tridimensional nativa característica y poder así ejercer su función biológica. El plegamiento constituye un importante proceso de reconocimiento intramolecular en la proteína, ya que conlleva el establecimiento de multitud de interacciones no covalentes entre los aminoácidos que la forman. La mayoría de las proteínas se pliegan espontáneamente y de forma autónoma, sin embargo, todavía no se conoce cómo ocurre este proceso exactamente.

La información básica necesaria para que una proteína se pliegue correctamente, adoptando su estructura tridimensional nativa o biológicamente activa, está contenida en su secuencia de aminoácidos (Epstein et al., 1963). Este principio fundamental se demuestra por el hecho de que muchas proteínas sencillas que han sido previamente desnaturalizadas pueden replegarse por sí mismas, de forma reversible, a su conformación nativa, si son devueltas a las condiciones iniciales (Anfinsen, 1973). No obstante, el mecanismo por el cual una cadena polipeptídica adopta, de entre todas las conformaciones posibles, aquella que le permite realizar la función para la que fue sintetizada se desconoce, y constituye el llamado *problema del plegamiento* de las proteínas, que es actualmente uno de los problemas centrales de estudio en Bioquímica.

La resolución del problema del plegamiento requiere la descripción cinética y termodinámica. El estudio cinético del plegamiento es muy difícil de abordar debido fundamentalmente a la rapidez con que ocurre el proceso, y además, a las limitaciones de las diferentes aproximaciones experimentales y teóricas disponibles. En cambio, el estudio termodinámico es habitualmente más sencillo, describiéndose los parámetros termodinámicos que controlan la reacción de plegamiento. La resolución del problema del plegamiento en su totalidad, presenta un enorme interés, no sólo desde el punto de vista del conocimiento básico, sino también por la gran cantidad de aplicaciones prácticas que abriría, como la predicción de la estructura tridimensional de cualquier

proteína a partir de su secuencia, o el diseño racional de proteínas con interés biomédico o industrial.

1.2.2. ESTADOS CONFORMACIONALES EN LA REACCIÓN DE PLEGAMIENTO.

La resolución completa del problema del plegamiento de una proteína implica conocer la estructura y la energética de todos los estados conformacionales implicados en la reacción de plegamiento: el estado inicial (estado desnaturalizado), el estado de transición, y el estado final (estado nativo o correctamente plegado). Además, bajo ciertas condiciones, algunas proteínas pueden presentar otros estados conformacionales intermedios (intermediarios de plegamiento). Los estados inicial y final del proceso de plegamiento son estables en equilibrio y se pueden caracterizar en mayor o menor detalle. Sin embargo, los intermediarios de plegamiento suelen estar asociados con procesos cinéticos y no siempre se pueden detectar: serán más fácilmente observados cuanto mayor sea su estabilidad relativa con respecto al estado precedente en la reacción considerada. La caracterización de los estados de transición implica medidas cinéticas (Fersht, 1999).

1.2.2.1. El estado desnaturalizado.

El estado conformacional de una proteína desnaturalizada es, idealmente, un ovillo estadístico (*random coil*) que carece de cualquier tipo de estructura regular (Dill y Shortle, 1991). La definición de ovillo estadístico, referida a cadenas de homopolímeros, es un conjunto de estructuras flexibles en continuo movimiento, con giros y torsiones aleatorios en diferentes puntos de la cadena polipeptídica. En esta conformación, los residuos no se encuentran totalmente expuestos al disolvente, sino que parte del tiempo interactúan al azar con otros residuos pero sin llegar a establecer interacciones estables entre ellos. El estado desnaturalizado presenta una gran entropía conformacional estabilizante debido a su capacidad de adoptar multitud de conformaciones diferentes con energías muy similares.

Cuando las proteínas se desnaturalizan, sus cadenas polipeptídicas no se comportan como ovillos estadísticos de forma estricta, ya que en realidad, las proteínas son heteropolímeros en los cuales unos aminoácidos interactúan preferentemente con otros, de modo que algunas conformaciones son más frecuentes que otras. Así pues, la cadena polipeptídica tiende a ser menos desordenada y más compacta de lo que se espera para un ovillo estadístico propiamente dicho. Desde el punto de vista termodinámico, los distintos estados desplegados de una proteína son indistinguibles

entre sí, por lo que se puede considerar que, a pesar de que el estado desnaturalizado presenta una conformación fluctuante, se trata de un único estado macroscópico (Pfeil y Privalov, 1976; Privalov et al., 1989; Privalov, 1992).

1.2.2.2. El estado nativo.

Es la conformación en la que la proteína se encuentra completamente plegada y por lo tanto, puede desempeñar su función biológica. En este estado, la proteína presenta una estructura global bien definida, ordenada y relativamente rígida. Se suele caracterizar por un elevado grado de compacidad, y por la existencia de un núcleo hidrófobo y una superficie hidrófila, originando una conformación única bien definida (o más bien un conjunto de conformaciones muy similares entre sí, estructural y energéticamente), con interacciones entre sus residuos aminoacídicos. La reacción de plegamiento implica la adquisición de todas estas características a partir de las propias del estado desnaturalizado.

El estado nativo es normalmente el más fácil de estudiar, tanto a nivel de su estructura como de su estabilidad. La estructura de la conformación nativa se puede determinar a nivel atómico por cristalografía de rayos X y por RMN; y su estabilidad se puede estudiar mediante una combinación de ingeniería de proteínas (Serrano et al., 1990), calorimetría (Privalov, 1992), y estudios en equilibrio (Pace y Scholtz, 1997).

1.2.2.3. Intermediarios de plegamiento: el glóbulo fundido.

La reacción de plegamiento de muchas proteínas es una transición simple, en la que tan sólo participan dos estados significativamente poblados: el estado nativo y el estado desnaturalizado; es lo que se conoce como *proceso de dos estados*. Sin embargo, las proteínas más complejas (habitualmente de mayor tamaño y formadas por varios dominios) suelen plegarse mediante un *proceso de tres o más estados*, que implica la presencia de uno o más intermediarios.

Estos intermediarios de plegamiento son conformaciones parcialmente plegadas, cuyas propiedades son intermedias entre las del estado nativo y las del desnaturalizado, y si son lo suficientemente estables como para poder ser detectados en equilibrio, se podrán estudiar con relativa facilidad y de forma aislada. Muchos de estos intermediarios presentan propiedades similares en proteínas diferentes y han sido denominados glóbulos fundidos (*molten globules*) en equilibrio (Ptitsyn, 1995). Las características fundamentales de los glóbulos fundidos son:

(A) Presencia de casi la misma compacidad que la forma nativa, con un tamaño sólo entre un 10 % y un 30 % más grande.

(B) Presencia de un elevado contenido de estructura secundaria.

(C) Ausencia de la mayor parte de la estructura terciaria específica, producida por el empaquetamiento de las cadenas laterales de los aminoácidos.

(D) Ausencia práctica de regiones hidrofóbicas bien empaquetadas, lo que hace que mucha de la superficie hidrófoba que debería estar enterrada en el interior de la proteína, permanezca accesible al disolvente.

Todavía existen muchas cuestiones relacionadas con los intermediarios de plegamiento que no tienen una respuesta clara, como el hecho de si se pueden considerar, desde el punto de vista físico-químico, un verdadero estado macroscópico de la macromolécula, cuyas propiedades termodinámicas puedan ser estudiadas; o si existen o no características estructurales comunes a todos los intermediarios. Además, también existe bastante controversia con respecto al papel que desempeñan en el plegamiento de las proteínas, es decir, hay opiniones dispares sobre si son o no indispensables para que la proteína se pliegue. Por lo tanto, es necesario continuar con estudios de proteínas muy diferentes para poder definir unas características generales en cuanto a la estructura y estabilidad de estos intermediarios, así como para el establecimiento de su función en el plegamiento de las proteínas.

1.2.3. PLEGAMIENTO DE PROTEÍNAS CON VARIOS DOMINIOS Y DE PROTEÍNAS OLIGOMÉRICAS.

El plegamiento de proteínas multidominio y oligoméricas ha sido menos estudiado que el de proteínas sencillas. De modo muy simplificado, puede decirse que el plegamiento de proteínas multidominio procede, en general, a través del plegamiento individual de los dominios, seguido de la interacción entre dominios y el posterior reordenamiento de la estructura.

El plegamiento de proteínas oligoméricas suele implicar primero el plegamiento total o parcial de cada monómero, seguido de su asociación y reordenamiento final de la estructura para dar lugar al oligómero nativo (Jaenicke y Lillie, 2000). En ciertas proteínas pequeñas o en dominios de oligomerización, el plegamiento y la asociación pueden estar acoplados en un proceso de dos estados (monómero disociado-oligómero plegado), sin intermediarios suficientemente poblados en el equilibrio. Sin embargo, esta ruta puede complicarse y seguir un proceso de tres o más estados, apareciendo uno o más intermediarios de plegamiento estables.

1.2.4. ESTUDIO DEL PLEGAMIENTO: TÉCNICAS EXPERIMENTALES.

El estudio del plegamiento de las proteínas desde el punto de vista termodinámico consiste en el análisis del equilibrio entre el estado desnaturalizado y los estados parcial o totalmente plegados. El objetivo es caracterizar con la mayor precisión posible cada uno de los estados conformaciones implicados en la reacción de plegamiento. Así pues, desde el punto de vista termodinámico, el estudio experimental del plegamiento se lleva a cabo induciendo la desnaturalización de las proteínas y estudiando mediante diversas técnicas biofísicas los cambios conformacionales ocurridos en este proceso de desnaturalización.

La desnaturalización de las proteínas se puede inducir alterando su entorno físico o químico. Los métodos más empleados son el cambio de pH, el cambio de temperatura, y la adición de agentes desnaturalizantes, que solubilizan las cadenas hidrófobas de las proteínas estabilizando el estado desnaturalizado. El cambio en las condiciones de la disolución desencadenará la reacción de desplegamiento. En proteínas oligoméricas, el equilibrio de desnaturalización depende de la concentración de proteína; de ahí la necesidad de realizar estos estudios de desplegamiento a distintas concentraciones.

Para llevar a cabo el estudio del proceso de desnaturalización de las proteínas, es necesaria la utilización de técnicas que proporcionen información acerca de la conformación de las distintas especies implicadas en dicho proceso. Se trata de técnicas que permiten medir propiedades físicas que dependen de la conformación de la proteína, es decir, propiedades que cambian de valor con la desnaturalización. Las técnicas espectroscópicas más utilizadas son la absorbancia, la fluorescencia, el DC, el FTIR y la RMN; además, también pueden utilizarse técnicas hidrodinámicas, como la cromatografía de exclusión por tamaño molecular (ver la descripción de las distintas técnicas biofísicas en el apartado 1.4.).

1.3. UNIÓN A MACROMOLÉCULAS: RECONOCIMIENTO INTERMOLECULAR.

1.3.1. LA INTERACCIÓN ENTRE MACROMOLÉCULAS.

Los fenómenos de reconocimiento intermolecular entre macromoléculas ocurren en la mayor parte de los procesos biológicos. Las proteínas llevan a cabo su función biológica interaccionando, de forma temporal o permanente, con ácidos nucleicos (en los procesos de replicación, transcripción y traducción), lípidos (proteínas de

membrana), glúcidos (como las lectinas), u otras proteínas. Particularmente, la interacción proteína-proteína se produce en muchos procesos bioquímicos, como reacciones enzimáticas o fenómenos de reconocimiento molecular específicos, tales como la interacción hormona-receptor o antígeno-anticuerpo (Wells, 1991; Schreiber y Fersht, 1995; Jones y Thornton, 1996; Xu et al., 1997). Además, la interacción proteína-proteína es responsable de la formación del gran número de proteínas oligoméricas existentes, así como del ensamblaje de numerosos complejos supramoleculares, como por ejemplo, la cápsida viral.

En general, el estudio de la interacción entre macromoléculas nos ayudará a:

- (i) comprender el proceso biológico en que tal interacción está implicada;
- (ii) determinar las anomalías que impiden el establecimiento óptimo de las interacciones y que provocan ciertas patologías importantes, como la enfermedad de Alzheimer o encefalopatías espongiiformes, entre otras;
- (iii) diseñar agentes terapéuticos con la capacidad de corregir dichas anomalías;
- (iv) desarrollar técnicas computacionales capaces de predecir la composición y características de las superficies de interacción entre moléculas de forma específica; y,
- (v) las interacciones proteína-proteína no son más que un caso de reconocimiento intermolecular, donde las interacciones implicadas son las mismas que en el plegamiento (pero aquí intramolecularmente), por tanto, un mayor conocimiento de dicho proceso nos puede ayudar a mejorar la comprensión del problema del plegamiento de las proteínas.

1.3.2. ESTUDIO DE LA INTERACCIÓN CON MACROMOLÉCULAS.

Desde un punto de vista biomédico, el estudio de la interacción proteína-proteína tiene una importante aplicación, ya que el reconocimiento productivo entre proteínas puede inhibirse mediante una distorsión o bloqueo de la interfase de interacción por acción de un compuesto exógeno. Una posibilidad atractiva para llevar a cabo este bloqueo consiste en la utilización de una forma biológicamente inactiva de una de las proteínas participantes (Babé et al., 1991; Zhang y Rollins, 1995), o de un péptido o una molécula cuya secuencia se corresponda con la de la interfase de interacción (Zhang et al., 1991; Divita et al., 1995; Schramm et al., 1996; Maroun et al., 2001), de forma que, el péptido o molécula presenta la capacidad potencial de unirse a la proteína impidiendo así una interacción proteína-proteína, es decir, puede actuar como inhibidor. Por lo tanto, el estudio de la interacción péptido-proteína o molécula-proteína presenta un enorme interés, ya que los conocimientos adquiridos de dicho estudio pueden utilizarse para el posterior diseño de fármacos.

La estrategia a seguir para estudiar la interacción molécula-proteína consta de dos etapas fundamentalmente. En primer lugar, debe identificarse la existencia de dicha interacción, y entre qué residuos se produce; para ello se pueden utilizar diferentes técnicas biofísicas que permiten detectar interacciones intermoleculares. Y en segundo lugar, una vez detectada la existencia de la interacción molécula-proteína, es necesario caracterizar la fortaleza de la misma mediante el empleo de diversas técnicas biofísicas que permiten la determinación de la constante de unión del proceso (ver apartado 1.4.).

1.4. TÉCNICAS BIOFÍSICAS.

1.4.1. TÉCNICAS UTILIZADAS PARA DETERMINAR LA ESTRUCTURA DE LAS PROTEÍNAS Y ESTUDIAR EL PLEGAMIENTO Y LA UNIÓN A MACROMOLÉCULAS.

Aunque actualmente existen muchas técnicas para caracterizar la estructura, el plegamiento y la unión macromolecular, aquí sólo nos limitaremos a describir aquellas utilizadas en este trabajo.

1.4.1.1. Técnicas espectroscópicas.

1.4.1.1.1. Absorbancia.

La práctica totalidad de las proteínas absorben luz en la región del ultravioleta cercano (250-350 nm) debido a las cadenas laterales de los residuos aromáticos que contienen: triptófano, tirosina y fenilalanina. De estos, la contribución más importante a la absorbancia de la proteína suele ser la del triptófano, debido a que posee el mayor coeficiente de extinción, seguida por la tirosina, y en menor medida la de la fenilalanina. Sin embargo, dado que el contenido en tirosinas de las proteínas es, en general, mayor que el de triptófanos, ambos aminoácidos suelen contribuir de forma parecida al espectro de absorción de las proteínas. Otra diferencia a tener en cuenta es que el triptófano absorbe luz de longitudes de onda ligeramente mayores (presenta un máximo de absorción centrado a 280 nm) que la tirosina (su máximo se centra a 275 nm) y la fenilalanina (en torno a 257 nm). Como los aminoácidos aromáticos no absorben por encima de 310 nm, la presencia de una línea de base descendente en la región de 310 a 400 nm es indicativa de que la proteína está agregada, ya que tal línea de base se origina como consecuencia de la dispersión de la luz por parte de partículas de gran tamaño. La absorción de los aminoácidos aromáticos también depende del entorno molecular en el que se encuentran, de modo que, una variación en el entorno

provoca cambios tanto en la anchura e intensidad de las bandas de absorbancia como en la longitud de onda de máxima absorción.

1.4.1.1.2. Fluorescencia.

El fenómeno de la fluorescencia consiste en la emisión rápida de luz de mayor longitud de onda por parte de una molécula que la ha absorbido previamente a una longitud de onda menor (de mayor energía). Las proteínas presentan la propiedad de la fluorescencia, aunque no contengan grupos prostéticos, puesto que los residuos aromáticos (triptófano y tirosina) son fluorescentes. La emisión de fluorescencia de las proteínas suele estar dominada por los residuos de triptófano, ya que tanto su coeficiente de extinción a la longitud de onda habitual de excitación como su rendimiento cuántico intrínseco son considerablemente mayores que los de la tirosina; además, los fotones absorbidos por las tirosinas son frecuentemente transmitidos a los triptófanos a través de un fenómeno cuántico denominado transferencia de energía por resonancia. Por todo ello, los triptófanos son los fluoróforos por excelencia de las proteínas (Tatischeff y Klein, 1975; Chen y Barkley, 1998; Lakowicz, 1999).

La espectroscopía de fluorescencia proporciona información acerca de la estructura terciaria de las proteínas alrededor de sus residuos aromáticos fluorescentes, y permite el estudio del plegamiento de las mismas, ya que la emisión de fluorescencia es muy sensible al entorno de los fluoróforos (más que la absorbancia). Al desnaturalizarse una proteína, ésta pierde su estructura terciaria, de forma que el entorno de los fluoróforos varía aumentando su polaridad, y esto a su vez provoca cambios en las características de la emisión fluorescente. En las proteínas que contienen triptófanos, los cambios producidos en la intensidad de la emisión de fluorescencia son importantes, y pueden consistir en un aumento o una disminución de la intensidad; el sentido de la variación va a depender de la proteína, dado que la emisión del triptófano en la proteína plegada puede ser mayor o menor que la del triptófano libre en disolución acuosa. Por el contrario, siempre se produce un desplazamiento del máximo de emisión desde longitudes de onda menores (entre 330 y 350 nm, cuando los triptófanos se encuentran enterrados en la estructura de la proteína) hacia mayores longitudes de onda (350-355 nm, cuando todos los triptófanos están completamente expuestos al disolvente). Sin embargo, en el caso de las proteínas que no contienen triptófanos pero sí residuos de tirosina, los cambios en la fluorescencia con la variación del entorno son principalmente cambios en la intensidad de emisión (observándose normalmente un aumento de la intensidad al desplegarse la proteína), mientras que el máximo de emisión

se mantiene alrededor de 303 nm, independientemente del entorno molecular de los fluoróforos (Eftink, 1994; Royer, 1995; Pace y Scholtz, 1997; Schmid, 1997; Lakowicz, 1999).

Para estudiar los cambios conformacionales también se pueden utilizar otras técnicas adicionales basadas en el fenómeno de la fluorescencia, como son la unión a sondas (ANS), el *quenching* de fluorescencia y la anisotropía.

(A) La unión a ANS permite detectar conformaciones proteicas parcialmente plegadas, puesto que el ANS es una sonda fluorescente que se une a regiones hidrofóbicas de las proteínas próximas entre sí que quedan expuestas al entorno con la desnaturalización (Goldberg et al., 1990; Semisotnov et al., 1991; Mann y Matthews, 1993). Cuando el ANS se une a regiones hidrofóbicas expuestas al disolvente, su intensidad de fluorescencia aumenta y se produce un desplazamiento de su máximo de emisión desde 520 nm (cuando el ANS está totalmente libre) hasta 480 nm (cuando el ANS se encuentra completamente unido) (Stryer, 1965; Turner y Brand, 1968).

(B) El *quenching* de fluorescencia consiste en la disminución de la intensidad de la emisión fluorescente de los fluoróforos a consecuencia de su interacción con una molécula ajena denominada *quencher* o desactivador. El *quenching* puede ser de dos tipos, dependiendo de cómo se produzca la interacción entre el fluoróforo y el *quencher*: dinámico, cuando el desactivador colisiona con el fluoróforo durante el tiempo de vida media del estado excitado, de forma que no se producen alteraciones químicas en las moléculas; o estático, si el desactivador se asocia al fluoróforo formando un complejo estable que impide la emisión de fluorescencia (Lakowicz, 1999). De esta manera, la intensidad del fenómeno de *quenching* será mayor cuanto mayor sea la accesibilidad de los fluoróforos a las moléculas de *quencher*. En consecuencia, el *quenching* de fluorescencia permite determinar la mayor o menor accesibilidad de los fluoróforos proteicos, que a su vez depende de la conformación en que se encuentra la proteína (Colucci et al., 1990).

(C) La anisotropía se basa en el principio de fotoselección, que consiste en que cuando una molécula se irradia con luz polarizada plana, sólo aquellas moléculas convenientemente orientadas la absorberán. Del análisis del plano de oscilación de la emisión fluorescente se puede determinar el grado de libertad en el movimiento del fluoróforo, ya que si la molécula se encuentra libre en disolución, los fotones emitidos tendrán planos de oscilación distintos de aquél de la luz incidente, tanto más distintos cuanto mayor sea su libertad de movimiento. Por lo tanto, la anisotropía de

fluorescencia proporciona información sobre la libertad de movimiento del fluoróforo, que depende fundamentalmente de su entorno, y éste a su vez de la conformación en que la proteína se encuentra (Lakowicz, 1999).

Adicionalmente, la espectroscopía de fluorescencia permite el estudio de los procesos de unión entre biomoléculas, tales como la asociación proteína-proteína o péptido-proteína, ya que la emisión de fluorescencia de los fluoróforos proteicos, principalmente la del triptófano, depende fuertemente de su entorno local, y éste cambia cuando se produce la interacción (bien porque los fluoróforos se localizan en el sitio de unión entre ambas moléculas, o bien porque aún situándose fuera del sitio de unión, la interacción provoca un cambio conformacional global que altera el microentorno de los fluoróforos). Al producirse la interacción, el espectro de fluorescencia de las moléculas implicadas puede cambiar, originándose variaciones en la longitud de onda de máxima emisión y/o en la intensidad de fluorescencia, que pueden utilizarse como observables para el seguimiento del proceso de unión. Además, para este tipo de estudios también se pueden aplicar otras técnicas basadas en la fluorescencia, como el *quenching* o la anisotropía de fluorescencia, midiendo respectivamente los cambios en la accesibilidad o en la libertad de movimiento de los fluoróforos que se generan a consecuencia de la unión (Eftink, 1997; Jameson y Seifried, 1999).

En ocasiones, para evitar extraer conclusiones erróneas, o bien para obtener confirmación experimental adicional, resulta necesario marcar a una de las moléculas implicadas en la interacción (usualmente suele marcarse la molécula más pequeña) con una sonda fluorescente cuyas propiedades de fluorescencia se vean afectadas por la unión. Esto suele realizarse cuando ambas moléculas que interaccionan poseen los mismos fluoróforos, ya que es muy difícil asignar los cambios en la fluorescencia a los residuos particulares. Por otra parte, la probabilidad de detectar un cambio como consecuencia de la interacción en el espectro de fluorescencia, decrece con el número total de residuos de triptófano, de modo que, el uso de sondas permite aumentar la sensibilidad en el seguimiento del proceso de unión.

1.4.1.1.3. Dicroísmo circular.

La espectroscopía de DC mide la absorción diferencial de la luz polarizada circularmente a la derecha y a la izquierda, por parte de las moléculas que presentan actividad óptica. Las señales de DC se observan en las mismas regiones espectrales en las que aparecen bandas de absorción, siempre que los cromóforos o su entorno molecular sean asimétricos. En las proteínas existen dos regiones principales donde se

producen bandas de DC, en el ultravioleta lejano y en el ultravioleta cercano, y cada una proporciona distinta información sobre la estructura de las proteínas (Freskgard et al., 1994; Pace y Scholtz, 1997; Kelly y Price, 2000).

En la región del ultravioleta lejano (170-250 nm) el principal cromóforo es el enlace peptídico. El enlace peptídico *per se* no es una estructura quiral, pero es la interacción entre grupos peptídicos adyacentes dispuestos en determinadas orientaciones la que determina que las proteínas presenten actividad óptica en esta región. En esta región del espectro se obtiene información sobre el contenido de estructura secundaria de la proteína, ya que los enlaces peptídicos dan señales distintas según estén dispuestos en hélices α , en láminas β o en segmentos no periódicos de la estructura (estructura desordenada) (Fig. 1.5).

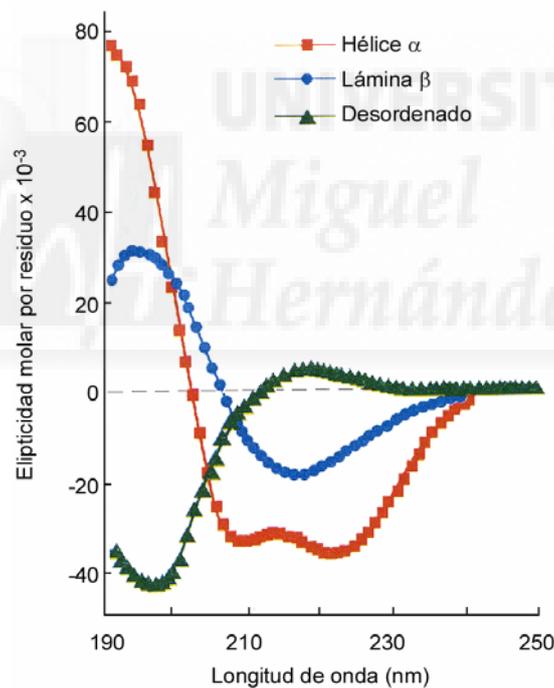


Figura 1.5 Espectros de DC en el ultravioleta lejano característicos de los diferentes tipos de estructura secundaria de las proteínas.

La señal más intensa entre 200 y 230 nm la dan las hélices α . Cada una de las tres estructuras mencionadas contribuye al espectro con la señal que le es característica, y una intensidad proporcional al número de residuos de la proteína que la adoptan. Cada proteína presenta un espectro que es, aproximadamente, una combinación lineal de los

espectros característicos de las distintas estructuras secundarias que contiene, por lo que es posible conocer, cualitativamente, qué tipo de estructura es mayoritaria en una proteína a partir de su espectro en esta región. Sin embargo, hay que tener en cuenta que los residuos aromáticos también pueden contribuir al espectro de DC en la región del ultravioleta lejano, cuando sus cadenas laterales, que no son quirales, se comportan como tales si así lo propicia su entorno proteico. Con esta salvedad, la actividad óptica de los enlaces peptídicos es de gran utilidad para conocer el contenido de estructura secundaria de las proteínas.

En la región del ultravioleta cercano (250-350 nm) las bandas de DC se deben exclusivamente a los residuos aromáticos que, atrapados en un entorno bastante fijo, adquieren asimetría. Así pues, la aparición de bandas de DC en el ultravioleta cercano refleja la existencia de residuos aromáticos inmovilizados en entornos asimétricos, de modo que, en esta región del espectro se obtiene información sobre la estructura terciaria alrededor de esos residuos.

La técnica de DC también puede utilizarse para el estudio del proceso de desnaturalización, tanto en la región del ultravioleta lejano como en el cercano. Cuando una proteína se desnaturaliza, su contenido en estructura secundaria suele disminuir, y esto puede monitorizarse haciendo medidas de DC en el ultravioleta lejano. Por otro lado, con la desnaturalización, la proteína pierde el entorno asimétrico de sus residuos aromáticos que suele ser típico de la conformación nativa (Schmid, 1997), y esto se refleja en el desvanecimiento de la señal de DC en el ultravioleta cercano. Así pues, las diferencias en el espectro de DC entre la conformación nativa y la desnaturalizada son muy grandes en ambas regiones espectrales, lo que hacen del DC una de las técnicas más usadas para los estudios de plegamiento de las proteínas (Fasman, 1990; Chiti et al., 1998 a, b)

Asimismo, la espectroscopía de DC es uno de los métodos más adecuados para el estudio de los procesos de unión entre biomoléculas (Blauer, 1974; Greenfield, 1975), ya que la interacción provoca cambios en la estructura secundaria y/o terciaria de las moléculas implicadas en la unión, que pueden ser estudiados por DC en el ultravioleta lejano o en el cercano, respectivamente. Otra forma de detectar la existencia de interacción por DC consiste en medir la estabilidad térmica, de modo que si se produce interacción, se observa una variación en la estabilidad del complejo formado con respecto a la estabilidad de sus componentes por separado.

1.4.1.1.4. Infrarrojo por transformada de Fourier.

La espectroscopía de FTIR mide la absorción de radiación infrarroja del enlace peptídico, que depende de la estructura secundaria en la que se encuentra implicado. Una de las principales bandas de absorción de las proteínas en el infrarrojo es la banda amida I, que aparece en la región de $1600\text{-}1700\text{ cm}^{-1}$, y está formada por otras bandas más simples, cada una de ellas con el máximo localizado en un valor distinto de frecuencia para cada tipo de estructura secundaria (Fig. 1.6). Esta banda amida I se puede descomponer matemáticamente en contribuciones características de cada elemento de estructura secundaria, cuya intensidad es proporcional al número de grupos que adoptan cada una de esas estructuras, y esto permite la cuantificación de la estructura secundaria de las proteínas (Surewicz et al., 1993; de Jongh et al., 1996).

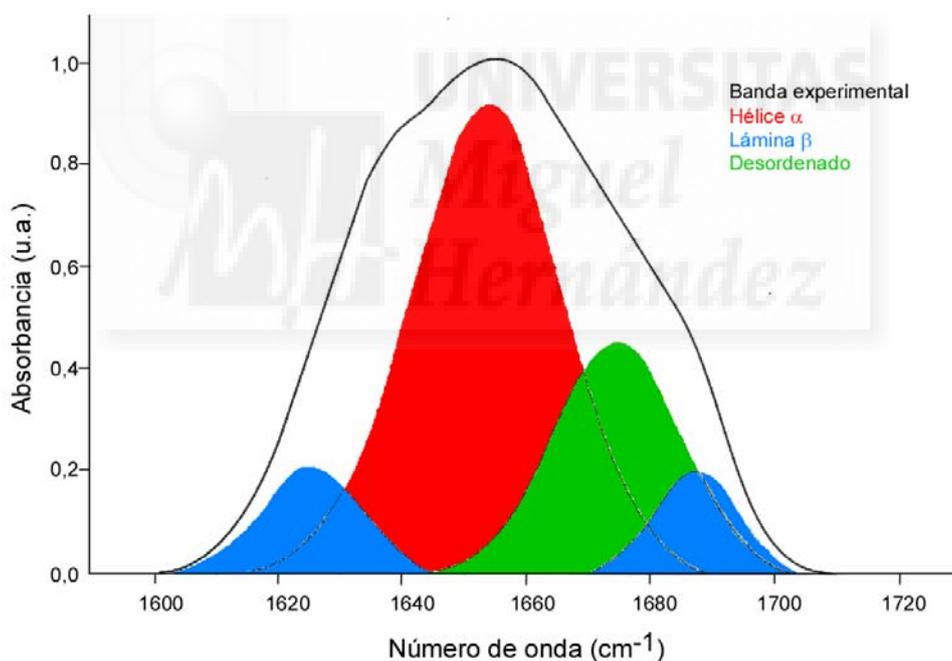


Figura 1.6 Bandas componentes de la amida I.

El FTIR presenta una serie de ventajas frente a otras técnicas para la determinación de la estructura de las proteínas, entre ellas que su uso no está limitado por el tamaño de la molécula a estudiar (como ocurre para la RMN), la turbidez de las muestras no supone ningún problema metodológico (al contrario de lo que sucede en la fluorescencia y el DC), y se pueden obtener rápida y fácilmente espectros de alta

calidad, proporcionando información directa sobre la estructura secundaria del esqueleto proteico. Sin embargo, también posee algunas desventajas, como la necesidad de utilizar grandes cantidades de muestra (mayores que para la fluorescencia y el DC, y del mismo orden que para la RMN), y la necesidad de sustituir el agua por agua deuterada, ya que la absorción del agua en el infrarrojo solapa con la mayoría de zonas que proporcionan información sobre las proteínas, y además, presenta mucha mayor intensidad (Mendelsohn y Mantsch, 1986).

La espectroscopía de FTIR permite, además, el seguimiento de la desnaturalización, ya que la radiación infrarroja es absorbida por los enlaces peptídicos de forma característica para cada tipo de estructura secundaria, de modo que al desnaturalizarse la proteína se alteran sus proporciones y, en consecuencia, el espectro de infrarrojo. Por otra parte, puesto que las cadenas laterales de algunos aminoácidos también absorben en esta región del espectro, se pueden seguir los cambios que ocurren en su absorbancia del infrarrojo con la desnaturalización.

1.4.1.1.5. Resonancia magnética nuclear.

Mediante la espectroscopía de RMN se determinan las posiciones exactas de cada uno de los átomos que forman la biomolécula. La técnica de RMN se basa en la absorción de ondas de radio por parte de determinados núcleos atómicos, que se produce cuando se cambia la orientación del momento magnético nuclear bajo la aplicación de un campo magnético externo. Los métodos de RMN utilizan las propiedades magnéticas de los núcleos atómicos dentro de una molécula para obtener una serie de restricciones de distancias entre átomos, a partir de las cuales se puede obtener la estructura tridimensional de la molécula proteica a nivel atómico. Así pues, esta técnica proporciona gran cantidad de información sobre la estructura secundaria y terciaria de las proteínas a nivel de residuo; por lo tanto, se trata de una técnica espectroscópica que proporciona, junto con la cristalografía de rayos X, información atómica detallada sobre las proteínas (Wüthrich, 1986).

La principal ventaja de la RMN respecto a la cristalografía de rayos X es que las proteínas se estudian en disolución, que por un lado representa mejor el entorno de las proteínas en el interior de las células, y por otro lado evita las dificultades derivadas de la cristalización, permitiendo además el estudio de procesos cinéticos. Entre las limitaciones de esta técnica destacan, que por el momento no se puede aplicar a proteínas mayores de 60-80 kDa, y que se necesitan concentraciones de proteína

elevadas (en torno a 1 mM), a las que algunas proteínas pueden presentar problemas de agregación.

El espectro más sencillo de RMN que se puede adquirir de una proteína, un espectro monodimensional de protón, proporciona las indicaciones más claras acerca de si una proteína está bien plegada o, por el contrario, parcial o completamente desnaturalizada. Si la proteína está bien plegada, en el espectro se observa una buena dispersión de la señal de los protones del enlace peptídico (que resuenan entre 6 y 10 ppm), y aparecen señales a campo alto (por debajo o alrededor de 0 ppm) que resultan de la interacción terciaria estable entre las cadenas laterales de aminoácidos alifáticos y residuos aromáticos. Cuando la proteína se desnaturaliza, disminuye drásticamente la dispersión de las señales de los protones amida y desaparecen las señales a campo alto, resultado de la pérdida de las interacciones terciarias características de la conformación nativa. Adicionalmente, los espectros monodimensionales permiten comprobar si una conformación parcialmente plegada es de tipo glóbulo fundido, ya que en los estados de glóbulo fundido son características las señales anchas; y si tal conformación parcialmente plegada forma agregados o no, ya que cuando existe agregación, la anchura de las señales y los desplazamientos químicos varían con la concentración (Wüthrich, 1986).

Otra importante aplicación de la técnica de RMN es que permite estudiar la unión entre biomoléculas. Cuando se produce una interacción entre distintas moléculas, se generan cambios en las señales del espectro de RMN (en los desplazamientos químicos, en la anchura de línea, y/o en la intensidad) con respecto a cada una de las moléculas por separado. En consecuencia, una manera de abordar el estudio por RMN de si existe o no interacción, consiste en comparar los espectros del mismo tipo (por ejemplo, los espectros monodimensionales) de las moléculas en estado libre, con respecto al espectro del complejo formado (Wüthrich, 1986).

1.4.1.2. Técnicas hidrodinámicas: cromatografía de exclusión por tamaño molecular.

Esta técnica, aunque menos utilizada que las anteriores para estudios conformacionales de las proteínas, es muy útil para complementar la información obtenida por los métodos espectroscópicos. La cromatografía de exclusión por tamaño molecular se basa en la separación de las moléculas en función de su tamaño. El proceso de separación se lleva a cabo mediante la utilización de una matriz con partículas esféricas de tamaño de poro determinado, empaquetadas en una columna y rodeadas de

disolvente (fase estacionaria). La fase móvil transporta las moléculas de la muestra a través de la columna. La muestra está constituida por moléculas de distinto tamaño, que puede ser superior, inferior o igual al tamaño del poro. Las moléculas de menor tamaño difundirán dentro de los poros de las partículas, con lo que se moverán más lentamente a través de la columna; mientras que las moléculas de mayor tamaño no podrán difundir dentro de los poros y su recorrido será menor, saliendo en primer lugar de la columna (esto es, con un V_e más pequeño). Por lo tanto, todas las moléculas eluyen en orden decreciente de tamaño.

El mecanismo de separación depende tanto del peso molecular como de la forma de la molécula, es decir, de su volumen hidrodinámico, de modo que, la cromatografía de exclusión por tamaño molecular permite detectar tanto los cambios conformacionales, como la existencia o no de fenómenos de agregación en las proteínas. Así, durante el proceso de desnaturalización, el volumen de la proteína va cambiando, de modo que, cuanto mayor sea el grado de desplegamiento de la misma, mayor será su volumen hidrodinámico y diferente su V_e en la columna (Ackers, 1970; Corbett y Roche, 1984).

Además, esta técnica hidrodinámica permite detectar interacciones entre biomoléculas. Su utilización para este fin se basa en la diferencia en peso molecular, y por tanto en volumen, entre el complejo formado y sus componentes por separado. La aparición de un pico de mayor peso molecular (y por ello de mayor volumen) que el de las moléculas consideradas individualmente, indica la existencia del complejo, y en consecuencia, confirma que se ha establecido interacción (Ackers, 1967; Valdés y Ackers, 1979).

1.4.2. TÉCNICAS UTILIZADAS PARA DETERMINAR CONSTANTES DE UNIÓN.

Desde el punto de vista experimental, la forma más utilizada para abordar este estudio consiste en determinar la constante de unión mediante titulación, manteniendo constante la concentración de una de las especies implicadas en el proceso de unión, y aumentando progresivamente la concentración de la otra especie. Los dos estados de la molécula o moléculas (tanto aisladas como unidas) deben tener propiedades diferentes. A las técnicas descritas anteriormente se pueden añadir dos técnicas adicionales, como la cromatografía de afinidad analítica y la ITC.

1.4.2.1. Cromatografía de afinidad analítica.

La cromatografía de afinidad analítica permite la medición de constantes de unión en procesos de asociación. El método a seguir consiste en inmovilizar de forma covalente uno de los componentes de la interacción a estudiar (normalmente la proteína) a una matriz cromatográfica, de modo que no se altere su capacidad de unión con el ligando (molécula en estudio). Entonces se hace pasar a través de dicha columna una disolución que contiene la molécula en cuestión. Si existe interacción entre la molécula y la proteína, se produce la unión específica entre ambas y la molécula queda retenida en la columna. La unión entre ambos componentes de la interacción suele ser el resultado de un equilibrio de asociación-disociación y, como tal, se trata de una interacción reversible, de manera que la molécula termina eluyendo de la columna con distinto V_e dependiendo del grado de retención, que depende a su vez de la fortaleza de la interacción. Así pues, de forma general, haciendo pasar a través de la columna disoluciones de la molécula a distintas concentraciones, y determinando su diferente V_e , se puede determinar la constante de unión molécula-proteína (Dunn y Chaiken, 1974; Taylor y Chaiken, 1977; Formosa et al., 1991).

No está muy claro si la cromatografía de afinidad analítica proporciona la verdadera constante de unión o tan sólo permite una estimación de la misma. Esto se debe a que para la aplicación de esta técnica se debe asumir que durante el experimento cromatográfico, la molécula unida está siempre en equilibrio con la molécula libre en disolución. No obstante, los valores de la constante de unión obtenidos mediante cromatografía de afinidad analítica coinciden con los proporcionados por otras técnicas, y esto le otorga validez (Formosa et al., 1991).

1.4.2.2. Calorimetría de titulación isotérmica.

La ITC es una de las técnicas que más se utiliza para caracterizar las interacciones entre moléculas biológicas, debido a su gran exactitud y precisión (Cooper y Johnson, 1994; Doyle, 1997; Pierce et al., 1999). La técnica de ITC se basa en la medición electrónica del calor absorbido o liberado durante el proceso de asociación a una temperatura determinada. La ventaja principal frente a otras técnicas es que permite determinar el valor exacto de la constante de unión, la estequiometría de la reacción, y la entalpía de la asociación, simultáneamente y de forma directa en un solo experimento. El procedimiento experimental general para la determinación de la constante de unión molécula-proteína consiste en realizar dos titulaciones independientes. Por un lado, en la celda de reacción se coloca la disolución de proteína

a una concentración determinada, y se van añadiendo secuencialmente cantidades constantes de la molécula. Se registra el calor necesario para mantener constante la temperatura de la celda después de cada inyección de molécula. Por otro lado, se realiza un experimento control similar, que consiste en ir añadiendo secuencialmente cantidades constantes de la molécula a la celda de reacción que sólo contiene tampón y no proteína. Así pues, de esta forma, el calor absorbido o liberado durante la reacción de asociación molécula-proteína se obtiene como la diferencia entre el calor de la reacción molécula-proteína, y el correspondiente calor de la dilución molécula-tampón. En los experimentos de ITC es fundamental que las condiciones del tampón sean idénticas en todas las disoluciones utilizadas (de la proteína y de la molécula), ya que pequeñas diferencias en las concentraciones salinas pueden alterar en gran medida el calor necesario para mantener constante la temperatura.

1.5. EL DOMINIO DE DIMERIZACIÓN DE LA PROTEÍNA DE LA CÁPSIDA DEL VIH-1 (CA-C).

1.5.1. EL VIRUS DE LA INMUNODEFICIENCIA HUMANA: ESTRUCTURA.

1.5.1.1. El virus de la inmunodeficiencia humana.

En 1981 se reconoció la existencia de una nueva enfermedad denominada SIDA, con un efecto devastador sobre la población mundial. Desde ese momento se iniciaron intensos estudios que culminaron en la identificación del VIH como el agente causante del SIDA (Barré-Sinoussi et al., 1983; Gallo et al., 1984; Levy et al., 1984). Hay dos tipos de virus, VIH-1 y VIH-2, que son genética y antigenéticamente diferentes. Mientras que el primero es el responsable de la pandemia (epidemia a nivel mundial), el segundo es endémico de África y es sumamente raro fuera de esta región. El VIH actúa, principalmente, mediante la infección de los linfocitos CD4⁺ del organismo huésped, que son esenciales en la respuesta inmunitaria, ocasionando su paulatina y lenta destrucción (Ho et al., 1995; Perelson et al., 1996; Wei et al., 1998; Turner y Summers, 1999).

El VIH es un retrovirus que pertenece al género de los lentivirus, y como tal, se caracteriza porque posee un genoma formado por un ácido ribonucleico (ARN) y una cápsida cónica con envuelta lipídica. Como en todos los retrovirus, el ARN viral debe transcribirse inversamente, por acción de la RT, a ADN durante el ciclo viral, para poder así integrarse en el núcleo de la célula huésped.

1.5.1.2. Estructura molecular del VIH-1.

Se puede considerar que el VIH-1 es una partícula esférica con un diámetro entre 80-110 nm. El virión maduro (Fig. 1.7; Turner y Summers, 1999) presenta en su exterior una envuelta formada por una bicapa lipídica, que deriva de la membrana de la célula huésped y está constituida por la inserción de trímeros de la glicoproteína viral de superficie (gp120), anclados a esta bicapa a través de la interacción con la glicoproteína viral transmembranal (gp41). En esta envuelta se encuentran también antígenos de histocompatibilidad, actina y ubiquitina (Arthur et al., 1992), derivados de la membrana de la célula que infectan. Por el interior, esta bicapa lipídica está recubierta de múltiples copias de MA (p17), y en el centro del virus está la cápsida cónica formada por aproximadamente 1500 unidades de CA (p24). Dentro de la cápsida se encuentran dos copias idénticas del ARN genómico de cadena única y de polaridad positiva estabilizadas por la asociación con múltiples unidades de NC (p7), y tres enzimas virales esenciales por la función que desempeñan en el ciclo viral: la PR, encargada del procesamiento de las proteínas virales que forman la estructura del virión; la RT, que retrotranscribe el ARN viral a ADN; y la IN, que lleva a cabo la integración del genoma viral retrotranscrito en el genoma celular. Además, también se empaquetan dentro de la cápsida las proteínas accesorias: Nef, Vif y Vpr. Estas proteínas desempeñan funciones reguladoras importantes: Nef estimula la tasa de replicación viral, Vif parece ser que participa en las primeras fases de la transcripción inversa y es necesaria para la producción de viriones infectivos, y Vpr dirige la localización en el núcleo celular del complejo viral de preintegración, de modo que actúa como acelerador del ciclo de replicación viral.

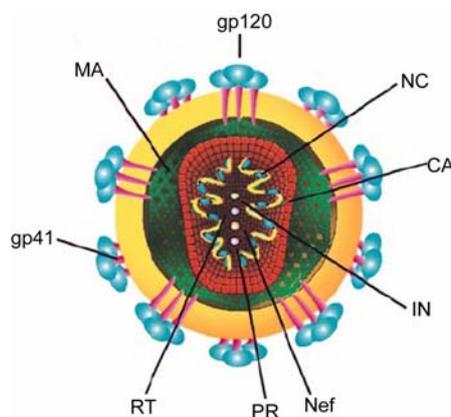


Figura 1.7 Estructura del virión maduro del VIH-1 (Turner y Summers, 1999). Para la descripción de las abreviaturas, ver el texto.

1.5.2. LA PROTEÍNA DE LA CÁPSIDA DEL VIH-1.

1.5.2.1. Maduración, morfología y composición de la cápsida del VIH-1.

El virión inmaduro (no infeccioso) del VIH-1 presenta, en el interior de la envuelta lipídica, una cápsida esférica (Gelderblom, 1991; Vogt, 1997) formada por aproximadamente unas 5000 copias de la poliproteína Gag (p55) (Briggs et al., 2004). Durante el proceso de maduración que sufre el virión inmaduro para dar lugar al virión maduro (con capacidad infecciosa), Gag es procesada proteolíticamente por acción de la enzima viral PR dando lugar a tres proteínas (MA, CA, y NC), y además, también se originan tres pequeños péptidos (p1, p2 y p6) que constituyen ribonucleoproteínas que se asocian al ARN viral (Mervis et al., 1988; Henderson et al., 1992). Como consecuencia del procesamiento proteolítico de la poliproteína Gag, la cápsida inmadura se disocia y tiene lugar un proceso de reorganización estructural dentro del virión, en el que MA permanece unida a la bicapa lipídica, NC se asocia con el ARN genómico formando un complejo nucleoproteico, y CA se asocia dando lugar a la cápsida madura (Fig. 1.7) que envuelve al complejo nucleoproteico y a las tres enzimas virales: PR, RT e IN (Vogt, 1997; Freed, 1998). Así, la cápsida madura del VIH-1 presenta una forma de cono truncado (Gelderblom et al., 1990; Briggs et al., 2003), y está formada por unas 1500 unidades idénticas de CA (Li et al., 2000; Briggs et al., 2003, 2004).

A pesar de que se han realizado numerosos estudios, todavía no se ha podido determinar con exactitud la morfología y estructura molecular de la cápsida inmadura y madura del VIH-1; no obstante, se han propuesto varios modelos para explicar el ensamblaje de la misma (Rosé et al., 1992; Dorfman et al., 1994; Misselwitz et al., 1995; Gamble et al., 1997; Frankel y Young, 1998; Berthet-Colominas et al., 1999; Ganser et al., 1999; Turner y Summers, 1999; Wilk y Fuller, 1999; Li et al., 2000; Göttlinger, 2001 a, b; Briggs et al., 2003). Sin embargo, no cabe duda de que la estructura de la cápsida del VIH-1 es el resultado de la asociación de las proteínas de la cápsida a través de varias interfases de oligomerización pequeñas y bien separadas unas de otras.

1.5.2.2. Estructura de la proteína de la cápsida del VIH-1.

CA, la proteína que forma la cápsida madura del VIH-1, contiene 231 aminoácidos y una masa molecular de 26 kDa. Está formada por dos dominios estructurales independientes: el dominio N-terminal, CA-N (residuos 1-145), y el dominio C-terminal, CA-C (residuos 146-231), conectados por una corta región flexible (Gitti et al., 1996; Momany et al., 1996; Gamble et al., 1996, 1997; Berthet-Colominas

et al., 1999) que permite distintos modos de asociación entre ellos. Se han obtenido las estructuras de alta resolución de ambos dominios aislados: la de CA-N mediante RMN (Gitti et al., 1996) y cristalografía de rayos X (Gamble et al., 1996), y la de CA-C mediante cristalografía de rayos X (Gamble et al., 1997; Worthylake et al., 1999). Además, también se ha resuelto la estructura de CA completa (Fig. 1.8; Berthet-Colominas et al., 1999).

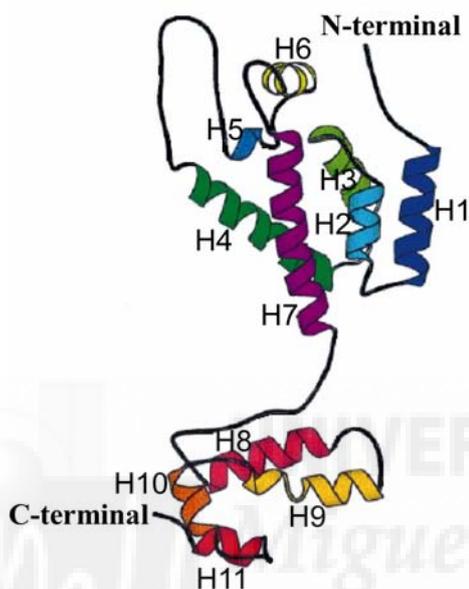


Figura 1.8 Modelo de cintas de la estructura del monómero de CA (Berthet-Colominas et al., 1999). Las distintas hélices están representadas con colores diferentes. En el dominio N-terminal: azul oscuro, hélice 1 (H1); azul claro, hélice 2 (H2); verde claro, hélice 3 (H3); verde oscuro, hélice 4 (H4); violeta claro, hélice 5 (H5); amarillo claro, hélice 6 (H6); y, violeta oscuro, hélice 7 (H7). Y en el dominio C-terminal: fucsia, hélice 8 (H8); amarillo oscuro, hélice 9 (H9); naranja, hélice 10 (H10); y, rojo, hélice 11 (H11).

CA-N está compuesto principalmente por siete hélices α , dos horquillas β y un bucle largo. CA-N del VIH-1 se encuentra en forma monomérica en disolución, incluso a las elevadas concentraciones (mM) utilizadas para la obtención de espectros de RMN. Se ha determinado que CA-N es fundamental para el ensamblaje de la cápsida madura y las fases posteriores del ciclo viral (Dorfman et al., 1994).

CA dimeriza en disolución a través de su dominio C-terminal, CA-C, con una constante de disociación de aproximadamente 18 μ M (Rosé et al., 1992). CA-C, la proteína objeto de este trabajo, contiene 86 aminoácidos (residuos 146-231 de CA

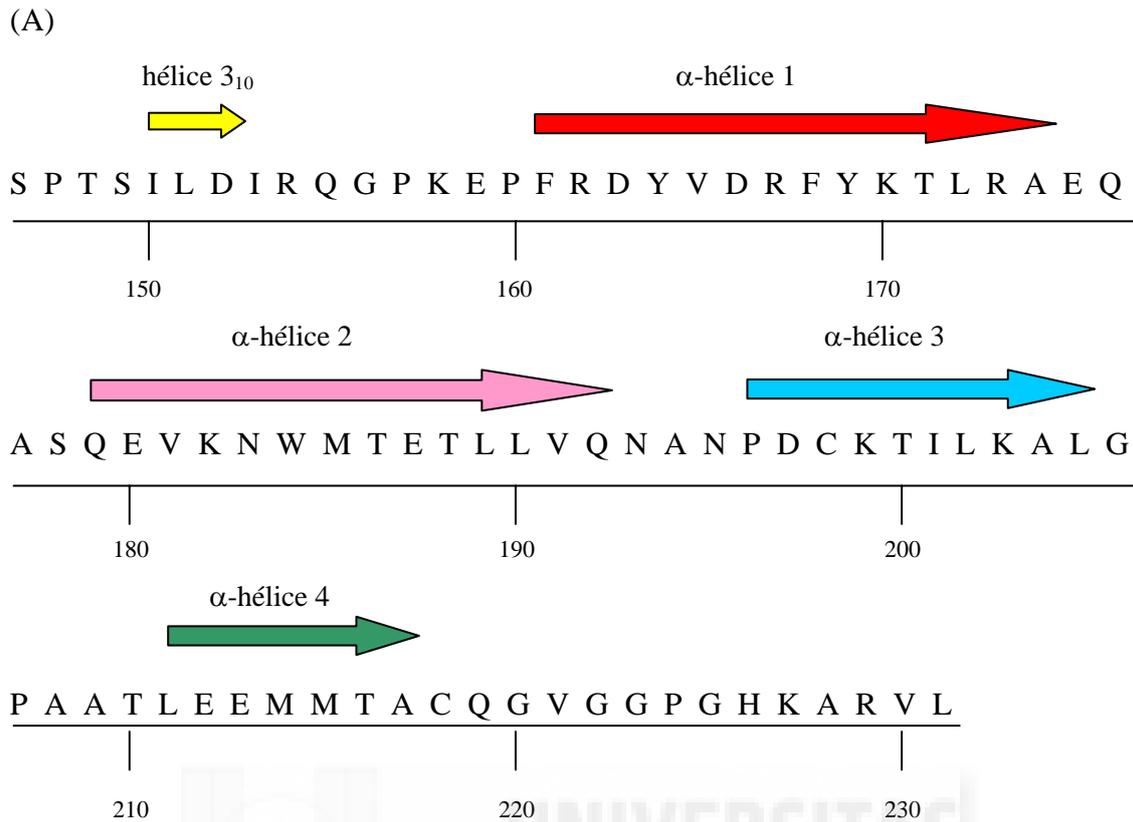
intacta; Fig. 1.9) y una masa molecular de 9.65 kDa. CA-C aislado es un dímero, tanto en la forma cristalizada como en disolución, y su constante de disociación tiene un valor de 10 μ M (Gamble et al., 1997), muy similar al de la proteína CA completa. Además, se ha visto que mutaciones puntuales en residuos localizados en la interfase de dimerización de CA-C, impiden la dimerización de CA completa (Gamble et al., 1997). Estas observaciones, junto con la comparación de la estructura de CA-C, tanto aislada como formando parte de CA, confirman que CA-C contiene íntegramente la interfase de dimerización de CA.

145	ATG Met	AGC Ser	CCT Pro	ACC Thr	AGC Ser	ATT Ile	CTG Leu	GAC Asp	ATA Ile	AGA Arg	154
155	CAA Gln	GGA Gly	CCA Pro	AAA Lys	GAA Glu	CCT Pro	TTT Phe	AGA Arg	GAC Asp	TAT Tyr	164
165	GTA Val	GAC Asp	CGG Arg	TTC Phe	TAT Tyr	AAA Lys	ACT Thr	CTA Leu	AGA Arg	GCC Ala	174
175	GAG Glu	CAA Gln	GCT Ala	TCA Ser	CAG Gln	GAG Glu	GTA Val	AAA Lys	AAT Asn	TGG Trp	184
185	ATG Met	ACA Thr	GAA Glu	ACC Thr	TTG Leu	TTG Leu	GTC Val	CAA Gln	AAT Asn	GCG Ala	194
195	AAC Asn	CCA Pro	GAT Asp	TGT Cys	AAG Lys	ACT Thr	ATT Ile	TTA Leu	AAA Lys	GCA Ala	204
205	TTG Leu	GGA Gly	CCA Pro	GCG Ala	GCT Ala	ACA Thr	CTA Leu	GAA Glu	GAA Glu	ATG Met	214
215	ATG Met	ACA Thr	GCA Ala	TGT Cys	CAG Gln	GGA Gly	GTA Val	GGA Gly	GGA Gly	CCC Pro	224
225	GGC Gly	CAT His	AAG Lys	GCA Ala	AGA Arg	GTT Val	TTG Leu	TAG ter			232

Figura 1.9 Secuencia del gen de CA-C y su producto de expresión. En negrita están indicados el codon de iniciación (ATG) y el de *STOP* (TAG).

Cada monómero de CA-C está formado por una hélice 3_{10} corta seguida por una hebra extendida y cuatro hélices α conectadas por lazos cortos (Gamble et al., 1997). La interfase de dimerización está principalmente formada por la interacción mutua de una hélice α de cada monómero, la hélice 2 (Ser178-Val191; es la hélice 9 de la estructura de CA completa, Fig. 1.8), alineadas de forma paralela (Fig. 1.10). Asimismo, algunos residuos situados en otros elementos de la estructura de cada monómero participan en la interfase (Gamble et al., 1997; Worthylake et al., 1999). CA-C contiene una secuencia de 20 aminoácidos (residuos 153-172) denominada MHR, que se encuentra altamente conservada entre los retrovirus (Craven et al., 1995), y forma un motivo estructural hebra extendida-giro-hélice, que no participa en la interfase de dimerización, pero forma una intrincada red de puentes de hidrógeno con la α -hélice 2 (Gamble et al., 1997). MHR es necesaria para el ensamblaje de la cápsida y las etapas posteriores del ciclo viral (Alin y Goff, 1996; Willems et al., 1997; Cairns y Craven, 2001). Se han encontrado varios residuos de MHR evolutivamente conservados que son esenciales para el plegamiento y la estabilidad del monómero de CA-C (Mateu, 2002); de ahí la implicación de MHR en el ensamblaje de la cápsida del VIH-1, y por lo tanto, en la infectividad del virus (del Álamo et al., 2003). CA-C tiene dos residuos de cisteína en su secuencia, Cys198 y Cys218, que forman un puente disulfuro intramolecular que se encuentra enormemente conservado en todas las cápsidas de los retrovirus. Se ha determinado que el puente disulfuro juega un papel importante en el ensamblaje y maduración de la cápsida, y por tanto, en la infectividad del VIH-1 (Mc Dermott et al., 1996; Khorasanizadeh et al., 1999; Worthylake et al., 1999). Es de destacar la presencia de un triptófano (Trp184) en la interfase de dimerización, y dos tirosinas (Tyr164 y Tyr169) enterradas en el núcleo hidrofóbico de cada monómero, muy lejos de la interfase de dimerización.

Algunos estudios han demostrado que la dimerización de CA a través de su dominio C-terminal, CA-C, es fundamental para la oligomerización de Gag (Dorfman et al., 1994; Franke et al., 1994; Mammano et al., 1994; Reicin et al., 1995), y para el ensamblaje y maduración de la cápsida del VIH-1 (Jowett et al., 1992; Wang y Barklis, 1993; Gamble et al., 1996; Borsetti et al., 1998; von Schwedler et al., 1998, 2003).



(B)

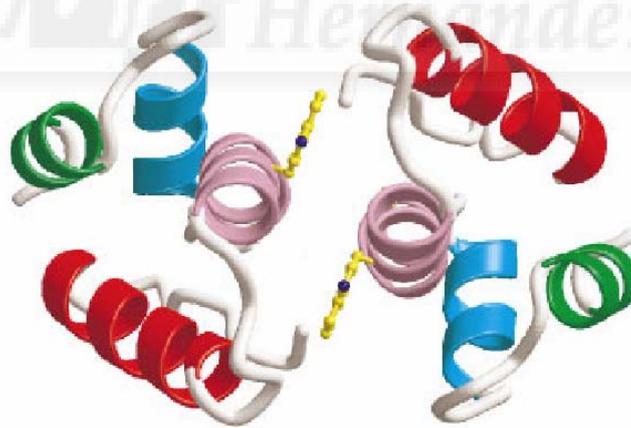


Figura 1.10 Estructura de CA-C. (A) Estructura secundaria de CA-C. Las distintas hélices están representadas con flechas sobre los aminoácidos correspondientes en la secuencia y con colores diferentes: amarillo, hélice 3₁₀ (residuos 150-152); rojo, α -hélice 1 (residuos 161-174); rosa, α -hélice 2 (residuos 179-192); azul, α -hélice 3 (residuos 196-205); y, verde, α -hélice 4 (residuos 211-217). (B) Modelo de cintas de la estructura del dímero de CA-C (Gamble et al., 1997). En la interfase de dimerización aparece representada la cadena lateral del Trp184. El código de colores es el mismo que en el apartado (A).

Se ha caracterizado la ruta de disociación-desnaturalización (o a la inversa, de plegamiento-asociación) del dímero de CA-C (Mateu, 2002), y ésta sigue un modelo de tres estados ($N_2 \leftrightarrow 2I \leftrightarrow 2D$), en el que el dímero nativo de CA-C (N_2) se disocia y despliega parcialmente dando lugar a un intermediario monomérico de baja estabilidad conformacional (I). Este intermediario mantiene intacta la estructura secundaria nativa, pero ha perdido parte de las interacciones terciarias intramonoméricas. En una segunda transición, el intermediario monomérico (I) se desnaturaliza cooperativamente dando lugar al monómero completamente desnaturalizado (D).

1.5.3. JUSTIFICACIÓN DE LA IMPORTANCIA DEL ESTUDIO ESTRUCTURAL Y TERMODINÁMICO DE CA-C.

Desde la identificación en 1983 del VIH como el agente causante del SIDA (Barré-Sinoussi et al., 1983), millones de personas han muerto, y por tanto, se han llevado a cabo tremendos esfuerzos para poder encontrar una solución a esta enfermedad.

La mayor parte de los medicamentos actuales de que se dispone para el tratamiento del VIH se basan en la inhibición de la función de alguna proteína específica del virus, utilizando como dianas terapéuticas, fundamentalmente, a las enzimas RT y PR (Naesens et al., 1997; De Clercq, 1998; Tavel et al., 1999), sólo dos de las quince proteínas codificadas en el genoma viral. Cuando se administran individualmente, estos medicamentos no son muy efectivos, pues retrasan la aparición de la enfermedad pero no la evitan en ningún caso, debido a la aparición de virus mutantes resistentes a los fármacos (De Clercq, 2002; Mansky, 2002; Frenkel y Tobin, 2004). La aparición de resistencia se debe, principalmente, a la baja fidelidad del proceso de replicación, ya que la RT carece de una actividad correctora (Coffin, 1995), pero también a la elevada tasa de mutación de los virus de ARN (Domingo y Holland, 1997; Domingo et al., 2001 a, b), al elevado número de rondas de replicación durante la infección (Coffin, 1995), y a la integración del genoma viral en el ADN celular en forma de provirus, dando lugar a sitios donde se acumulan los virus no susceptibles a los fármacos (Chun et al., 1997; Finzi et al., 1997; Wong et al., 1997). Cuando estos medicamentos se utilizan en combinaciones adecuadas (terapia combinada), la aparición de virus mutantes resistentes se reduce en gran medida, pero no se elimina (Gulick et al., 1997; Pillay, 2000; Richman, 2001). Por todo ello, se hace necesario continuar la investigación de otros componentes virales como objetivos potenciales de fármacos diversos.

Una diana potencial es la proteína CA, debido a su papel esencial en la morfogénesis del virión. Los agentes que afectan tanto al ensamblaje de la cápsida como a su estabilidad son fuertes inhibidores potenciales de la replicación viral: la inhibición del ensamblaje de CA mediante la introducción de mutaciones tiene un efecto letal (Gross et al., 1997; von Schwedler et al., 1998; Tang et al., 2001) y la alteración de la estabilidad de la cápsida reduce la replicación del virus (Forshey et al., 2002). Estos resultados apoyan una potencial estrategia antiviral contra el VIH-1, basada en la utilización de compuestos capaces de interferir en el reconocimiento entre los monómeros de CA, y como consecuencia, en el ensamblaje de la cápsida. Sin embargo, el diseño racional de cualquier tipo de inhibidor interfásico requiere un conocimiento detallado de la estructura de la proteína diana, y una caracterización termodinámica de las interacciones que están implicadas en la superficie de interacción, en este caso la interfase entre ambos monómeros de CA, que como se ha indicado en el apartado 1.5.2.2., está íntegramente contenida en su dominio C-terminal, CA-C.

1.6. OBJETIVOS.

El presente trabajo se puede dividir en tres grandes capítulos:

El primero se centra en la caracterización termodinámica de las formas monomérica y dimérica de CA-C, mediante la utilización de varias técnicas biofísicas: fluorescencia, DC, FTIR, RMN, y cromatografía de exclusión por tamaño molecular, a diferentes condiciones de pH y temperatura. El objetivo es, por un lado, describir la estabilidad de las formas monomérica y dimérica de CA-C y, por otro, caracterizar cualquier posible intermediario que pueda aparecer en la reacción de plegamiento-asociación de la proteína. Este estudio se realizó para comprender en profundidad las interacciones intramoleculares que posibilitan la formación del dímero de CA-C.

En el segundo capítulo se ha llevado a cabo el diseño de un péptido que comprende la única hélice α que forma la interfase de dimerización de CA-C (α -hélice 2) y se han estudiado, tanto su estructura como sus propiedades de asociación a CA-C, mediante diversas técnicas biofísicas: fluorescencia, DC, RMN, cromatografía de exclusión por tamaño molecular, cromatografía de afinidad analítica, e ITC. El objetivo de este capítulo es determinar si la α -hélice 2, por sí sola, es suficiente para inhibir la dimerización de CA-C. La utilidad principal de este estudio es que posibilita

la descripción de las interacciones intermoleculares, y el diseño como cabeza de serie de un péptido con posibles aplicaciones farmacéuticas.

Y por último, el tercer capítulo se centra en el diseño, mediante de técnicas de ingeniería de proteínas, de miniproteínas de CA-C que tengan la capacidad de heterodimerizar con la proteína silvestre, e inhibir de este modo la dimerización natural de la cápsida, y por tanto, la infectividad del VIH.



***CARACTERIZACIÓN
TERMODINÁMICA
DE CA-C***



Para describir la estabilidad de las formas monomérica y dimérica de CA-C, y caracterizar cualquier posible intermediario que pueda aparecer en la reacción de plegamiento-asociación de la proteína, así como para comprender los factores energéticos que gobiernan la dimerización, se ha llevado a cabo la caracterización termodinámica, tanto del monómero como del dímero de CA-C, mediante la utilización de diversas técnicas biofísicas: fluorescencia, DC, FTIR, RMN, y cromatografía de exclusión por tamaño molecular, a diferentes condiciones de pH y temperatura.

Los resultados indican que el equilibrio de desnaturalización térmica de CA-C se compone de tres transiciones térmicas reversibles e independientes: (i) la primera transición ($T_m \sim 307$ K), que se observa por FTIR y RMN, supone la disociación del dímero de CA-C, dando lugar a un intermediario monomérico plegado; (ii) la segunda transición ($T_m \sim 325$ K), observada por fluorescencia del ANS, anisotropía y DC en el ultravioleta cercano, implica el desplegamiento parcial del intermediario monomérico plegado; y, (iii) la tercera transición ($T_m \sim 332$ K), que se observa por DC en el ultravioleta lejano, absorbancia, FTIR y RMN, conlleva el desplegamiento global del intermediario monomérico parcialmente plegado. Además, también se han determinado los parámetros termodinámicos de la disociación y desnaturalización de las formas dimérica y monomérica de CA-C a pH fisiológico. Se ha observado que CA-C, tanto en forma monomérica como dimérica, experimenta una transición conformacional a pH ácido, formando una especie con las características propias de los llamados glóbulos fundidos, cuyas propiedades son similares a las que presenta el intermediario monomérico parcialmente plegado observado en el equilibrio de desnaturalización térmica de CA-C (paso (ii) arriba indicado).

2.1. MATERIALES Y MÉTODOS.

2.1.1. PRODUCCIÓN DE CA-C.

2.1.1.1. Expresión de CA-C.

Para la expresión de CA-C silvestre se utilizó el plásmido recombinante pET21b(+)-CAC cedido por el Dr. Mauricio G. Mateu (CSIC-UAM, Madrid). Está basado en el vector de expresión pET21b(+) (Novagen), y contiene el segmento que codifica el dominio C-terminal de CA (residuos 146-231) de la cepa BH10 del VIH-1, flanqueado por señales de iniciación y terminación y los sitios de restricción únicos *NdeI* y *BamHI*.

El plásmido pET21b(+)CAC se transformó en la cepa *E. coli* BL21 (DE3). Una colonia transformante recién obtenida se inoculó en 15 ml de LB suplementado con ampicilina a una concentración final de 100 µg/ml, y se incubó a 37 °C en agitación a 250 rpm hasta alcanzar una densidad óptica a 600 nm de 0.7-0.8. Para eliminar los posibles productos de desecho bacterianos y maximizar el rendimiento de la expresión proteica, las células de este preinóculo se separaron del medio de cultivo mediante centrifugación a 6000 rpm y se resuspendieron en 15 ml de LB fresco con ampicilina a 100 µg/ml, conservándolo a 4 °C hasta el día siguiente. La expresión de CA-C a gran escala se llevó a cabo inoculando 3 ml del preinóculo anterior en cuatro erlenmeyers de 2 litros que contenían 500 ml de LB suplementado con ampicilina a una concentración final de 100 µg/ml. Estos cultivos se incubaron a 37 °C en agitación a 250 rpm hasta que alcanzaron una densidad óptica a 600 nm de 0.7-0.8, y en ese momento se indujo la expresión mediante la adición de IPTG a una concentración final de 1 mM. Los cultivos se mantuvieron en las mismas condiciones (temperatura y velocidad de agitación) durante 3 horas más. A continuación, las células se separaron del medio de cultivo mediante centrifugación a 6000 rpm durante 15 minutos a 4 °C en una centrífuga Beckman Coulter J2-HS con un rotor JA-10, y se guardaron congeladas a -20 °C hasta su utilización.

2.1.1.2. Purificación de CA-C.

El protocolo utilizado para la purificación de CA-C (Mateu, 2002) es una modificación del descrito por Hill y colaboradores (Worthylake et al., 1999). Las células, obtenidas como se indica en el apartado 2.1.1.1., se resuspendieron y sonicaron a 4 °C en tampón 10 mM Tris-HCl, 0.1 mM EDTA, pH 8.0. El extracto celular se centrifugó a 16000 rpm durante 1 hora a 4 °C en una centrífuga Beckman Coulter J2-HS con un rotor JA-20. Entonces se realizó una precipitación por fraccionamiento con sulfato amónico (Fluka) a 4 °C del extracto celular soluble, obteniéndose CA-C en la fracción entre el 40 % y el 70 % de saturación de sulfato amónico. La proteína precipitada se redisolvió en 25 mM fosfato potásico pH 6.0 y se dializó frente al mismo tampón a 4 °C, utilizando membranas de diálisis con un tamaño de exclusión de 3500 Da (Spectrapore).

A continuación, la disolución de proteína obtenida se cargó a temperatura ambiente en una columna de intercambio catiónico de 5 ml (SP-Sepharose, Amersham

Biosciences), previamente equilibrada en 25 mM fosfato potásico pH 6.0, acoplada a un aparato de cromatografía AKTA-FPLC con un detector de absorbancia a 280 nm (Amersham Biosciences). La proteína retenida se eluyó con un gradiente lineal de 0 a 0.5 M de NaCl, con un flujo de 1 ml/min, apareciendo el máximo de elución a 120-150 mM NaCl. Las fracciones que contenían CA-C se mezclaron y dializaron frente a 25 mM fosfato potásico pH 6.0 a 4 °C. Entonces, la muestra dializada se cargó a temperatura ambiente en una columna de afinidad de 5 ml (Heparin-Sepharose, Amersham Biosciences), previamente equilibrada en 25 mM fosfato potásico pH 6.0. La proteína retenida se eluyó con un gradiente lineal de 0 a 0.5 M de NaCl, con un flujo de 1 ml/min, apareciendo el máximo de elución a 150-180 mM NaCl. Las fracciones que contenían CA-C se mezclaron y, posteriormente, la mezcla se cargó a temperatura ambiente a una columna de exclusión por tamaño molecular (Superdex 75 16/60, Amersham Biosciences), previamente equilibrada en 25 mM fosfato potásico, 150 mM NaCl, pH 7.0. La proteína se eluyó con un flujo de 1 ml/min. Las fracciones que contenían CA-C se mezclaron, y esta mezcla se dializó frente a H₂O (desionizada y purificada utilizando un sistema Millipore) a 4 °C. Las alícuotas obtenidas se congelaron inmediatamente en nitrógeno líquido y se almacenaron a -20 °C hasta su utilización. La repetida congelación y descongelación no provocó agregación apreciable ni un cambio significativo en la concentración de proteína. CA-C fue completamente soluble, incluso a concentraciones muy elevadas (del orden de mM). El rendimiento de la purificación osciló entre 10 y 30 mg de proteína por litro de cultivo.

2.1.1.3. Determinación de la pureza y concentración de CA-C.

La pureza de las preparaciones de CA-C se determinó mediante SDS-PAGE, y espectrometría de masas. El estudio por espectrometría de masas se llevó a cabo en el Servicio de Masas del Centro de Investigaciones Biológicas-CSIC (Madrid) por Alicia Prieto.

La concentración de proteína (monómero de CA-C) se determinó a partir de medidas de absorbancia a 280 nm, utilizando un espectrofotómetro ultravioleta/visible Beckman DU 530 y una cubeta de cuarzo de 1 cm de paso óptico (Hellma). El coeficiente de extinción molar a 280 nm (ϵ_{280}) utilizado se calculó a partir de la fórmula (Pace y Scholtz, 1997):

$$\varepsilon_{280} = \alpha n_{\text{Trp}} + \beta n_{\text{Tyr}} + \chi n_{\text{s-s}}$$

donde n es el número de residuos de triptófano, tirosina o puentes disulfuro, $\alpha = 5440$, $\beta = 1480$, y $\chi = 130$ unidades de absorbancia ($\text{M}^{-1} \text{cm}^{-1}$). De esta manera, el valor teórico para el monómero de CA-C fue: $\varepsilon_{280} = 8530 \text{ M}^{-1} \text{cm}^{-1}$.

2.1.2. FLUORESCENCIA.

Los experimentos de fluorescencia se realizaron en un espectrofluorímetro Cary Eclipse (Varian) conectado a una celda Peltier, utilizando una cubeta de cuarzo de 1 cm de paso óptico (Hellma).

2.1.2.1. Medidas en el estado estacionario.

2.1.2.1.1. Fluorescencia intrínseca de triptófano y tirosina.

Las muestras de proteína se excitaron a 280 y a 295 nm para poder discernir el comportamiento fluorescente de triptófanos y tirosinas (Schmid, 1997). Los espectros de emisión se recogieron entre 300 y 400 nm. Las rendijas de excitación y emisión fueron ambas de 5 nm. La señal se adquirió cada 1 nm, y el tiempo de integración fue de 1 s. Todos los espectros se corrigieron restando la línea de base adecuada. Los experimentos se realizaron a 298 K. Las concentraciones de proteína utilizadas fueron 20 y 200 μM , entre pH 2 y 12.

Se midió el pH de las muestras antes y después de cada experimento, con un pH-metro de Radiometer (Copenhage) que llevaba incorporado un electrodo ultrafino (Aldrich). Las sales y ácidos empleados en la preparación de los tampones fueron: ácido fosfórico (pH 2.0-3.0), ácido fórmico (pH 3.0-4.0), ácido acético (pH 4.0-5.5), NaH_2PO_4 (pH 6.0-7.0), Tris ácido (pH 7.5-9.0), Na_2CO_3 (pH 9.5-11.5), y Na_3PO_4 (pH 11.5-12.0). Para todas las muestras se usó una concentración de tampón de 25 mM.

2.1.2.1.2. Fluorescencia del ANS.

La unión a ANS se midió tomando espectros de fluorescencia de muestras con una concentración de proteína de 2 μM , a diferentes pHs (en el rango de pH 2 a 12), y en presencia de 100 μM de ANS. Los tampones se prepararon como se indica en el apartado 2.1.2.1.1. La disolución madre de ANS se preparó en agua, utilizando un coeficiente de extinción molar de $6.8 \times 10^3 \text{ M}^{-1} \text{cm}^{-1}$ a 370 nm (Mann y Matthews, 1993). Se excitó a 370 nm, y los espectros de emisión se recogieron entre 430 y 700 nm. La anchura de la rendija se fijó en 5 nm, tanto para la excitación

como para la emisión. La señal se adquirió cada 1 nm, y el tiempo de integración fue de 1 s. Todos los espectros se tomaron a 298 K. A 200 μM de proteína y 500 μM de ANS no se pudieron realizar experimentos, ya que el fotomultiplicador se saturaba a esas concentraciones de sonda.

2.1.2.1.3. Quenching de fluorescencia de triptófano y tirosina.

Se examinó el *quenching* de la fluorescencia intrínseca de la proteína por yoduro y acrilamida (Lakowicz, 1999) a diferentes valores de pH. Se excitó a 280 nm (para determinar el *quenching* asociado a los residuos de triptófano y tirosina), y a 295 nm (para determinar exclusivamente el *quenching* asociado al triptófano). Los espectros de emisión se recogieron entre 300 y 400 nm. La anchura de la rendija se fijó en 5 nm, tanto para la excitación como para la emisión. La señal se adquirió cada 1 nm, y el tiempo de integración fue de 1 s. Todos los espectros se tomaron a 298 K.

Para ambos *quencher*s, el rango de concentración utilizado fue de 0 a 0.7 M, en intervalos de 0.1 M, y la concentración de proteína fue 20 μM en todos los casos. En los experimentos en los que se utilizó como *quencher* el ión yoduro (I^-) en forma de yoduro potásico (KI), la fuerza iónica se mantuvo constante mediante la adición de KCl; además, se añadió $\text{Na}_2\text{S}_2\text{O}_3$ a una concentración final de 0.1 M con el fin de evitar la formación de I_3^- .

Los parámetros que definen el *quenching* dinámico y el estático por acrilamida se obtuvieron mediante el ajuste de los datos de intensidad de fluorescencia a diferentes longitudes de onda (en el intervalo de 335 a 338 nm) a la ecuación de Stern-Volmer (2.1), que incluye un término exponencial que define el *quenching* estático (Lakowicz, 1999):

$$F_0/F = (1 + K_{sv}[X])e^{v[X]} \quad (2.1),$$

donde K_{sv} es la constante de Stern-Volmer para el *quenching* dinámico; v es la constante de *quenching* estático; F_0 es la intensidad de fluorescencia en ausencia de *quencher*; y, F es la intensidad de fluorescencia medida a una concentración de *quencher* dada, $[X]$ (M). El *quenching* por yoduro no mostró un componente estático significativo, por lo que el término exponencial se eliminó para el ajuste a la ecuación (2.1).

2.1.2.2. Medidas de desnaturalización térmica.

2.1.2.2.1. Fluorescencia intrínseca de triptófano y tirosina.

Los experimentos de desnaturalización térmica de CA-C siguiendo la emisión de fluorescencia de los residuos de triptófano y tirosina, se realizaron a dos concentraciones de proteína, 20 y 200 μM , en el rango de pH 4 a 10, excitando a 280 nm, y recogiendo la intensidad de la emisión fluorescente tanto a 335 como a 350 nm. La anchura de la rendija se fijó en 5 nm, tanto para la excitación como para la emisión. Se tomaron datos cada 0.2 K, desde 290 hasta 363 K, y el tiempo de integración fue de 1 s. Se utilizó una velocidad de calentamiento de 30 K/h y 60 K/h. Cada experimento de desnaturalización térmica se repitió tres veces con muestras nuevas.

2.1.2.2.2. Fluorescencia del ANS.

Los experimentos se realizaron a dos concentraciones de proteína: 20 y 80 μM , con 100 y 160 μM de ANS, respectivamente, en el rango de pH 4 a 11. Para comprobar si había dependencia de las curvas de desnaturalización con la velocidad de calentamiento, se siguieron dos procedimientos diferentes:

(A) Aumentando la temperatura de forma manual, desde 290 hasta 363 K, de modo que se tomó un espectro cada 3 K, esperando previamente 5 minutos para que la temperatura se estabilizara. La longitud de onda de excitación fue de 370 nm, y los espectros de emisión se recogieron entre 430 y 600 nm. Las rendijas de excitación y emisión fueron ambas de 5 nm. La señal se adquirió cada 1 nm, y el tiempo de integración fue de 1 s.

(B) Aumentando la temperatura automáticamente, con una velocidad de calentamiento de 60 K/h y tomando un punto cada 0.2 K, desde 290 hasta 363 K. Se utilizaron distintas longitudes de onda de excitación: 360, 370 y 380 nm, y de emisión: 480 y 520 nm. La anchura de la rendija se fijó en 5 nm, tanto para la excitación como para la emisión, y el tiempo de integración fue de 1 s.

No se observaron diferencias entre las curvas de desnaturalización obtenidas por ambos procedimientos. Cada experimento de desnaturalización térmica se repitió tres veces con muestras nuevas.

2.1.2.2.3. Anisotropía.

La anisotropía en el estado estacionario, $\langle r \rangle$, se define por (Lakowicz, 1999):

$$\langle r \rangle = \frac{I_V - G I_H}{I_V + 2G I_H} \quad (2.2),$$

donde I_V e I_H son, respectivamente, las componentes paralela y perpendicular de la luz de fluorescencia polarizada; y, G es un factor de corrección para las distintas respuestas de fluorescencia a los dos componentes polarizados. El factor G fue 0.46 y 0.56 para la longitud de onda de emisión 330 y 350 nm, respectivamente. La anisotropía de fluorescencia se registró en una disposición en forma de L. Los experimentos se realizaron a dos concentraciones de proteína, 20 y 200 μM , a pH 7. Las muestras se excitaron a 280 nm, y la intensidad de emisión fluorescente se recogió a 330 y 350 nm. No se apreciaron diferencias entre las curvas obtenidas a ambas longitudes de onda de emisión; no obstante, las medidas tomadas a 350 nm y 20 μM de CA-C tuvieron una menor relación señal a ruido. La anchura de la rendija se fijó en 10 nm, tanto para la excitación como para la emisión y el tiempo de integración fue de 5 s. La temperatura se fue incrementando de forma manual, desde 278 hasta 363 K, tomando datos cada 3 K, esperando previamente 5 minutos para que la temperatura se estabilizara. Cada valor de anisotropía obtenido es la media de cuatro mediciones independientes para cada temperatura. Cada experimento de desnaturalización térmica se repitió dos veces con muestras nuevas.

2.1.3. DICROÍSMO CIRCULAR Y ABSORBANCIA.

2.1.3.1. Medidas en el estado estacionario.

Los espectros de DC se adquirieron en un espectropolarímetro Jasco J810 equipado con una celda termostaticada, conectada a un baño de agua Neslab RTE-111. El espectropolarímetro fue calibrado periódicamente con ácido (+)10-camforsulfónico.

Se tomaron espectros de DC en el ultravioleta lejano de CA-C a 20 y 200 μM , a diferentes pHs (en el rango de pH 2 a 12). Los tampones se prepararon como se indica en el apartado 2.1.2.1.1. Los espectros se adquirieron entre 190 y 250 nm, con una velocidad de barrido de 50 nm/min, una anchura de banda de 1 nm, y un tiempo de respuesta de 2 s, siendo promediados sobre cuatro experimentos a 298 K.

Las desnaturalizaciones químicas por GdmHCl se siguieron mediante DC en el ultravioleta lejano, utilizando 20 μ M de CA-C en tampón 25 mM fosfato potásico pH 7, y a una temperatura determinada en el rango de 278 a 313 K. La concentración exacta de la disolución madre de GdmHCl se calculó a partir de su índice de refracción (Pace y Scholtz, 1997). La titulación con GdmHCl se realizó a concentraciones entre 0 y 7 M, en intervalos de 0.25 M, y las muestras se dejaron equilibrar al menos una hora antes de llevar a cabo los experimentos. No se observaron diferencias entre las muestras preparadas la noche anterior y aquellas preparadas una hora antes. Las desnaturalizaciones químicas fueron reversibles, como lo demuestra el hecho de que las curvas sigmoidales obtenidas a partir de la dilución de muestras de proteína a 7 M de GdmHCl fueron superponibles a las de las muestras donde no había GdmHCl. Cada experimento de desnaturalización química fue repetido tres veces con muestras nuevas. Los espectros se adquirieron como se indica en el párrafo anterior.

Para la adquisición de espectros en el ultravioleta lejano se utilizó una cubeta de cuarzo de 0.1 cm de paso óptico (Hellma). Todos los espectros se corrigieron mediante la sustracción de la línea de base correspondiente.

La elipticidad molar por residuo, $[\Theta]$, se calculó según:

$$[\Theta] = \frac{\Theta}{(10lcN)} \quad (2.3),$$

donde Θ es la elipticidad experimental obtenida (mdeg); l es el paso óptico de la celda (cm); c es la concentración de proteína (M); y, N es el número de aminoácidos (86 en CA-C).

2.1.3.2. Medidas de desnaturalización térmica.

Los experimentos de desnaturalización térmica, tanto en el ultravioleta lejano como en el cercano, se realizaron en un espectropolarímetro Jasco J-710, en colaboración con Marta Bueno, que trabaja en el laboratorio del Dr. Javier Sancho (Departamento de Bioquímica y Biología Molecular y Celular, Facultad de Ciencias, Universidad de Zaragoza).

Las desnaturalizaciones térmicas en el ultravioleta lejano se adquirieron siguiendo la elipticidad a 222 nm, desde 278 hasta 363 K, utilizando velocidades de calentamiento de 30 K/h y 60 K/h, una anchura de banda de 1 nm, y un tiempo de respuesta de 8 s, en una cubeta circular (Hellma) dotada de una camisa termostatazada

por agua, con 0.1 cm de paso óptico. Las concentraciones de proteína utilizadas fueron 20 y 200 μM , en el rango de pH 4 a 10. La reversibilidad de las transiciones térmicas se comprobó realizando un nuevo experimento tras enfriar hasta 278 K la muestra previamente desnaturizada, y comparando éste con el espectro obtenido en el primer barrido. Cada experimento de desnaturización térmica se repitió dos veces con muestras nuevas. En ningún caso se observó precipitación después del calentamiento. La posibilidad de desviación del espectropolarímetro se descartó tomando espectros del tampón antes y después de cada experimento de temperatura, no encontrándose diferencia entre ambos.

Los experimentos de desnaturización térmica en el ultravioleta cercano a 280 nm se midieron de forma simultánea por DC y absorbancia, utilizando una cubeta circular (Hellma) dotada de una camisa termostatizada por agua, con 1 cm de paso óptico. Las concentraciones de proteína utilizadas fueron 20 y 200 μM , a distintos pHs. Los parámetros de adquisición de los experimentos fueron los mismos que los indicados para el ultravioleta lejano. Cada experimento de desnaturización térmica se repitió dos veces con muestras nuevas.

2.1.4. CROMATOGRAFÍA DE EXCLUSIÓN POR TAMAÑO MOLECULAR.

Se usó una columna analítica de exclusión por tamaño molecular (Superdex 75 H/R 10/30, Amersham Biosciences) acoplada a un aparato de cromatografía AKTA-FPLC (Amersham Biosciences) a 298 K, y detectando la absorbancia a 280 nm. La columna se calibró utilizando el conjunto de patrones de calibración de bajo peso molecular (Amersham Biosciences). Los patrones utilizados y sus correspondientes radios de Stokes fueron: ribonucleasa A (16.4 Å), quimiotripsinógeno (20.9 Å), ovoalbúmina (30.5 Å), y albúmina de suero bovino (35.5 Å) (Hinkle et al., 1999). Se utilizó un flujo de 1 ml/min, y se cargaron alícuotas de 100 μl en la columna después de ser equilibrada con 4 volúmenes de columna en el tampón de elución correspondiente según el pH a analizar (en el rango de pH 2 a 12). La concentración del tampón fue de 25 mM, con 150 mM de NaCl para evitar interacciones no específicas entre la columna y la proteína. Las concentraciones de CA-C utilizadas fueron 20 y 200 μM . El V_e obtenido a cada pH es el resultado de tres medidas independientes con muestras nuevas.

2.1.5. INFRARROJO POR TRANSFORMADA DE FOURIER.

Los espectros se adquirieron con un aparato Bruker FTIR-66S equipado con un detector DTGS y termostatizado con un baño de agua Braun. El interior del aparato fue constantemente purgado con aire purificado y desecado para evitar las contribuciones del vapor de agua ambiental a los espectros de infrarrojo. Se tomaron 50 espectros por muestra, promediados y apodizados con una función de Happ-Genzel, y finalmente sometidos a una transformada de Fourier para dar una resolución de 2 cm^{-1} . La proteína se secó en un concentrador Speed-Vac (Savant, Farmingdale) y se disolvió en el tampón correspondiente preparado con agua deuterada (Aldrich). Los valores de pH indicados no incluyen la corrección por el efecto isotópico. Las muestras de proteína se colocaron entre dos cristales de CaF_2 , usando un separador de $50\ \mu\text{m}$ en una celda desmontable Harrick.

Se realizaron dos series de experimentos con CA-C seguidos por FTIR:

(A) Para estudiar la banda de absorción de las tirosinas, se tomaron espectros de CA-C a una concentración de $1\ \text{mM}$, en el rango de pH 2 a 12, y a una temperatura de $278\ \text{K}$.

(B) Se llevaron a cabo experimentos de desnaturalización térmica: (i) de CA-C a $1\ \text{mM}$ y pH 7, en el intervalo de temperatura entre 278 y $363\ \text{K}$, y en base a los resultados obtenidos (apartado 2.2.2.6.), se llevó a cabo un estudio de la transición a baja temperatura de CA-C a $1\ \text{mM}$, en el rango de pH 4 a 9, calentando sólo hasta $325\ \text{K}$; y, (ii) de CA-C entre 0.53 y $4.5\ \text{mM}$, a pH 7. En todos los experimentos de desnaturalización térmica, la velocidad de calentamiento fue de $50\ \text{K/h}$, tomando un espectro cada $2.5\ \text{K}$. Las transiciones térmicas se analizaron siguiendo el valor de la anchura a $\frac{3}{4}$ de altura de la banda amida I'.

Todos los espectros se corrigieron restando la contribución del tampón correspondiente.

2.1.6. RESONANCIA MAGNÉTICA NUCLEAR.

Se realizaron experimentos monodimensionales con un espectrómetro Bruker DRX-500 a $298\ \text{K}$ adquiriéndose 32000 puntos y utilizando la secuencia WATERGATE (Piotto et al., 1992) para eliminar la señal del agua residual. Se tomaron 512 espectros, y la anchura espectral fue $6000\ \text{Hz}$ en todos los casos. La concentración de CA-C fue $200\ \mu\text{M}$ en tampón $25\ \text{mM}$ fosfato pH 7 en $90\ \%\ \text{H}_2\text{O}/10\ \%\ ^2\text{H}_2\text{O}$, o en tampón $25\ \text{mM}$ acetato deuterado pH 3 en $90\ \%\ \text{H}_2\text{O}/10\ \%\ ^2\text{H}_2\text{O}$. No se hicieron

correcciones de los valores de pH por los efectos isotópicos. También se realizaron experimentos con una concentración de proteína de 1 mM, y no se observaron diferencias, ni en los desplazamientos químicos ni en la variación con la temperatura, cuando se comparan con los espectros adquiridos a menor concentración (200 μ M).

Los espectros se procesaron utilizando el *software* BRUKER-XWINNMR. Para mejorar la resolución, se aplicó una función ventana exponencial y correcciones de línea de base polinomiales. Los datos de cada espectro monodimensional final contenían 64000 puntos. Los desplazamientos químicos están referenciados respecto a TSP.

2.1.7. ANÁLISIS DE LAS CURVAS DE DESNATURALIZACIÓN.

La energía de emisión promedio, $\langle \lambda \rangle$, en los espectros de fluorescencia se calculó como (Royer, 1995):

$$\langle \lambda \rangle = \frac{\sum_i^n \frac{I}{\lambda_i} I_i}{\sum_i^n I_i} \quad (2.4),$$

donde I_i es la intensidad de fluorescencia medida a una longitud de onda dada λ_i . Este parámetro se utilizó para seguir los cambios en la exposición de los residuos aromáticos en los experimentos de desnaturalización por pH, ya que al tener en cuenta todas las longitudes de onda del espectro de emisión, es capaz de detectar cambios pequeños, y como su valor final resulta de una media de muchos valores, el error experimental contribuye en menor medida.

Los experimentos de desnaturalización por pH se analizaron asumiendo que ambas especies (protonada y desprotonada) contribuyen por igual a la propiedad física en estudio, X (energía de emisión promedio, longitud de onda de máxima emisión, elipticidad, V_e , o máximo de la banda de absorción de FTIR), utilizando la siguiente ecuación:

$$X = \frac{(X_a + X_b 10^{n(pH-pK_a)})}{(1 + 10^{n(pH-pK_a)})} \quad (2.5),$$

donde n es el coeficiente de Hill, que proporciona información sobre la cooperatividad de la transición; X_a es la propiedad física observada a pHs bajos (forma ácida); X_b es la propiedad física observada a pHs altos (forma básica); y, pK_a es el punto medio aparente de la titulación. El valor de pK_a se obtuvo a partir de tres titulaciones

independientes con muestras nuevas. En todos los experimentos realizados en este trabajo, el valor de n estuvo próximo a 1.0, como ocurre en otras proteínas (Forsyth y Robertson, 2000).

Los datos de desnaturalización química (obtenidos siguiendo el cambio en la elipticidad a 222 nm) y térmica, se ajustaron a la ecuación para dos estados:

$$X = \frac{(X_N + X_D e^{(-\Delta G/RT)})}{(1 + e^{(-\Delta G/RT)})} \quad (2.6),$$

donde X_N y X_D son las correspondientes fracciones del estado plegado y desplegado, respectivamente, para las que se asume una relación lineal con la concentración de agente desnaturalizante, y que se definen como: $X_N = \alpha_N + \beta_N[D]$, y $X_D = \alpha_D + \beta_D[D]$; X es la propiedad física observada; ΔG es la energía libre del proceso de desnaturalización; R es la constante de los gases ($1.987 \text{ cal mol}^{-1} \text{ K}^{-1}$); y, T es la temperatura (K).

Las curvas de desnaturalización química a diferentes temperaturas se analizaron siguiendo el modelo de extrapolación lineal: $\Delta G = m([D]_{1/2} - [D])$ (Pace, 1986; Pace y Laurents, 1989; Pace y Scholtz, 1997), donde $[D]$ es la concentración de agente desnaturalizante; $[D]_{1/2}$ es la concentración de agente desnaturalizante en el punto medio de la transición; y, m es la pendiente de la zona de transición entre el estado plegado y desplegado, que informa de la dependencia de ΔG con respecto a la concentración de agente desnaturalizante.

Para las curvas de desnaturalización térmica, el cambio de la energía libre con la temperatura en la ecuación (2.6) viene dado por la ecuación de Gibbs-Helmholtz:

$$\Delta G(T) = \Delta H_m \left(1 - \frac{T}{T_m} \right) - \Delta C_p \left[(T_m - T) + T \ln \left(\frac{T}{T_m} \right) \right] \quad (2.7),$$

donde ΔH_m es el cambio en la entalpía de van't Hoff; T_m es la temperatura media de desnaturalización; y, ΔC_p es la diferencia de capacidad calorífica entre los estados nativo y desnaturalizado. La sustitución de esta expresión en la ecuación (2.6) permitió la determinación de los parámetros termodinámicos ΔH_m , T_m y ΔC_p en los experimentos de desnaturalización térmica. En los experimentos de FTIR, donde se observó un comportamiento dependiente de la concentración, el cambio de la energía libre con la temperatura se calculó mediante la siguiente ecuación (Backmann et al., 1998):

$$\Delta G(T) = \Delta H_m \left(1 - \frac{T}{T_m} \right) - \Delta C_p \left[(T_m - T) + T \ln \left(\frac{T}{T_m} \right) \right] - RT \ln(2 C_t) \quad (2.8),$$

donde C_t es la concentración molar de proteína expresada en equivalentes de dímero.

Los ajustes de los datos por mínimos cuadrados a las diferentes ecuaciones se llevaron a cabo empleando la opción de ajuste general de curvas del programa KaleidaGraph (*software* de Abelbeck). Asimismo, también se realizó un ajuste global de los datos de DC en el ultravioleta lejano y cercano, a una concentración de proteína de 20 y 200 μM (Irún et al., 2001). Se escogieron estos experimentos para realizar el ajuste global porque presentaban una mejor relación señal a ruido. Por ambos procedimientos de ajuste utilizados se obtuvieron los mismos resultados, dentro del error experimental.

2.2. RESULTADOS.

2.2.1. DESNATURALIZACIÓN POR pH.

Se utilizaron diversas técnicas biofísicas, que aportan información complementaria sobre el desplegamiento de la proteína. Teniendo en cuenta que la constante de disociación determinada para CA-C tiene un valor de 10 μM (Gamble et al., 1997; Mateu, 2002; del Álamo et al., 2003), los experimentos se realizaron a distintas concentraciones de proteína. Para los experimentos de FTIR y RMN se usó una concentración de 1 mM (concentración a la que CA-C se encuentra al 100 % en forma dimérica); y, para el resto de las técnicas: 200 μM (el 90 % de CA-C en forma dimérica), 20 μM (el 40 % de CA-C en forma monomérica), y 2 μM (el 75 % de CA-C en forma monomérica).

2.2.1.1. Fluorescencia intrínseca de triptófano y tirosina.

En cada monómero de CA-C hay tres residuos fluorescentes: el Trp184, situado en la interfase de dimerización; y, las dos tirosinas, Tyr164 y Tyr169, enterradas en el núcleo hidrofóbico de cada monómero, muy lejos de la interfase de dimerización. Se tomaron espectros de fluorescencia de CA-C a diferentes pHs, y se siguieron los cambios tanto en la longitud de onda de máxima emisión como en la energía de emisión promedio.

El espectro de emisión de fluorescencia de CA-C a 200 μM y pH fisiológico está dominado por la fluorescencia del único residuo de triptófano, ya que mostró un máximo a 340 nm y no se observaron hombros aparentes a 305 nm (máximo de emisión de las tirosinas). Conforme se fue incrementando el pH por encima de pH neutro, los espectros se desplazaron hacia mayores longitudes de onda (Fig. 2.1 A), situándose el máximo en torno a 350 nm, debido a la desnaturalización básica de la proteína. No se

pudo determinar el pK_a de esta titulación a causa de la ausencia de línea de base a pHs básicos, pero su valor aparentemente alto sugiere que se puede corresponder con la desprotonación de residuos de lisina o arginina (Cantor y Schimmel, 1980). Al ir disminuyendo el pH, también se observó un desplazamiento del máximo hacia mayores longitudes de onda, debido a la desnaturalización ácida de la proteína, con un pK_a de 4.3 ± 0.2 . Este valor es similar a la constante de ionización de las cadenas laterales del ácido aspártico, ácido glutámico y del grupo carboxilo C-terminal (Cantor y Schimmel, 1980). Al representar la energía de emisión promedio frente al pH se observó un perfil consistente con el anterior, con dos titulaciones: una a pH bajo con un pK_a de 4.2 ± 0.2 y otra a pH alto (Fig. 2.1 A).

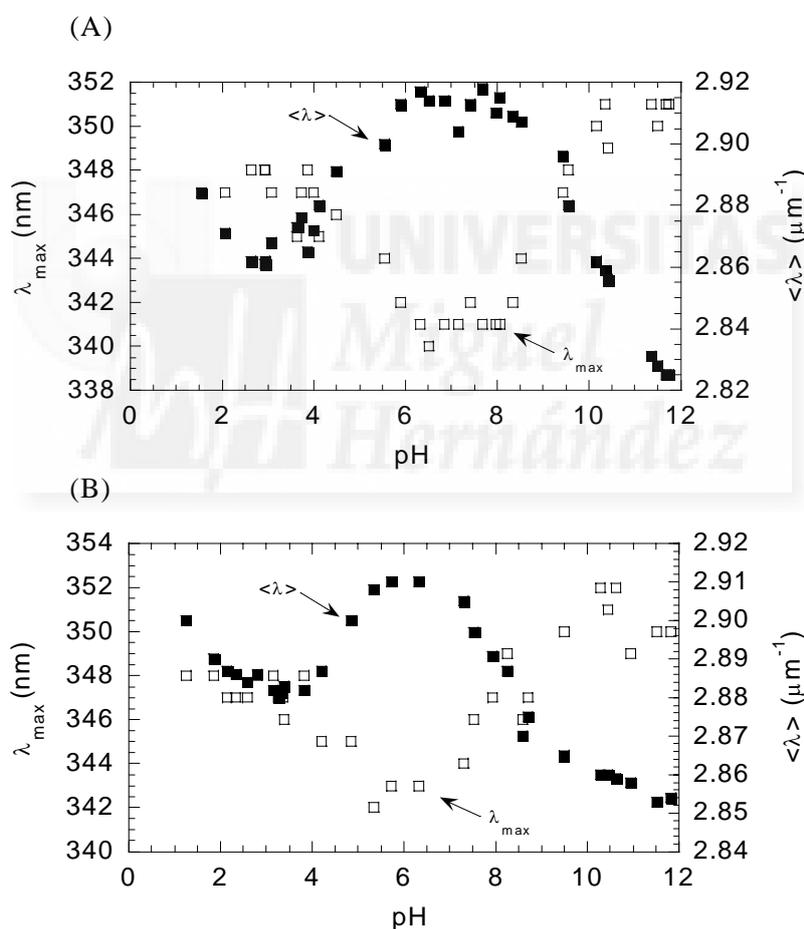


Figura 2.1 Desnaturalización por pH seguida mediante fluorescencia. Comportamiento de la longitud de onda de máxima emisión (cuadrados blancos, eje izquierdo) y de la energía de emisión promedio (cuadrados negros, eje derecho) frente al pH, a dos concentraciones distintas de CA-C: (A) 200 μM y (B) 20 μM . Los espectros se adquirieron con una longitud de onda de excitación de 280 nm, y a 298 K.

Los experimentos hechos a 20 μM de CA-C mostraron el mismo comportamiento con el pH que los realizados a 200 μM , tanto para el máximo como para la energía de emisión promedio (Fig. 2.1 B). Sin embargo, el espectro de emisión de fluorescencia de CA-C a 20 μM y pH fisiológico mostró un máximo a 343 nm, que se encuentra desplazado hacia el rojo con respecto al observado a 200 μM (340 nm); esto se debe a la exposición del triptófano al disolvente como consecuencia de la monomerización. El valor de pK_a de la transición ácida determinado a partir del máximo de emisión fue de 4.2 ± 0.2 , y el determinado a partir de la energía de emisión promedio de 4.1 ± 0.2 , ambos similares a los valores obtenidos a mayor concentración.

2.2.1.2. Fluorescencia del ANS.

La unión a ANS se usó para seguir la exposición al disolvente de las regiones hidrofóbicas de la proteína, y detectar la presencia de conformaciones parcialmente plegadas (Stryer, 1965; Semisotnov et al., 1991), tras la desnaturalización inducida por pH. El experimento se realizó con una concentración de CA-C de 2 μM , a distintos pHs, y en presencia de 100 μM de ANS, registrándose los cambios en el espectro de fluorescencia del ANS, tanto en la longitud de onda de máxima emisión como en la energía de emisión promedio.

En presencia de CA-C a bajo pH se observó un aumento en la intensidad de emisión fluorescente del ANS, y el espectro tenía un máximo de emisión a 510 nm (Fig. 2.2). Esto sugiere que CA-C une a ANS, pero esta unión no es lo suficientemente

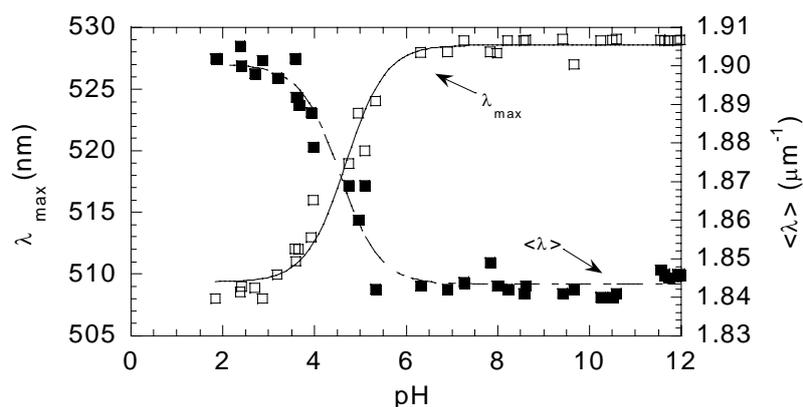


Figura 2.2 Desnaturalización por pH seguida mediante unión a ANS. Comportamiento de la longitud de onda de máxima emisión (cuadrados blancos, eje izquierdo) y de la energía de emisión promedio (cuadrados negros, eje derecho) frente al pH. La concentración de CA-C fue de 2 μM . Las líneas son los ajustes a la ecuación (2.5). Los experimentos se realizaron a 298 K.

fuerte, ya que cuando el ANS se encuentra totalmente unido su máximo se sitúa a 480 nm. Conforme se fue aumentando el pH, se produjo una disminución en la intensidad de emisión y un desplazamiento del máximo hacia 530 nm. La energía de emisión promedio mostró un comportamiento consistente con el anterior, ya que su valor fue disminuyendo al aumentar el pH (Fig. 2.2). Para ambos observables se obtuvo un pK_a de 4.4 ± 0.2 , similar, dentro del error experimental, a los valores determinados a partir de la fluorescencia intrínseca de la proteína.

2.2.1.3. Quenching de fluorescencia de triptófano y tirosina.

Para evaluar la accesibilidad al disolvente de los residuos de triptófano y tirosina de la proteína, se realizaron experimentos de *quenching* de fluorescencia por yoduro y acrilamida a distintos pHs. Los pHs se escogieron en base a los resultados obtenidos en los experimentos de fluorescencia en el estado estacionario (Fig. 2.1): pH 3, en el que CA-C se encuentra en la región ácida; pH 7, entre la titulación ácida y la básica; y, pH 11, en la región básica.

Los datos obtenidos en los experimentos de *quenching* por acrilamida, a una longitud de onda de excitación de 295 nm, se ajustaron siguiendo un comportamiento exponencial, excepto a pH neutro. El valor de la constante de Stern-Volmer, K_{sv} , permaneció constante, dentro del error, a pHs ácidos y básicos, pero fue menor a pH neutro (Tabla 2.1), lo que sugiere que el triptófano de la proteína está más expuesto al disolvente a pHs extremos. Los experimentos control, que se realizaron en presencia de 5 M de GdmHCl a pH 7, condiciones en las que proteína está totalmente desnaturalizada (Mateu, 2002), dieron lugar a valores de K_{sv} similares a los obtenidos a pHs ácidos y básicos. El valor del parámetro ν fue similar tanto a pHs extremos como en presencia de 5 M de GdmHCl, pero a pH 7 fue nulo, indicando la ausencia de un componente estático en el *quenching*.

En los experimentos en los que se excitó a 280 nm, el valor de K_{sv} siguió la misma tendencia descrita, siendo mayor a pHs extremos y en presencia de 5 M de GdmHCl que a pH neutro (Tabla 2.1). El valor de ν permaneció constante, dentro del error, a pHs extremos y en presencia de 5 M de GdmHCl, siendo mayor que a pH neutro. De forma general, los valores de K_{sv} y de ν fueron mayores cuando se excitó a 280 nm que cuando se hizo a 295 nm, probablemente debido a que a 280 nm se excitan tanto el triptófano como la tirosina, de modo que se está observando el *quenching* de

ambos residuos, mientras que a 295 nm sólo se observa el *quenching* del triptófano.

En los experimentos en los que se utilizó el yoduro como *quencher*, los datos se ajustaron siguiendo un comportamiento lineal, y se obtuvieron resultados similares a ambas longitudes de onda de excitación (Tabla 2.1): el valor de K_{sv} fue mayor a pHs ácidos y en presencia de 5 M de GdmHCl que a pH neutro o básico. La similitud del valor de K_{sv} a pH fisiológico, en el que la proteína se encuentra totalmente plegada, y a pH básico, en el que la proteína se encuentra desnaturalizada (como se vio por fluorescencia en el estado estacionario, Fig. 2.1), se puede explicar por el establecimiento de interacciones electrostáticas entre el ión yoduro (I^-) y la proteína (a pHs básicos su carga neta es negativa). El valor de K_{sv} fue mayor a pH 3 que en presencia de 5 M de GdmHCl; este hecho puede deberse a la descomposición parcial del $Na_2S_2O_3$, que es inestable a pH bajo (Rubinson y Rubinson, 1998), y esto conlleva a la formación de I_3^- , que absorbe intensamente a 280 nm, y por tanto se reduce la intensidad de la luz usada en la excitación. No se observaron diferencias en los valores de K_{sv} obtenidos a 280 y 295 nm, como ocurre para la acrilamida, probablemente debido: (i) al mayor tamaño de ésta con respecto al yoduro; o, (ii) a la carga negativa del ión yoduro (I^-).

En conclusión, los experimentos realizados con ambos *quenchers* sugieren que los residuos de triptófano y tirosina de la proteína están más expuestos al disolvente a pHs extremos.

Tabla 2.1 Constantes de *quenching* de CA-C por acrilamida y KI, a distintas longitudes de onda de excitación^a.

	Acrilamida				KI	
	280 nm		295 nm		280 nm	295 nm
	$K_{sv} (M^{-1})$	$\nu (M^{-1})$	$K_{sv} (M^{-1})$	$\nu (M^{-1})$	$K_{sv} (M^{-1})$	$K_{sv} (M^{-1})$
pH 3	5.5 ± 0.6	1.2 ± 0.2	5.4 ± 0.2	0.67 ± 0.06	3.3 ± 0.2	3.4 ± 0.2
pH 7	4.4 ± 0.3	0.6 ± 0.1	3.2 ± 0.5	0.0 ± 0.1	0.9 ± 0.1	0.8 ± 0.1
pH 11	6.1 ± 0.2	1.2 ± 0.1	5.0 ± 0.8	0.7 ± 0.2	1.4 ± 0.1	1.4 ± 0.1
5 M GdmHCl	5.9 ± 0.6	1.0 ± 0.2	4.9 ± 0.6	0.4 ± 0.2	2.6 ± 0.2	2.7 ± 0.2

^aLos errores indicados son errores del ajuste a la ecuación (2.1) de la intensidad de fluorescencia a 338 nm con (acrilamida) o sin (KI) el término exponencial. Se obtuvieron resultados similares a 335, 336 y 337 nm. Todos los experimentos se realizaron con una concentración de CA-C de 20 μ M, y a 298 K. Los experimentos con 5 M de GdmHCl se hicieron a pH 7.

2.2.1.4. Dicroísmo circular en el ultravioleta lejano.

Los espectros de DC en el ultravioleta lejano de CA-C, tanto a 20 como a 200 μM , mostraron sendos mínimos a 208 y 222 nm, característicos de estructura en hélice α , y de acuerdo con su estructura tridimensional (Gamble et al., 1997). No se apreciaron cambios en la forma del espectro a distintos pHs, lo que sugiere que la estructura secundaria no se altera significativamente; sin embargo, el valor de $[\Theta]_{222}$, sí que fue dependiente del pH (a 208 nm se observó la misma tendencia, pero la relación señal a ruido fue menor).

Para una concentración de CA-C de 200 μM y en el intervalo de pH 7 a 10, el valor de $[\Theta]_{222}$ fue constante, en torno a $-13000 \text{ deg cm}^2 \text{ d mol}^{-1}$ (Fig. 2.3 A). Conforme el pH fue aumentando, el valor absoluto de la elipticidad fue haciéndose progresivamente más pequeño, pero no se pudo determinar el pK_a de esta titulación a causa de la ausencia de línea de base a pHs básicos. Por otro lado, al ir disminuyendo el pH por debajo de pH neutro, también se observó un decrecimiento del valor absoluto de la elipticidad, mostrando un comportamiento sigmoidal, con un pK_a de 4.5 ± 0.2 .

Los experimentos realizados con una concentración de proteína de 20 μM (Fig. 2.3 B), mostraron la misma variación de $[\Theta]_{222}$ con el pH que la observada a 200 μM , con un pK_a de 4.4 ± 0.3 para la transición ácida; este valor es idéntico, con error experimental, al obtenido a 200 μM , y similar a los valores observados por otras técnicas.

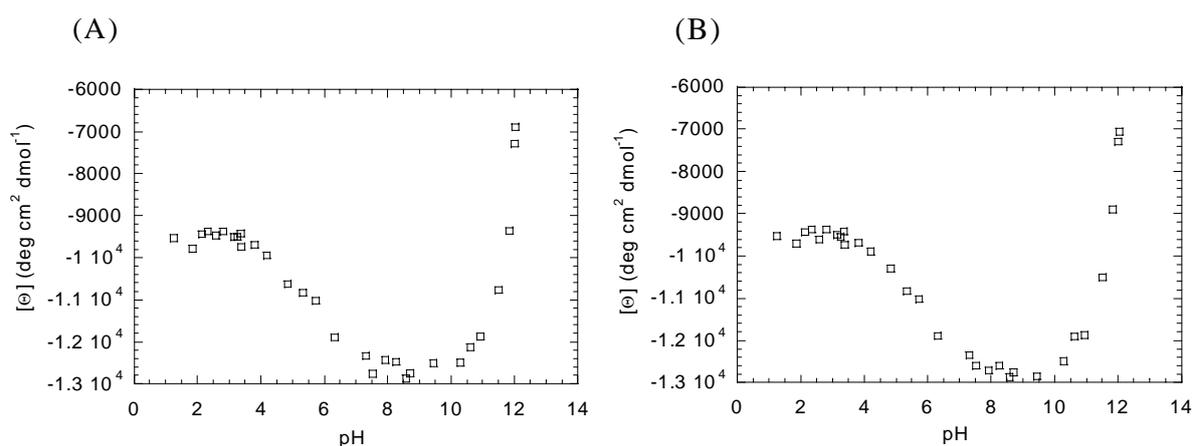


Figura 2.3 Desnaturalización por pH seguida mediante DC en el ultravioleta lejano. Elipticidad molar por residuo a 222 nm frente al pH, a dos concentraciones distintas de CA-C: (A) 200 μM , y (B) 20 μM . Los experimentos se realizaron a 298 K.

2.2.1.5. Cromatografía de exclusión por tamaño molecular.

En los experimentos de cromatografía de exclusión por tamaño molecular de CA-C se obtuvo un solo pico a cualquier pH y a cualquier concentración de proteína utilizada (20 o 200 μM). Esto indica que el equilibrio entre la forma monomérica y dimérica de CA-C es rápido comparado con el tiempo de elución (Corbett y Roche, 1984), por lo tanto, las diferencias en el V_e entre ambas concentraciones (Fig. 2.4) deben reflejar el promedio de ambas especies (monomérica y dimerica) a las concentraciones utilizadas.

En líneas generales, tanto a 20 como a 200 μM de CA-C, se observó el mismo comportamiento cuando se siguió el V_e frente al pH (Fig. 2.4): el V_e fue aumentando a medida que disminuía el pH, lo que sugiere que la proteína (en su forma monomérica o dimérica) posee menor radio hidrodinámico a pHs bajos. El V_e a pHs ácidos es muy distinto del volumen de exclusión molecular de la columna (8.0 ml), indicando que las especies que se observan a pH ácido no son oligómeros. Para esta titulación ácida, los valores de pK_a obtenidos fueron de 4.3 ± 0.1 (a 20 μM de CA-C), y de 4.1 ± 0.1 (a 200 μM de CA-C), que son similares, dentro del error experimental, a los obtenidos por otras técnicas. También se observó que tanto a pHs altos (por encima de pH 11) como a pHs bajos (por debajo de pH 2.5), las curvas a ambas concentraciones tienden a unirse.

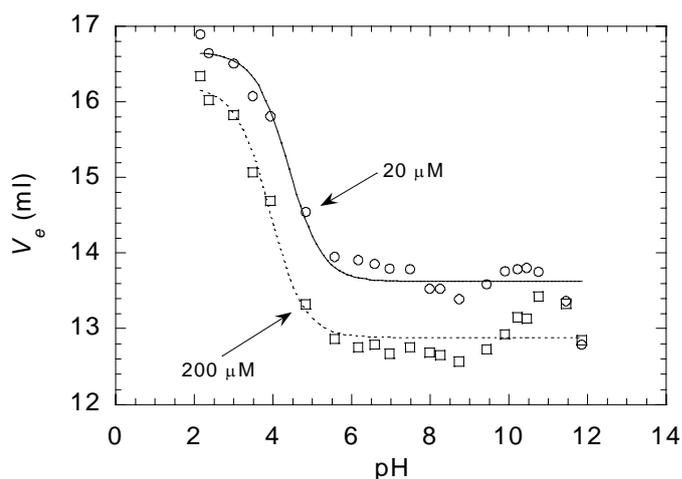


Figura 2.4 Desnaturalización por pH seguida mediante cromatografía de exclusión por tamaño molecular. V_e de CA-C a 20 μM (círculos blancos) y a 200 μM (cuadrados blancos) frente al pH. Las líneas (continua para 20 μM y discontinua para 200 μM de CA-C) son los ajustes a la ecuación (2.5). Los experimentos se realizaron a 298 K.

2.2.1.6. Infrarrojo por transformada de Fourier.

La técnica de FTIR proporciona información acerca de las cadenas laterales de algunos aminoácidos, entre ellos la tirosina. El residuo de tirosina, a pHs ácidos y neutros, presenta una banda de absorción con un máximo centrado en 1515 cm^{-1} , y cuando pierde su protón fenólico a pH básico convirtiéndose en tirsinato, su banda de absorción se desplaza hacia frecuencias menores (1505 cm^{-1}) (Surewicz et al., 1993; de Jongh et al., 1996).

En los experimentos de FTIR se observó que a pH ácido el máximo de la banda de absorción de la tirosina se situó próximo a 1515 cm^{-1} (Fig. 2.5), y con el aumento de pH, esta banda se desplazó ligeramente hacia un valor de frecuencia mayor, dando lugar a una transición con un pK_a de 4.3 ± 0.2 . Este aumento del máximo de la banda de absorción de la tirosina con el pH sugiere que la titulación se debe a la desprotonación de algún residuo ácido situado cerca de al menos uno de los dos residuos de tirosina de CA-C (Tyr164 o Tyr169).

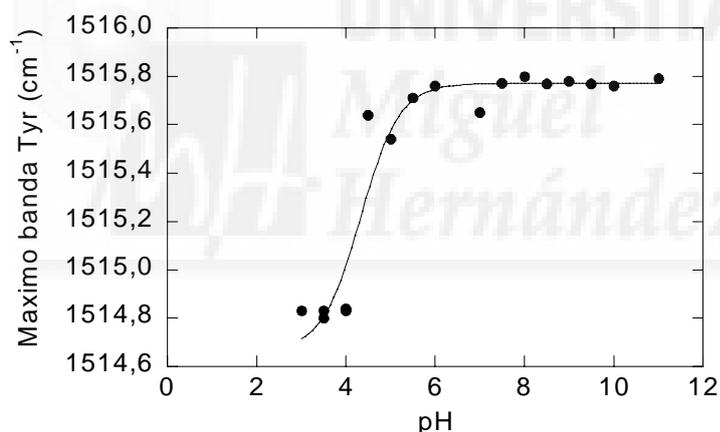


Figura 2.5 Desnaturalización por pH seguida mediante FTIR. Máximo de la banda de absorción de la tirosina frente al pH. La línea es el ajuste a la ecuación (2.5). Los experimentos se realizaron con una concentración de CA-C de 1 mM y a 278 K.

2.2.2. DESNATURALIZACIÓN TÉRMICA.

Con el fin de obtener los parámetros termodinámicos que caracterizan la reacción de plegamiento de CA-C, se realizaron experimentos en los que se siguió la desnaturalización térmica de la proteína mediante diversas técnicas biofísicas.

2.2.2.1. Fluorescencia intrínseca de triptófano y tirosina.

Las desnaturalizaciones térmicas seguidas por la emisión de fluorescencia de CA-C a 200 μM en el rango de pH 5 a 9 dieron lugar a una transición sigmoideal completamente reversible (Fig. 2.6; Tabla 2.2). Sin embargo, a pH 10 la transición no fue reversible, y a pH 4 se observó un decrecimiento lineal de la intensidad de fluorescencia con la temperatura (Fig. 2.6). En los experimentos realizados con una concentración de CA-C de 20 μM se observó, para la mayoría de los pHs estudiados, la ausencia de línea de base en la región de la proteína nativa, probablemente debido a la gran cantidad de monómero presente. Esta ausencia de línea de base impidió la determinación de los parámetros termodinámicos de la desnaturalización térmica y, en consecuencia, no se pudo concluir si la transición térmica seguida por fluorescencia tenía un comportamiento dependiente de la concentración de proteína. Por lo tanto, sólo se pudo realizar el análisis de van't Hoff de las curvas de desnaturalización obtenidas a 200 μM de CA-C en el rango de pH 5 a 9, cuyos termogramas mostraron líneas de base largas para la proteína tanto en la región nativa como en la desnaturalizada; sin embargo, debido a la gran pendiente de ambas líneas de base, no se pudo calcular el valor de ΔC_p . No obstante, las grandes pendientes no afectan a la determinación de T_m , ya que puede calcularse a partir de la segunda derivada de la función. En general, se observó que la T_m disminuyó a pHs muy ácidos y básicos, es decir, la termoestabilidad de CA-C disminuyó a pHs extremos (Tabla 2.2).

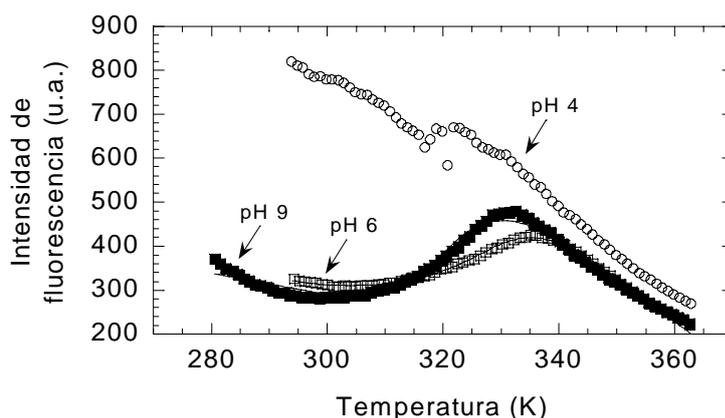


Figura 2.6 Desnaturalización térmica seguida mediante fluorescencia. Intensidad de fluorescencia a 350 nm frente a la temperatura para CA-C a 200 μM , y a distintos pHs: pH 4 (círculos blancos), pH 6 (cuadrados blancos) y pH 9 (cuadrados negros). La escala en el eje de ordenadas es arbitraria. Las líneas son los ajustes a la ecuación (2.6), teniendo en cuenta que la energía libre viene dada por la ecuación (2.7).

Tabla 2.2 Temperaturas medias de desnaturalización (T_m) (K) de CA-C a distintos pHs, obtenidas mediante diversas técnicas biofísicas^a.

Técnica	Condiciones										Concentración de CA-C (μM)
	pH										
	3	4	5	6	7	8	9	10			
Fluorescencia		ND	324.0 \pm 0.2	325.7 \pm 0.3	325.5 \pm 0.2	324.5 \pm 0.3	324.8 \pm 0.5	NR		200	
Unión a ANS		ND	ND	318 \pm 1	325 \pm 1	325 \pm 6	322 \pm 2	ND		20	
Anisotropía					325.2 \pm 0.5					80	
					325 \pm 4 ^b					20	
					326 \pm 2					200	
DC en el UV-lejano		328 \pm 3	333.0 \pm 0.3	332.5 \pm 0.5	332.1 \pm 0.3	331.7 \pm 0.3	331.5 \pm 0.3	NR		20	
		330 \pm 1	331.0 \pm 0.5	334.6 \pm 0.3	332.6 \pm 0.6	332.6 \pm 0.5	331.8 \pm 0.5	NR		200	
DC en el UV-cercano		ND			326 \pm 5 ^b	326 \pm 7 ^b				20	
		ND			327 \pm 0.9	327.5 \pm 0.9				200	
Absorbancia		331 \pm 0.8	330.5 \pm 0.5			334 \pm 0.7	328.6 \pm 0.7			200	
FTIR					304.5 \pm 0.6 ^c					530	
		NR	NR	304.7 \pm 0.3	307.2 \pm 0.8 ^c	301.2 \pm 0.6	NR			1000	
					307.1 \pm 0.8 ^c					1100	
					307.9 \pm 0.7 ^c					2000	
					309.8 \pm 0.8 ^c					4500	
RMN					334 \pm 5 ^e					200	

^aLos errores provienen del ajuste a la ecuación (2.6), teniendo en cuenta que la energía libre viene dada por la ecuación (2.7) para todas las técnicas, excepto para FTIR que viene dada por la ecuación (2.8). Las desnaturalizaciones térmicas seguidas por fluorescencia se repitieron tres veces a cada pH, y la desviación estándar de las tres medidas fue ± 0.6 K. Las desnaturalizaciones térmicas seguidas por la unión a ANS se repitieron tres veces a cada pH, y la desviación estándar fue ± 3 K. Cada valor de anisotropía obtenido es la media de cuatro mediciones independientes para cada temperatura, con una desviación estándar de ± 0.4 K; cada experimento de desnaturalización térmica se repitió dos veces a cada concentración, y la desviación estándar fue ± 4 K. Los experimentos seguidos por DC en el ultravioleta lejano y cercano, y por absorbancia, se repitieron dos veces a cada pH, y la desviación estándar fue ± 0.4 K, para DC en el ultravioleta lejano; ± 3 K (20 μM), y ± 0.5 K (200 μM), para DC en el ultravioleta cercano; ± 0.5 K, para absorbancia. Los experimentos seguidos por FTIR y RMN, sólo se adquirieron una vez, debido a la gran cantidad de proteína necesaria para dichas medidas. ND, sin determinar; NR, transición térmica no reversible.

^bLas transiciones a 20 μM de CA-C tienen grandes errores asociados al ajuste de la ecuación (2.6), debido a la baja relación señal a ruido.

^c T_m s de la transición a baja temperatura observada por FTIR.

^dLa T_m se ha determinado a partir de la señal del indol.

^eLa T_m se ha determinado a partir de la señal de un metilo, que posee pocos puntos experimentales en la línea de base de la región de la proteína desnaturalizada.

2.2.2.2. Fluorescencia del ANS.

Las desnaturalizaciones térmicas seguidas mediante la unión a ANS de CA-C en el rango de pH 6 a 9 dieron lugar a una transición sigmoideal completamente reversible (Tabla 2.2). Los valores de T_m no fueron dependientes de la concentración de proteína, entre 20 y 80 μM de CA-C a pH 7. No se pudo calcular el valor de ΔC_p debido a los grandes errores asociados al ajuste de ΔH_m . La termoestabilidad de CA-C disminuyó, sin embargo, a pHs extremos.

Los experimentos que se realizaron a pH 4, 5, 10 y 11, no mostraron transiciones sigmoideales (Fig. 2.7). A pH 4 y 5, condiciones en las que CA-C une ANS a 298 K (Fig. 2.2), se observó que la intensidad de emisión fluorescente fue decreciendo desde 290 hasta 320 K, y desde esa temperatura permaneció constante hasta 360 K. Esta curva, probablemente, se corresponde con la región final de una transición sigmoideal, y su comportamiento sugiere que las regiones hidrofóbicas de la proteína, que a 298 K están lo suficientemente próximas entre sí para permitir la unión de ANS, se alejan unas de otras cuando la proteína se desnatura. Por el contrario, a pH 10 y 11, condiciones en las que CA-C no une ANS a 298 K (Fig. 2.2), se observó que la intensidad de emisión fluorescente fue aumentando linealmente con la temperatura; esto es probablemente debido a que la proteína sufrió modificaciones a consecuencia de las altas temperaturas y pHs muy alcalinos.

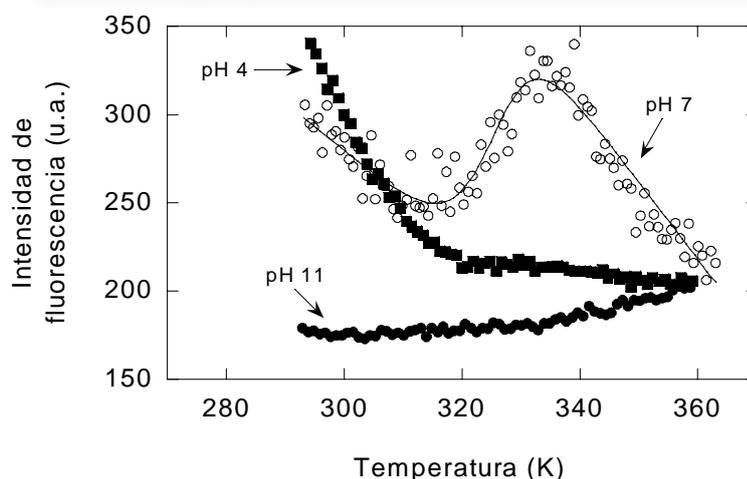


Figura 2.7 Desnaturalización térmica seguida mediante unión a ANS. Intensidad de fluorescencia a 520 nm frente a la temperatura para CA-C a 20 μM , con 100 μM de ANS, y a distintos pHs: pH 4 (cuadrados negros), pH 7 (círculos blancos) y pH 11 (círculos negros). La escala en el eje de ordenadas es arbitraria. La línea es el ajuste a la ecuación (2.6), teniendo en cuenta que la energía libre viene dada por la ecuación (2.7).

2.2.2.3. Anisotropía.

Los experimentos de desnaturalización térmica seguidos por el cambio en la anisotropía se realizaron a dos concentraciones de proteína: 20 y 200 μM , a pH 7 (Fig. 2.8). Los valores de T_m no fueron dependientes de la concentración de proteína, y fueron similares a los obtenidos por fluorescencia y unión a ANS (Tabla 2.2).

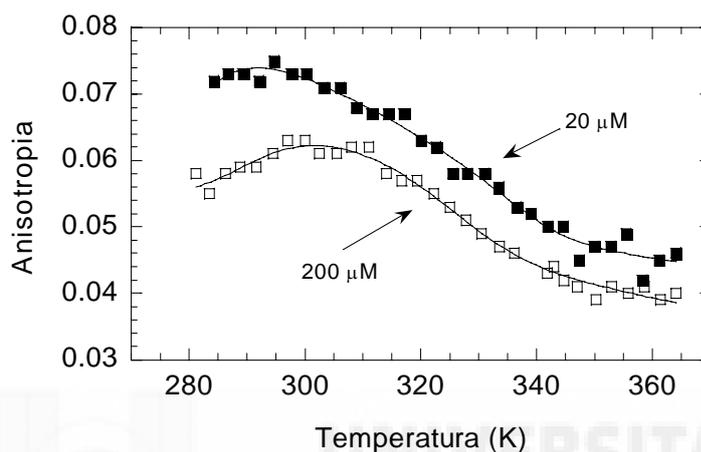


Figura 2.8 Desnaturalización térmica seguida mediante anisotropía. Comportamiento de la anisotropía frente a la temperatura para CA-C a 20 μM (cuadrados negros) y 200 μM (cuadrados blancos), a pH 7. Las líneas son los ajustes a la ecuación (2.6), teniendo en cuenta que la energía libre viene dada por la ecuación (2.7).

2.2.2.4. Dicroísmo circular en el ultravioleta lejano.

Las desnaturalizaciones térmicas de CA-C seguidas midiendo la variación de la elipticidad a 222 nm con la temperatura en el rango de pH 4 a 9 (Fig. 2.9), mostraron un comportamiento sigmoideal, completamente reversible y no dependiente de la concentración (Tabla 2.2), como se había observado previamente a pH 7.3 (Mateu, 2002). El hecho de que las transiciones no fueran dependientes de la concentración es indicativo de que se está observando la desnaturalización de la forma monomérica de CA-C. Sólo a pH 10 la transición térmica no fue reversible, al igual que ocurría por fluorescencia, y probablemente fue debido a que la proteína sufrió modificaciones a consecuencia de las altas temperaturas y pH muy alcalino. La termoestabilidad de CA-C disminuyó a pHs extremos (al igual que ocurría por fluorescencia y unión a ANS), y fue superior a la obtenida mediante fluorescencia, unión a ANS y anisotropía (Tabla 2.2). Con estos datos de T_m a diferentes pHs fue

posible obtener, a partir de ΔH_m y T_m , el valor del cambio en la capacidad calorífica de la reacción de desplegamiento. Así, se obtuvo un valor de ΔC_p de $1.8 \pm 0.5 \text{ kcal mol}^{-1} \text{ K}^{-1}$ (Fig. 2.10).

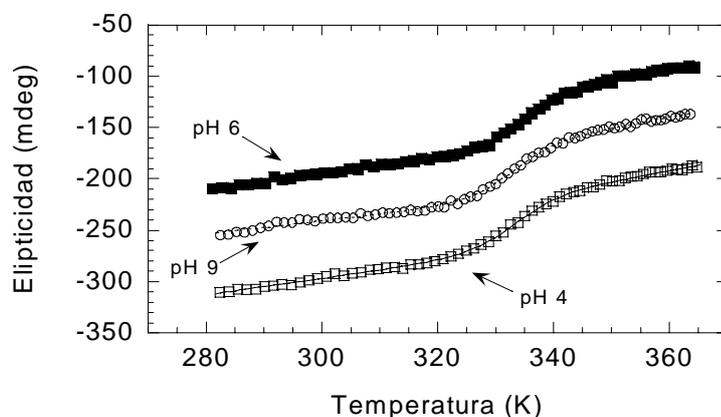


Figura 2.9 Desnaturalización térmica seguida mediante DC en el ultravioleta lejano. Elipticidad experimental a 222 nm frente a la temperatura para CA-C a 200 μM , y a distintos pHs: pH 4 (cuadrados blancos), pH 6 (cuadrados negros) y pH 9 (círculos blancos). Las líneas son los ajustes a la ecuación (2.6), teniendo en cuenta que la energía libre viene dada por la ecuación (2.7).

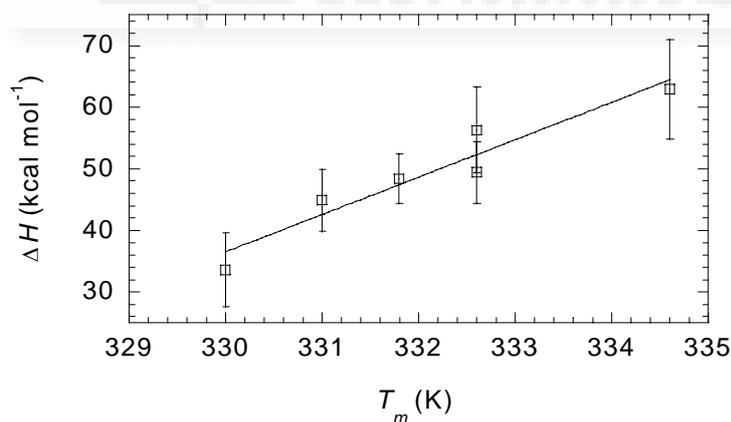


Figura 2.10 Determinación de ΔC_p de la forma monomérica de CA-C a partir de experimentos de desnaturalización térmica seguidos por DC en el ultravioleta lejano. Comportamiento de ΔH_m frente a T_m para CA-C a 200 μM . Las barras de error son los errores del ajuste de los datos experimentales representados en la Fig. 2.9 a la ecuación (2.7).

2.2.2.5. Dicroísmo circular en el ultravioleta cercano y absorbancia.

Se realizaron experimentos de desnaturalización térmica de CA-C a dos concentraciones diferentes (20 y 200 μM) y a tres pHs seleccionados (pH 4, 7 y 8), siguiendo el cambio de la elipticidad a 280 nm con la temperatura. Las desnaturalizaciones térmicas de CA-C a 200 μM (Fig. 2.11 A; Tabla 2.2) y a pH 7 y 8 mostraron una transición sigmoidal completamente reversible, mientras que a pH 4 no se pudieron determinar los valores de ΔH_m ni de T_m , debido a la gran pendiente de la línea de base en la región de la proteína nativa. La estabilidad térmica de CA-C medida por DC en el ultravioleta cercano fue inferior a la obtenida en el ultravioleta lejano, y similar a la obtenida por fluorescencia, unión a ANS y anisotropía.

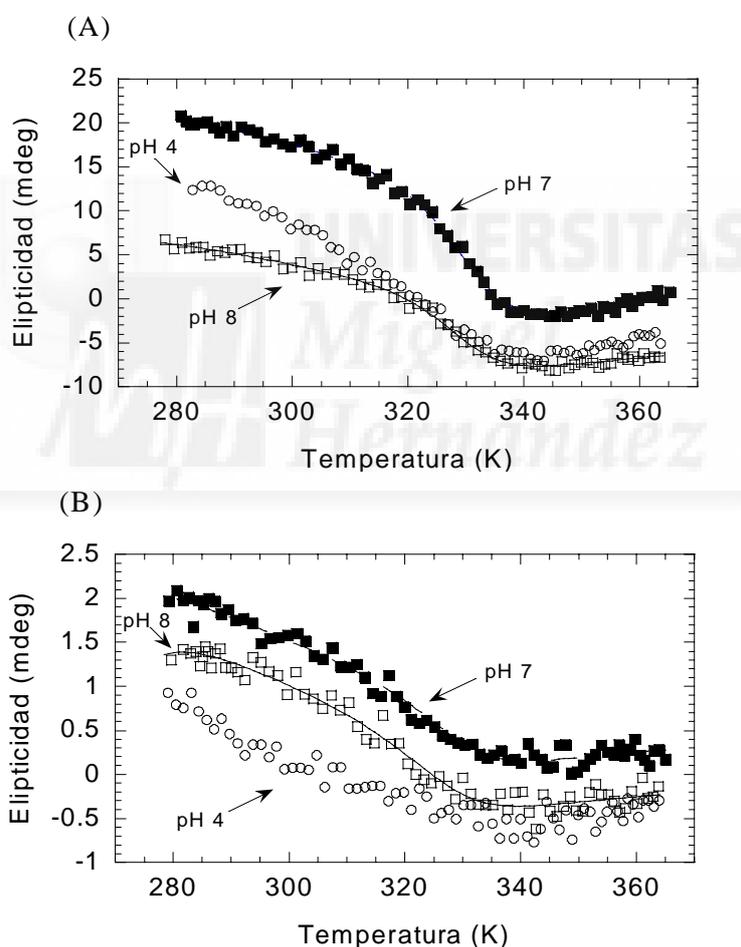


Figura 2.11 Desnaturalización térmica seguida mediante DC en el ultravioleta cercano.

Elipticidad experimental a 280 nm frente a la temperatura a distintos pHs: pH 4 (círculos blancos), pH 7 (cuadrados negros) y pH 8 (cuadrados blancos), y a dos concentraciones de CA-C: (A) 200 μM y (B) 20 μM . Las líneas son los ajustes a la ecuación (2.6), teniendo en cuenta que la energía libre viene dada por la ecuación (2.7).

Cuando los experimentos se realizaron con una concentración de proteína de 20 μM (Fig. 2.11 B; Tabla 2.2) y a cualquiera de los pHs seleccionados, no se pudieron determinar valores fiables para ΔH_m , y los valores obtenidos de T_m tenían asociados grandes errores a consecuencia de las pendientes de las líneas de base en la región de la proteína nativa, probablemente debidas a la gran cantidad de monómero. Por todo ello, a pesar de que los valores de T_m sugieren un comportamiento independiente de la concentración de CA-C, esta conclusión debe aceptarse con cuidado.

También se realizaron experimentos de desnaturalización térmica en el ultravioleta cercano siguiendo la variación de la absorbancia a 280 nm con la temperatura, de CA-C a 200 μM y pHs 4, 5, 8 y 9 (Fig. 2.12; Tabla 2.2). En todos los casos se obtuvieron transiciones sigmoidales completamente reversibles. La estabilidad térmica de CA-C medida por absorbancia fue similar a la obtenida por DC en el ultravioleta lejano. Debido a la falta de suficiente línea de base en la región de la proteína desnaturalizada, no se pudieron determinar valores fiables de ΔH_m , y esto impidió el cálculo de ΔC_p . No se pudo seguir la desnaturalización de CA-C a 20 μM midiendo la absorbancia a 280 nm debido a la baja relación señal a ruido.

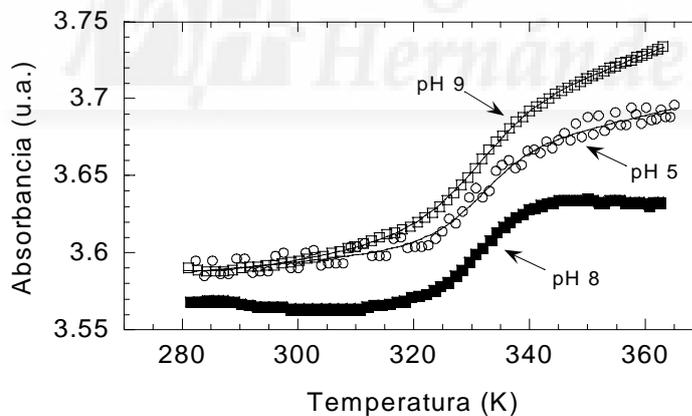


Figura 2.12 Desnaturalización térmica seguida mediante absorbancia en el ultravioleta cercano. Absorbancia a 280 nm frente a la temperatura para CA-C a 200 μM , y a distintos pHs: pH 5 (círculos blancos), pH 8 (cuadrados negros) y pH 9 (cuadrados blancos). La escala en el eje de ordenadas es arbitraria y las curvas de desnaturalización térmica se han desplazado unas respecto a otras para una mayor claridad. Las líneas son los ajustes a la ecuación (2.6), teniendo en cuenta que la energía libre viene dada por la ecuación (2.7).

2.2.2.6. Infrarrojo por transformada de Fourier.

Los experimentos de desnaturalización térmica seguidos por FTIR que se realizaron con una concentración de proteína de 1 mM a pH 7, mostraron dos transiciones diferentes (Fig. 2.13):

(A) Una transición con una $T_m \sim 307$ K, que no se había observado por ninguna de las técnicas utilizadas con anterioridad, y que tenía lugar con un pequeño cambio en la anchura de la banda. Esta transición sólo fue reversible cuando se calentó hasta temperaturas no superiores a 325 K.

(B) Una transición con una $T_m \sim 335$ K, similar a la obtenida por DC en el ultravioleta lejano y absorbancia (Tabla 2.2), y que tenía lugar con la mayoría del cambio en la anchura de la banda. Esta transición no fue reversible, debido a las modificaciones sufridas por la proteína a las concentraciones tan altas y las elevadas temperaturas.

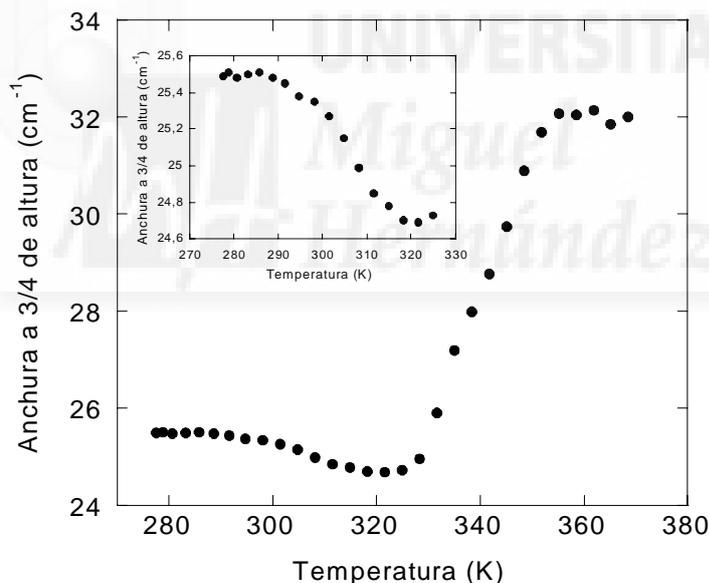


Figura 2.13 Desnaturalización térmica seguida mediante FTIR. Anchura a $\frac{3}{4}$ de altura de la banda amida I' frente a la temperatura para CA-C a 1 mM y pH 7. En la figura insertada se muestra una expansión de la transición a baja temperatura.

Como la desnaturalización térmica que ocurría a alta temperatura ($T_m \sim 335$ K) parecía ser la misma que la observada por otras técnicas (DC en el ultravioleta lejano y absorbancia), se llevó a cabo un estudio de la primera transición con el pH (en el rango de pH 4 a 9), pero calentando sólo hasta 325 K y utilizando una concentración de CA-C

de 1 mM (Fig. 2.14 A; Tabla 2.2). Únicamente se observó reversibilidad a pH 6, 7 y 8. Asimismo, se observó una dependencia de T_m con la concentración de proteína usada a pH 7 (Fig. 2.14 B; Tabla 2.2), ya que el valor de T_m fue aumentando con la concentración de CA-C.

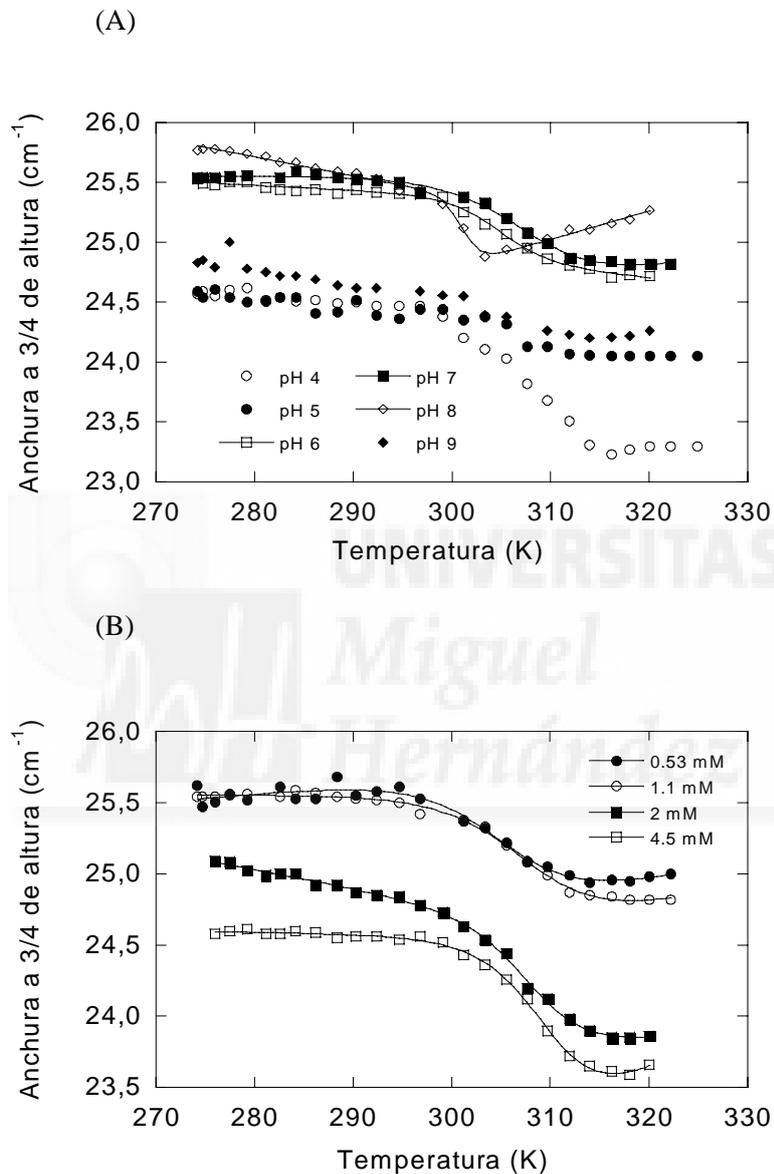


Figura 2.14 Desnaturalización térmica seguida mediante FTIR. (A) Anchura a $\frac{3}{4}$ de altura de la banda amida I' frente a la temperatura para la transición a baja temperatura de CA-C a 1 mM y a distintos pHs (indicados en la leyenda). (B) Anchura a $\frac{3}{4}$ de altura de la banda amida I' frente a la temperatura para la transición a baja temperatura de CA-C a distintas concentraciones (indicadas en la leyenda) y a pH 7. Las líneas son los ajustes a la ecuación (2.6), teniendo en cuenta que la energía libre viene dada por la ecuación (2.8).

La variación de T_m con la concentración de proteína puede utilizarse para determinar los parámetros termodinámicos de la transición (Fig. 2.15). En un sistema oligomérico, la dependencia de T_m con C_t (siendo C_t la concentración molar total de proteína referida a la masa molar del monómero), viene dada por la siguiente ecuación (Marky y Breslauer, 1987; Steif et al., 1993):

$$\frac{1}{T_m} = -\frac{R}{\Delta H^0} \ln(C_t) + \frac{\Delta S^0}{\Delta H^0} \quad (2.9),$$

donde ΔH^0 y ΔS^0 son el cambio en la entalpía y entropía, respectivamente, independientes de la temperatura, y en condiciones estándar, a una temperatura de igual valor que T_m . Para CA-C, los parámetros termodinámicos de la reacción de

disociación fueron (con $\frac{R}{\Delta H^0} = (2.7 \pm 0.4) \times 10^{-5} \text{ K}^{-1}$, y $\frac{\Delta S^0}{\Delta H^0} = (30.7 \pm 0.2) \times 10^{-4} \text{ K}^{-1}$):

$\Delta H^0 = 74 \pm 8 \text{ kcal (mol de unidad cooperativa)}^{-1}$ y $\Delta S^0 = 230 \pm 30 \text{ cal K}^{-1} \text{ (mol de unidad cooperativa)}^{-1}$ a una concentración de proteína de 1 M (condiciones estándar) (Fig. 2.15).

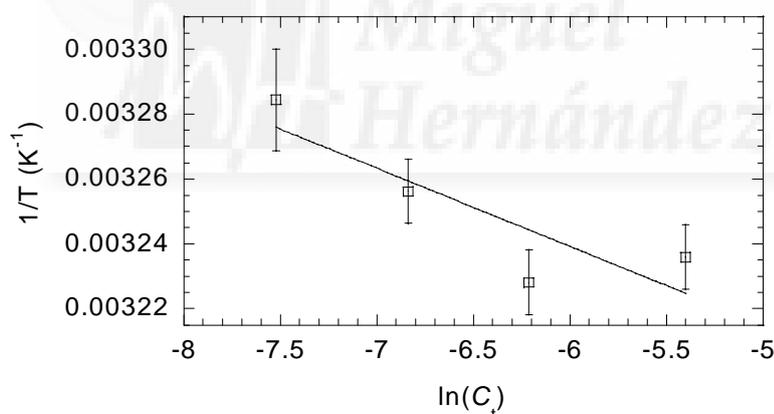


Figura 2.15 Determinación de los parámetros termodinámicos de la forma dimérica de CA-C a pH 7, a partir de experimentos de desnaturalización térmica seguidos por FTIR. Variación de T_m con la concentración de CA-C para la transición a baja temperatura. Las barras de error son los errores del ajuste a la ecuación (2.8). La línea es el ajuste a la ecuación (2.9).

2.2.2.7. Resonancia magnética nuclear.

Se tomaron espectros monodimensionales de RMN, tanto a 298 como a 320 K, con una concentración de proteína de 200 μM , y a dos pHs distintos, que se escogieron en base a los resultados obtenidos previamente: pH 7, en el que CA-C se encuentra

plegada; y, pH 3, en el que la proteína ha perdido parte de su estructura terciaria pero conserva la estructura secundaria (apartados 2.2.1.1. y 2.2.1.4.).

El espectro adquirido a pH 3 y 298 K presentó un mayor número de señales de protones amida agrupadas entre 8.0 y 8.5 ppm que las propias a pH 7; además, sólo había dos señales de metilos a campo alto (Fig. 2.16 C). Este resultado sugiere que CA-C posee menos estructura a pH 3 que a pH 7 (Wüthrich, 1986).

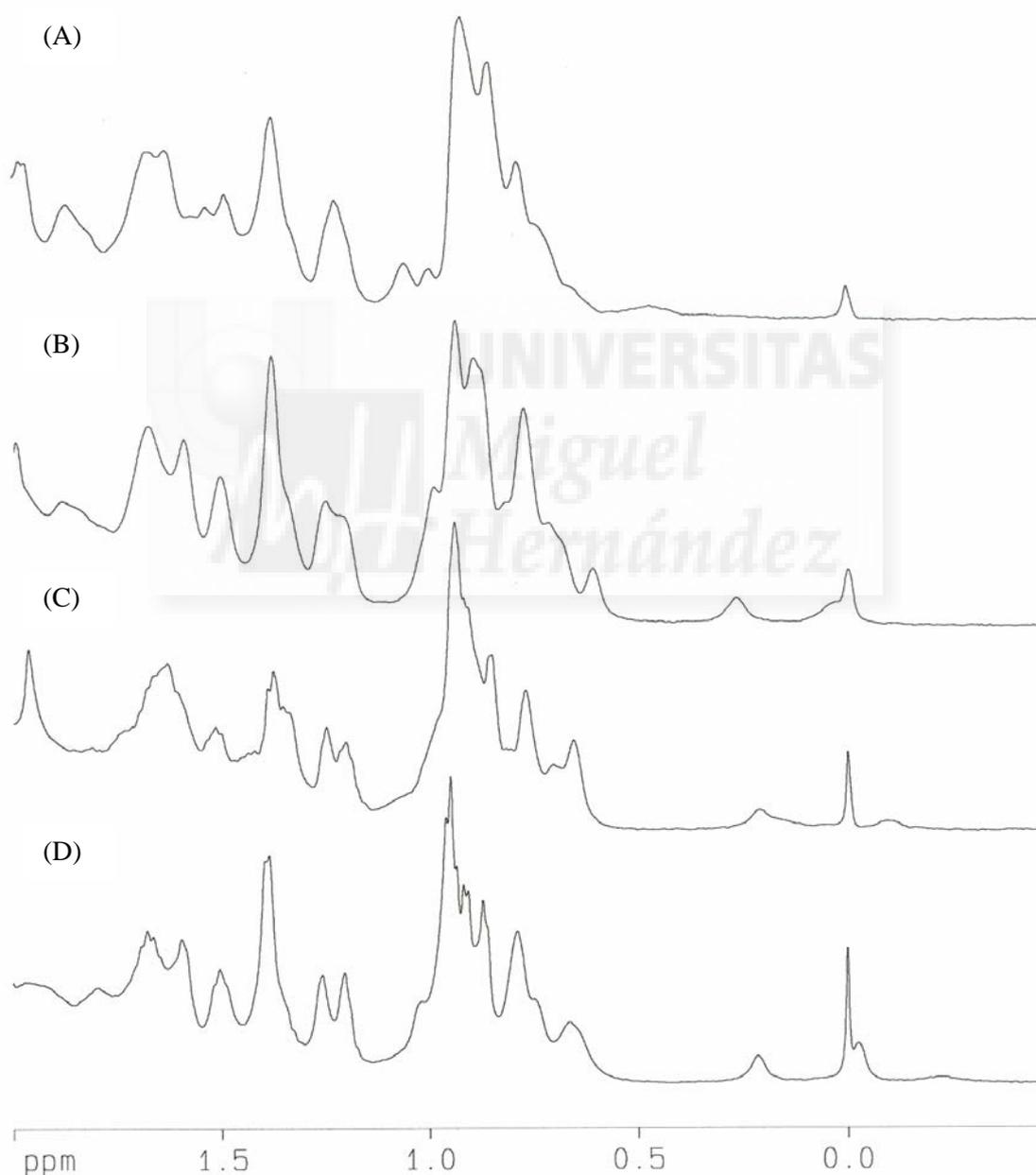


Figura 2.16 Región de campo alto de los protones metilo del espectro de RMN de CA-C a 200 μ M. (A) pH 3 y 320 K; (B) pH 7 y 320 K; (C) pH 3 y 298 K; y, (D) pH 7 y 298 K.

El espectro a pH 7 y 320 K presentó todos los protones amida agrupados entre 8.0 y 8.5 ppm, y la mayoría de los protones metilo a campo alto había desaparecido (Fig. 2.16 B), sugiriendo que la mayor parte de la estructura terciaria se había perdido con el aumento de temperatura. Por otro lado, a pH 3 y 320 K, desaparecieron tanto la pequeña dispersión de los protones amida como la señal de uno de los dos protones metilo a campo alto observados a 298 K (Fig. 2.16 A). Así pues, los experimentos de RMN confirman que CA-C es menos estable a pH ácido.

Para estudiar los cambios con la temperatura, se siguieron las señales de los protones metilo e indol, ya que ambas están bien separadas del resto de protones del espectro.

En los experimentos de desnaturalización térmica de CA-C a 200 μ M y pH 7, los protones, según su comportamiento con la temperatura, se pueden clasificar en tres grupos:

(A) *Clase I*: son los protones a -0.26, -0.20 y -0.037 ppm (Fig. 2.17 A), que fueron visibles a temperaturas bajas (en el intervalo entre 278 y 288 K), pero sus señales desaparecieron a partir de 310, 290 o 300 K, respectivamente.

(B) *Clase II*: es el protón metilo que aparece a 0.2 ppm (Fig. 2.17 B), cuyo desplazamiento químico fue aumentando ligeramente con la temperatura en el intervalo entre 278 y 320 K, y a partir de 320 K mostró un comportamiento sigmoidal. A pesar de que había pocos datos experimentales en la línea de base en la región de la proteína desnaturalizada, se realizó el ajuste a las ecuaciones (2.6) y (2.7), y se obtuvo un valor de T_m de 334 ± 5 K, similar a los obtenidos por DC en el ultravioleta lejano y absorbancia (Tabla 2.2).

(C) *Clase III*: es el protón indol del Trp184 (Fig. 2.17 C), cuyo desplazamiento químico fue disminuyendo con la temperatura, desapareciendo su señal en el intervalo entre 290 y 315 K.

Por el contrario, en las desnaturalizaciones térmicas de CA-C a 200 μ M y pH 3, sólo se observaron dos clases de protones:

(A) *Clase I*: es el protón que aparece a -0.20 ppm (Fig. 2.17 D), cuyo desplazamiento químico fue aumentando con la temperatura hasta 310 K, y a partir de esa temperatura su señal desapareció. Esta señal parece ser la misma que se observa a pH 7 a -0.26 ppm (Fig. 2.17 A), ya que ambas presentan un desplazamiento químico y un comportamiento con la temperatura muy similar.

(B) *Clase II*: compuesta por el protón metilo que aparece a 0.2 ppm (Fig. 2.17 E) y el protón indol del Trp184 (Fig. 2.17 F). El protón metilo probablemente es el mismo que se observó a pH 7 (Fig. 2.17 B). La señal del protón indol del triptófano a pH 3 fue visible en todo el rango de temperatura estudiado, al contrario de lo observado a pH 7 (Fig. 2.17 C).

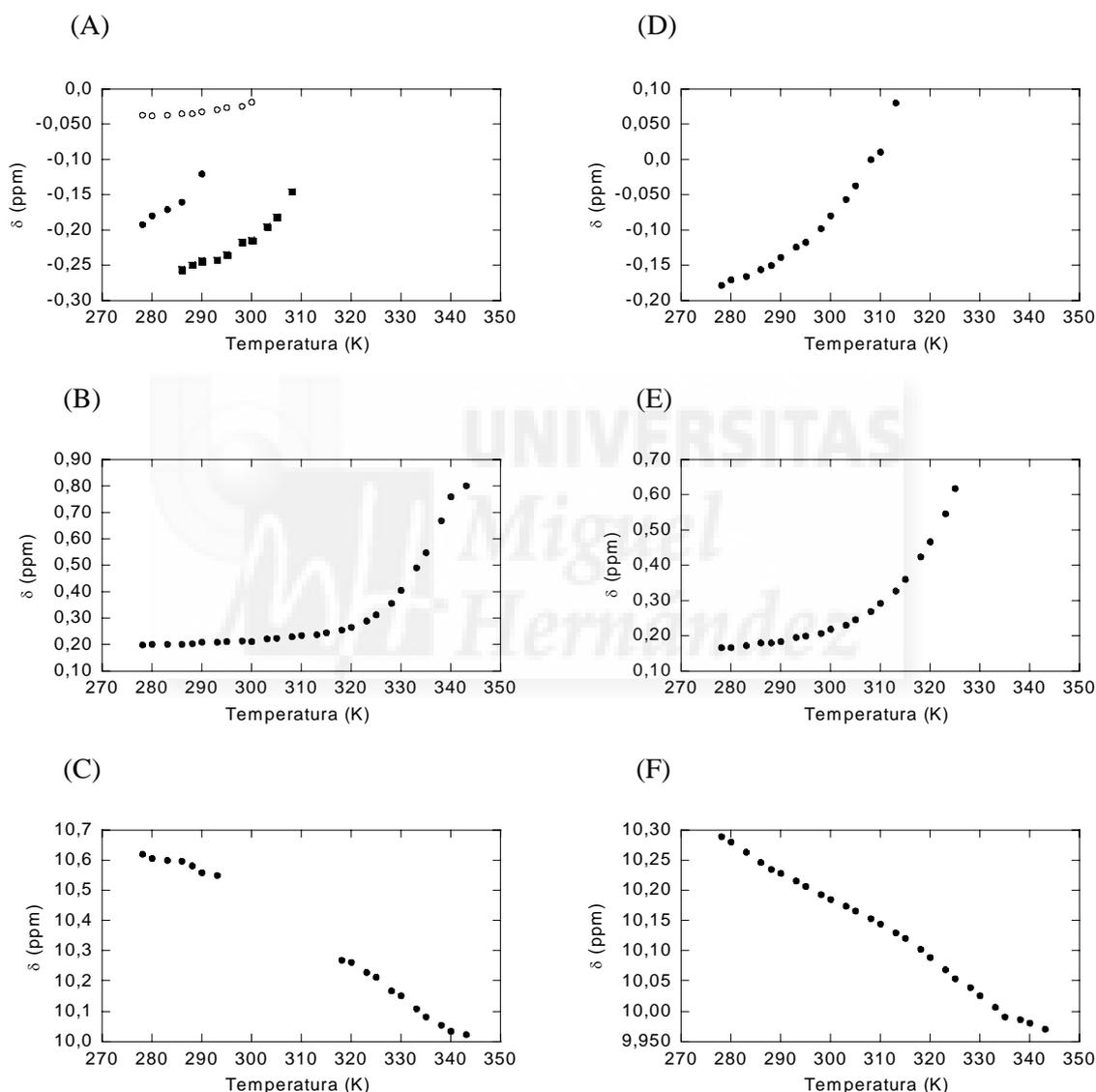


Figura 2.17 Desnaturalización térmica seguida mediante RMN. Variación de los desplazamientos químicos de los distintos protones con la temperatura, para CA-C a 200 μ M y a dos pHs distintos. A pH 7: (A) Clase I, (B) Clase II, y (C) Clase III. A pH 3: (D) Clase I, (E) Clase II (protón metilo), y (F) Clase II (protón indol).

2.2.3. ESTABILIDAD CONFORMACIONAL Y PARÁMETROS TERMODINÁMICOS DE LA REACCIÓN DE DESPLEGAMIENTO DE LA FORMA MONOMÉRICA DE CA-C A pH 7.

El valor de ΔC_p que se obtuvo a partir de los experimentos de DC en el ultravioleta lejano ($\Delta C_p = 1.8 \pm 0.5 \text{ kcal mol}^{-1} \text{ K}^{-1}$; Fig. 2.10) es sólo aproximado, a consecuencia de los grandes errores asociados al ajuste de ΔH_m . Por lo tanto, es necesario calcular el valor de ΔC_p mediante la utilización de otras aproximaciones. Con esta finalidad, se utilizó una aproximación, desarrollada por Pace y Laurents (Pace y Laurents, 1989), que consiste en medir el valor de ΔG a un pH determinado y a diferentes temperaturas por medio de desnaturalizaciones químicas, analizándolas siguiendo el modelo de extrapolación lineal (Clarke y Fersht, 1993; Pace y Scholtz, 1997), y combinar estos datos con los obtenidos en las desnaturalizaciones térmicas al mismo pH.

Se realizaron experimentos de desnaturalización química por GdmHCl seguidos mediante DC en el ultravioleta lejano, utilizando 20 μM de CA-C a pH 7, y a una temperatura determinada en el rango de 278 a 313 K; en estas condiciones se está observando el desplegamiento de la forma monomérica de la proteína (Mateu, 2002). A partir del análisis de las curvas de desnaturalización química siguiendo el modelo de extrapolación lineal, se hallaron los valores de energía libre a cada temperatura estudiada. Se obtuvo un valor de m constante, dentro del error, en todo el rango de temperatura analizado ($m = 1.3 \pm 0.1 \text{ kcal mol}^{-1} \text{ M}^{-1}$, Fig. 2.18 A); sin embargo, esto no es lo esperado por el modelo de extrapolación lineal (Zweifel y Barrick, 2002). Este comportamiento probablemente se debe a que en los experimentos de desnaturalización química seguidos por DC, las líneas de base de la proteína, tanto en la región nativa como en la desnaturalizada, presentan grandes pendientes, como se ha observado en otras proteínas (Pace y Scholtz, 1997; Neira y Gómez, 2004).

Al representar los valores de $[\text{GdmHCl}]_{1/2}$ y ΔG frente a la temperatura, se obtuvo una curva muy similar para ambos parámetros (Fig. 2.18 B). Se realizó el ajuste de la curva de estabilidad (ΔG frente a la temperatura) a pH 7, utilizando la siguiente ecuación:

$$\Delta G(T) = \Delta H(T_o) - T \cdot \Delta S(T_o) + \Delta C_p \cdot \left[T - T_o - T \cdot \ln \left(\frac{T}{T_o} \right) \right] \quad (2.10),$$

que es similar a la ecuación (2.7), pero en este caso T_o , la temperatura de referencia, se tomó a 298 K. El uso de esta ecuación evita errores en el ajuste debidos al gran número de datos experimentales en torno a la T_m . A partir del análisis de la curva de estabilidad, se obtuvieron los parámetros termodinámicos de la desnaturalización de la forma monomérica de CA-C: un valor de ΔH_m de 54 ± 2 kcal mol⁻¹, T_m de 332.7 ± 0.1 K (que coincide con los valores determinados anteriormente por DC en el ultravioleta lejano, absorbancia, FTIR y RMN), y ΔC_p de 1.14 ± 0.06 kcal mol⁻¹ K⁻¹ (similar, dentro del error, al obtenido mediante DC en el ultravioleta lejano).

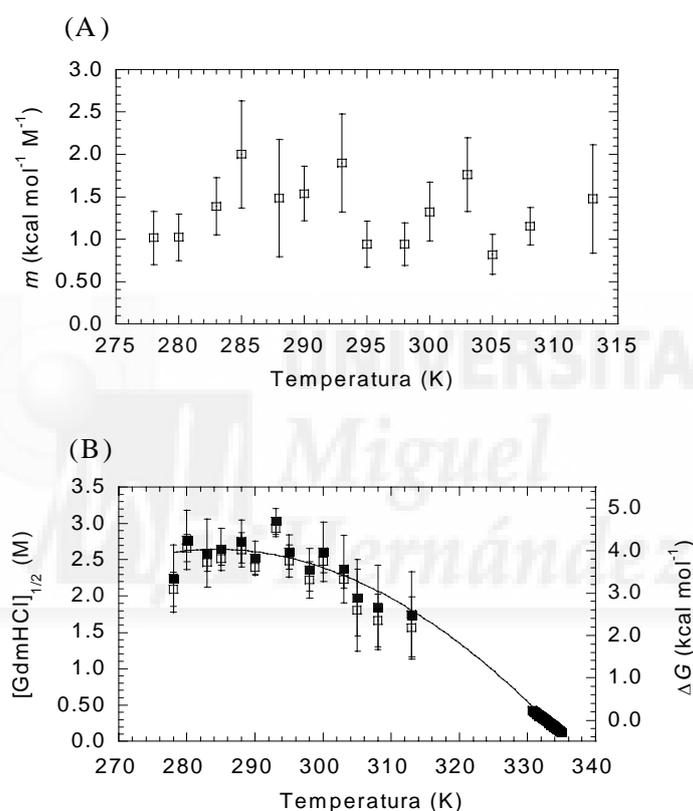


Figura 2.18 Determinación de los parámetros termodinámicos de la forma monomérica de CA-C, a partir de experimentos de desnaturalización química por GdmHCl a pH 7 seguidos por DC en el ultravioleta lejano. (A) Variación del valor de m con la temperatura. Las barras de error son los errores del ajuste a la ecuación del modelo de extrapolación lineal (apartado 2.1.7.). (B) Variación del valor de $[GdmHCl]_{1/2}$ (cuadrados blancos, eje izquierdo) y de ΔG (cuadrados negros, eje derecho) con la temperatura. Los valores de ΔG se obtuvieron utilizando el valor medio de m a todas las temperaturas estudiadas. Las barras de error son los errores del ajuste a la ecuación del modelo de extrapolación lineal (los errores fueron mayores a las temperaturas más altas, ya que las líneas de base de la proteína en la región nativa eran más estrechas). La línea es el ajuste por mínimos cuadrados a la ecuación (2.10).

Como puede observarse, los resultados de los experimentos de desnaturalización química son similares a los obtenidos en los experimentos de desnaturalización térmica. Por lo tanto, este hecho valida la utilización del modelo de extrapolación lineal; además, sugiere que mediante las medidas de desnaturalización, tanto química como térmica, se está observando el mismo estado desnaturalizado de CA-C.

2.3. DISCUSIÓN.

2.3.1. CAMBIOS CONFORMACIONALES DE CA-C CON EL pH.

2.3.1.1. Cambios conformacionales de CA-C a pH ácido.

CA-C, tanto en forma monomérica como dimérica, experimenta una transición conformacional a pH ácido, con un pK_a de 4.3 ± 0.2 (es el valor medio de todos los obtenidos por las diferentes técnicas utilizadas: Figs. 2.1, 2.2, 2.3, 2.4 y 2.5). Es importante señalar que este pK_a es un valor aparente, ya que es el valor medio de los pK_a s de las curvas de titulación de los residuos de CA-C en estado tanto plegado como desnaturalizado (Anderson et al., 1990). El bajo valor determinado para pK_a sugiere que la titulación se debe a un grupo ácido, como el ácido aspártico, el ácido glutámico o el grupo carboxilo C-terminal (Cantor y Schimmel, 1980). Por otro lado, tal y como se observó midiendo el cambio en la banda de absorción de las tirosinas por FTIR (Fig. 2.5), esta titulación se debe a la desprotonación de algún residuo ácido situado cerca de al menos uno de los dos residuos de tirosina de CA-C (Tyr164 o Tyr169). Proponemos que este residuo podría ser el Asp166, ya que se encuentra muy próximo a ambos residuos de tirosina de la proteína, tanto en la estructura primaria como en la secundaria, formando parte de la primera hélice α .

La tendencia mostrada en los experimentos de cromatografía de exclusión por tamaño molecular sugiere que la disociación del dímero ocurre a pH muy bajo, ya que a estos pHs se observa un aumento del V_e (Fig. 2.4). Entonces, se puede pensar que la transición ácida de CA-C está asociada con las etapas iniciales de la disociación del dímero. Sin embargo, a pesar de que no se puede afirmar que alguna fracción del dímero de CA-C se disocia con la disminución del pH, sí se puede decir que una fracción significativa de dímero todavía está presente en el rango de pH 3 a 4, como se deduce a partir de dos hechos característicos: (i) el diferente V_e a baja y alta concentración de CA-C (20 y 200 μ M) (Fig. 2.4); y, (ii) la presencia a pH 4 de una

transición térmica a baja temperatura no reversible en el espectro de FTIR (Fig. 2.14 A; Tabla 2.2).

En la transición ácida, CA-C pierde parte de su estructura terciaria (como se observó por fluorescencia, Fig. 2.1), pero parte se conserva (como se vio en los espectros de RMN, Fig. 2.16). Además, en los experimentos de cromatografía de exclusión por tamaño molecular hubo un cambio en la compacidad de la molécula (Fig. 2.4), y los experimentos de unión a ANS indican la exposición al disolvente de regiones hidrofóbicas de CA-C pertenecientes a cada monómero (Fig. 2.2). Después de la transición ácida, los monómeros de CA-C todavía conservan parte de su estructura secundaria, como se concluye a partir de los valores de elipticidad a 222 nm a pHs bajos (Fig. 2.3), y por la presencia de protones a campo alto en los experimentos de RMN a pH 3 (Fig. 2.16). La similitud en los desplazamientos químicos de algunas señales a campo alto entre los espectros a pH 7 y pH 3 (Fig. 2.16), sugiere que esta estructura secundaria residual presente en las especies a pH ácido es nativa. La desnaturalización térmica de la estructura secundaria residual a pH 4 (próximo al pK_a calculado) sigue un comportamiento sigmoidal, como se vio por DC en el ultravioleta lejano (Fig. 2.9; Tabla 2.2) y por FTIR (Fig. 2.14 A; Tabla 2.2); por el contrario, la desnaturalización térmica de la estructura terciaria residual a ese pH no sigue un comportamiento sigmoidal, como se observó en los experimentos de fluorescencia (Fig. 2.6), aunque probablemente se debe a la exposición del Trp184 al disolvente, cuyo entorno se puede haber alterado a consecuencia de la transición conformacional a pH ácido. Con la finalidad de comprobar esta hipótesis y medir la accesibilidad del triptófano, se llevaron a cabo experimentos de *quenching* por yoduro y acrilamida, y con ambos *quenchers* se observó que el triptófano a pHs bajos (pH 3) se encuentra tan expuesto como cuando la proteína está completamente desnaturalizada (Tabla 2.1). Por otro lado, la exposición del triptófano al disolvente también se confirmó mediante los experimentos de RMN a pH 3, en los que el desplazamiento químico del protón indol de CA-C, a la menor temperatura estudiada (278 K), fue 10.27 ppm (Fig. 2.17 F), muy similar al valor observado para los modelos de ovillo estadístico (*random coil*), que es 10.22 ppm (Wüthrich, 1986). Por lo tanto, podemos concluir que la espectroscopía de fluorescencia no es una sonda apropiada para estudiar la transición térmica de la estructura terciaria a pH bajo. El comportamiento descrito en este trabajo para CA-C a pH ácido es similar al mostrado por CA al mismo pH (Misselwitz et al., 1995). Debido a la ausencia de

estudios conformacionales de CA-N, sugerimos que el comportamiento de CA a pH ácido es debido a CA-C.

La estructura de CA-C a pH ácido, tal y como se ha comentado anteriormente, presenta unos rasgos que son característicos de los llamados glóbulos fundidos (*molten globules*) (Ptitsyn, 1995), aunque con algunas diferencias. En primer lugar, con la unión a ANS, ni los cambios en la intensidad de fluorescencia, ni en la longitud de onda de máxima emisión, fueron tan intensos como se esperaría (Stryer, 1965); este comportamiento probablemente se debe a que la superficie hidrofóbica de la proteína, que queda expuesta al entorno con la desnaturalización al disminuir el pH, o bien es pequeña, o bien no tiene mucha tendencia a unir ANS. Y en segundo lugar, la desnaturalización térmica de la estructura secundaria de CA-C a pH ácido sigue un comportamiento sigmoidal, distinto a lo observado en los primeros ejemplos descritos de *molten globules* (Ptitsyn, 1995). Sin embargo, como este comportamiento sigmoidal también se ha apreciado con posterioridad en otros *molten globules*, se está aceptando la idea de que estas especies podrían ser muy similares a las proteínas nativas, tanto en su estructura como en su energética (Maldonado et al., 1998; Irún et al., 2001).

2.3.1.2. Cambios conformacionales de CA-C a pH básico.

Como se vio por cualquiera de las técnicas biofísicas utilizadas, las propiedades espectroscópicas de CA-C en el rango de pH 6 a 8 permanecieron constantes. Sin embargo, las medidas de fluorescencia y DC indicaron la presencia de una nueva transición conformacional por encima de pH 8 (Figs. 2.1 y 2.3). Esta transición a pHs básicos se caracterizó por la presencia de cambios en la estructura secundaria de la proteína (Fig. 2.3), y en el entorno de los residuos de triptófano y tirosina (Fig. 2.1; Tabla 2.1); sin embargo, en los experimentos de unión a ANS no se observó la exposición al disolvente de regiones hidrofóbicas de CA-C (Fig. 2.2). Esta transición probablemente se deba a la titulación de residuos de lisina o arginina, o incluso a la titulación de la His226, pero no puede estar asociada a la titulación de residuos de tirosina, ya que la banda de absorción de la tirosina en el FTIR es muy sensible a la pérdida del protón fenólico (Barth y Zscherp, 2002), y no se observaron cambios en esta banda a pHs básicos (Fig. 2.5).

2.3.2. DESNATURALIZACIÓN TÉRMICA DE CA-C.

2.3.2.1. Disociación de la forma dimérica de CA-C.

La transición a baja temperatura observada por FTIR en los experimentos de desnaturalización térmica fue claramente dependiente de la concentración (Fig. 2.14 B),

y por lo tanto, se debe corresponder con la disociación de la forma dimérica de CA-C. Además, a partir de los experimentos de RMN a pH 7, hay dos hechos que sugieren que se produce dicho proceso de disociación: (i) la señal del protón indol desapareció en el intervalo de temperatura de 290 a 315 K (Fig. 2.17 C), de modo que el Trp184, que es un residuo clave en la interfase de dimerización (Gamble et al, 1997; Mateu, 2002), debe estar implicado en algún tipo de equilibrio conformacional lento en la escala de tiempos de la RMN; y, (ii) algunas señales de protones metilo desaparecieron a partir de 290 K (Fig. 2.17 A), aunque no el resto de las señales del espectro, y en consecuencia, sólo algunas regiones particulares de la proteína estaban implicadas en dicho equilibrio conformacional. Un comportamiento similar se ha observado, para un triptófano, en la reacción de desnaturalización del dominio monomérico B1 de la proteína G (Ding et al., 2004). En conclusión, los datos obtenidos sugieren que, tanto por FTIR como por RMN, se observó la disociación de CA-C.

2.3.2.2. Estados conformacionales de la forma monomérica de CA-C.

Como se puede concluir de la variación de T_m con el pH (Tabla 2.2), en líneas generales la estabilidad térmica de la forma monomérica de CA-C aumenta desde pH 4 hasta una zona en torno a pH 6, y a partir de ahí disminuye a medida que el pH aumenta. Las técnicas utilizadas se pueden dividir en dos grupos, según el valor de T_m obtenido: DC en el ultravioleta lejano, absorbancia y RMN dieron un valor de T_m mayor (aproximadamente en 7 K) que el hallado por unión a ANS, anisotropía y DC en el ultravioleta cercano. El hecho de que no coincidan las curvas de desnaturalización de la proteína obtenidas mediante el seguimiento de diferentes propiedades espectroscópicas, es una evidencia de la acumulación de un intermediario de plegamiento en el proceso de desnaturalización (Luo et al., 1995, 1997). Por otra parte, el hecho de que se obtenga distinto valor para T_m mediante DC en el ultravioleta cercano y absorbancia, que son dos técnicas que proporcionan información sobre el entorno de los residuos aromáticos, no es nuevo, sino que también se ha observado en otras proteínas (Maldonado et al., 1998; Irún et al., 2001; Campos et al., 2004). Esto se puede explicar considerando que el espectro de absorbancia proporciona información, principalmente, sobre los cambios en la exposición al disolvente de los residuos aromáticos (Match et al., 1995), que a su vez están muy influenciados por los cambios en la estructura secundaria de las proteínas; sin embargo, la técnica de DC en el ultravioleta cercano informa sobre los cambios que ocurren en el entorno asimétrico de los residuos

aromáticos, que prácticamente no se altera con las modificaciones en la estructura secundaria.

Entonces, en base a los resultados obtenidos se puede decir que, para la forma monomérica de CA-C, hay al menos tres estados conformacionales (ver Fig. 2.19): (i) la forma monomérica observada por encima de 307 K, que resulta de la disociación del dímero; (ii) la especie obtenida a partir de una reorganización conformacional, cuyo valor de T_m fue hallado por unión a ANS, anisotropía y DC en el ultravioleta cercano; y, (iii) el estado desnaturalizado final, resultado de la transición cuya T_m fue obtenida mediante DC en el ultravioleta lejano, absorbancia y RMN. En todos los casos, como las transiciones están bien definidas y separadas unas de otras en unos 7 K, cada transición se ajustó a una curva de dos estados (ecuaciones (2.6) y (2.7); Tabla 2.2). El hecho de que por diferentes técnicas biofísicas se obtengan los mismos valores para T_m , permite descartar que las distintas T_m s obtenidas sean resultado del proceso de ajuste; además, para confirmarlo, se realizó un ajuste global de los datos de DC en el ultravioleta lejano y cercano a una concentración de proteína de 20 y 200 μM . Este análisis mostró que hay tres estados, con $T_{m1} = 329 \pm 3 \text{ K}$ y $T_{m2} = 335 \pm 3 \text{ K}$, ambos valores similares a los obtenidos a partir del ajuste individual de cada transición a una curva de dos estados. El hecho de que la especie monomérica parcialmente plegada (que surge de una reorganización conformacional) una ANS, mientras que el monómero que resulta de la disociación no lo haga, sugiere que la forma monomérica reordenada puede ser similar a la observada a pH ácido a 298 K. Además, es importante señalar que la especie monomérica cuyo ensamblaje da lugar a la formación del dímero, no presenta las características de los llamados glóbulos fundidos, al contrario de lo que se ha observado en otros oligómeros (Ptitsyn, 1995; Jaenicke y Lillie, 2000).

2.3.2.3. Equilibrio de desnaturalización térmica de CA-C.

Los datos de desnaturalización térmica sugieren que el desplegamiento de la forma dimérica de CA-C transcurre mediante un mecanismo de cuatro estados. La reacción de disociación se produce a temperaturas en torno a 307 K, dando lugar a un intermediario monomérico plegado, y éste experimenta una transición conformacional a temperaturas alrededor de 325 K, para finalmente desnaturalizarse a temperaturas en torno a 332 K (Fig. 2.19), de forma que la disociación del dímero y la desnaturalización del monómero no están acoplados. Es importante tener en cuenta que los resultados obtenidos sólo se han observado en equilibrio, y que la secuencia de especies que se sugiere puede que no se observe cuando se describa la ruta de plegamiento cinético de

CA-C. Estas series de intermediarios térmicos coinciden con lo encontrado en experimentos de desnaturalización química seguidos por fluorescencia y DC en el ultravioleta lejano, que siguen exclusivamente la disociación del dímero-reorganización conformacional del monómero y la desnaturalización del monómero, respectivamente (Mateu, 2002; Fig 2.19).

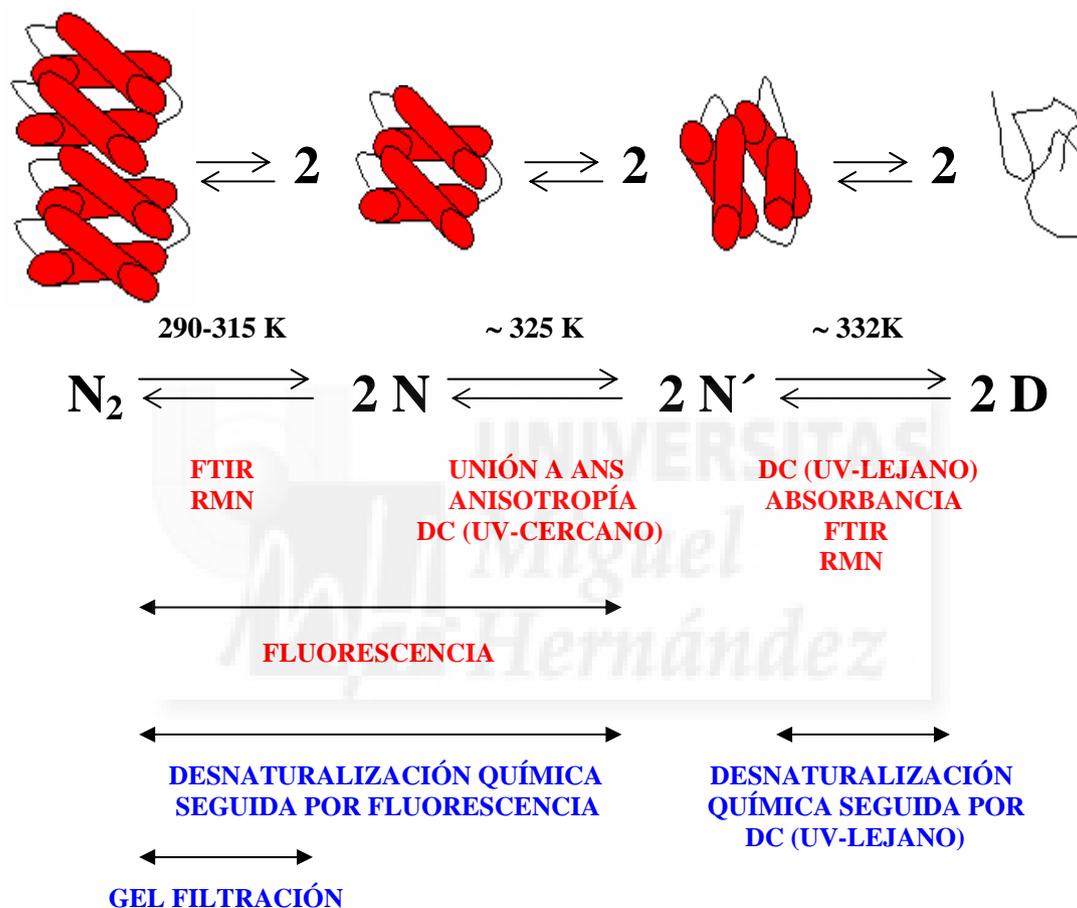


Figura 2.19 Equilibrio de desnaturalización térmica de CA-C. Esquema de las diferentes especies observadas en el equilibrio de desplegamiento de CA-C a pH 7, y las distintas técnicas utilizadas para detectarlas. N₂, dímero nativo; N, monómero plegado; N', monómero parcialmente plegado; D, monómero desnaturalizado. Sobre las flechas se indican las T_m obtenidas para cada transición; y, bajo las flechas se señalan las técnicas utilizadas en este trabajo (en rojo), o en otros trabajos (Mateu, 2002; del Álamo et al., 2003) (en azul), por las que se detectan las distintas transiciones. Los dibujos de la parte superior sólo tienen un propósito ilustrativo.

Como no se puede concluir sin ambigüedad que la transición térmica observada en los experimentos de fluorescencia sea independiente de la concentración, tampoco se puede asignar esta transición de forma particular al proceso de disociación y/o reordenamiento conformacional. Probablemente, la desnaturalización térmica seguida por fluorescencia informa sobre varios procesos, tal y como se ha discutido en otros estudios sobre la dependencia de la desnaturalización química con la concentración seguida por fluorescencia (Mateu, 2002). En base a los resultados obtenidos a partir de los experimentos de unión a ANS, anisotropía, DC en el ultravioleta lejano y FTIR, se puede sugerir que la desnaturalización térmica por fluorescencia está midiendo la disociación del dímero y el reordenamiento que tiene lugar posteriormente.

2.3.3. PARÁMETROS TERMODINÁMICOS DE LA DISOCIACIÓN Y DESNATURALIZACIÓN DE LAS FORMAS DIMÉRICA Y MONOMÉRICA DE CA-C: COMPARACIÓN CON MODELOS TEÓRICOS.

El valor de ΔC_p observado en la desnaturalización de la proteína es el resultado de los cambios en la hidratación de los distintos grupos, que están enterrados en el estado nativo completamente plegado y quedan expuestos al disolvente con la desnaturalización (Freire, 1995; Robertson y Murphy, 1997). Teniendo en cuenta que los cambios en el área accesible al disolvente, ASA, de un monómero de CA-C con la desnaturalización, calculados a partir de su estructura de rayos X, son: $\Delta ASA_{\text{nopolar}} = 4298 \text{ \AA}^2$ y $\Delta ASA_{\text{polar}} = 1385 \text{ \AA}^2$ (Gamble et al., 1997; Worthylake et al., 1999), y utilizando la siguiente expresión:

$$\Delta C_p = (0.45 \pm 0.02) (\Delta ASA_{\text{nopolar}}) + (-0.26 \pm 0.03) (\Delta ASA_{\text{polar}}) \quad (2.11),$$

se obtiene un valor de ΔC_p de $1.6 \pm 0.2 \text{ kcal mol}^{-1} \text{ K}^{-1}$, similar al determinado por DC en el ultravioleta lejano ($1.8 \pm 0.5 \text{ kcal mol}^{-1} \text{ K}^{-1}$), y por la aproximación de Pace y Laurents ($1.14 \pm 0.06 \text{ kcal mol}^{-1} \text{ K}^{-1}$; Pace y Laurents, 1989); además, el valor medido también es similar al determinado para otras proteínas de tamaño parecido (Robertson y Murphy, 1997).

Por otra parte, el valor de m obtenido ($1.3 \pm 0.1 \text{ kcal mol}^{-1} \text{ M}^{-1}$, que es el valor medio de todas las temperaturas estudiadas) (Fig. 2.18 A), también se puede comparar con aquellos calculados a partir de aproximaciones teóricas (Myers et al., 1995) mediante: $m (\text{cal mol}^{-1} \text{ M}^{-1}) = (368 \pm 0.11 \Delta ASA_{\text{total}})$. Como para la forma monomérica de CA-C, $\Delta ASA_{\text{total}} = 5683 \text{ \AA}^2$, entonces $m = 0.993 \text{ kcal mol}^{-1} \text{ M}^{-1}$, que es ligeramente inferior al obtenido a partir de los datos experimentales. Este hecho también se ha

observado en otras proteínas helicoidales, para las que se han predicho valores teóricos inferiores a los obtenidos experimentalmente (Padmanabhan et al., 1997).

El uso de la expresión anterior (2.11) también permite estimar el valor de ΔC_p para la reacción de disociación de CA-C, asumiendo que después de la disociación, los monómeros conservan la mayor parte de la estructura nativa que tenían cuando formaban parte del dímero. Esta es una aproximación realista que se basa en varias líneas de evidencia: (i) la similitud entre los espectros de DC en el ultravioleta lejano de las formas monomérica y dimérica (Mateu, 2002); (ii) la presencia de protones con desplazamientos químicos similares en el espectro de RMN de las formas monomérica y dimérica (resultados no publicados); y, (iii) los pequeños cambios en la anchura de la banda amida I' del espectro de FTIR en la transición a baja temperatura (Fig. 2.13). Los cambios estimados en el área accesible a la superficie para cada monómero después de la formación del dímero, a partir de la estructura de rayos X, son: $\Delta ASA_{\text{nopolar}} = 611 \text{ \AA}^2$ y $\Delta ASA_{\text{polar}} = 245 \text{ \AA}^2$ (Worthylake et al., 1999), y a partir de ellos se obtiene: $\Delta C_p = 211 \pm 10 \text{ cal mol}^{-1} \text{ K}^{-1}$ por monómero, que es un valor demasiado pequeño para ser detectado por la mayoría de las técnicas utilizadas en este trabajo. Los parámetros termodinámicos de la reacción de disociación de CA-C, obtenidos por FTIR, son: $\Delta H^0 = 74 \pm 8 \text{ kcal (mol de unidad cooperativa)}^{-1}$ y $\Delta S^0 = 230 \pm 30 \text{ cal K}^{-1} \text{ (mol de unidad cooperativa)}^{-1}$ a una concentración de proteína de 1 M (condiciones estándar), que también son valores pequeños en comparación con los correspondientes para otras proteínas oligoméricas pequeñas, en las que la disociación y desnaturalización de los dímeros sucede a la vez (Neet y Timm, 1994; Backmann et al., 1998; Jaenicke y Lillie, 2000). Estos pequeños valores se deben a que, en CA-C, las magnitudes de ΔH^0 , ΔS^0 y ΔC_p , reflejan exclusivamente la disociación del dímero.

En experimentos previos, la energía libre de la reacción “disociación del dímero-reordenamiento conformacional del monómero de CA-C” se obtuvo a partir de medidas de desnaturalización química por GdmHCl seguidas por fluorescencia a 298 K, dando un valor de $12.1 \text{ kcal mol}^{-1}$ a una concentración de proteína de 1 M (condiciones estándar) (Mateu, 2002). También se obtuvo la energía libre del paso específico de disociación a 298 K, a partir de la dilución de la proteína en análisis de gel filtración, dando un valor de $6.9 \text{ kcal mol}^{-1}$ a una concentración de proteína de 1 M (condiciones estándar) (del Álamo et al., 2003). Si se asume, en base a los cálculos anteriores, que el valor de ΔC_p para la reacción de disociación es pequeño, utilizando los valores medidos

de ΔH^0 y ΔS^0 , la energía libre de la disociación fue: $\Delta G = 5 \pm 8 \text{ kcal mol}^{-1}$ a 298 K. Aunque el gran error en este último valor impide una comparación cuantitativa, estos cálculos indican que, tanto la disociación como una reorganización conformacional del monómero, pueden contribuir significativamente a la transición combinada de disociación-reordenamiento medida por fluorescencia en los experimentos de desnaturalización química (Mateu, 2002; del Álamo et al., 2003).

2.3.4. POSIBLES IMPLICACIONES BIOLÓGICAS.

Aún no está establecido con exactitud el papel que desempeña el pH en la maduración del VIH *in vivo*, pero se sabe que muchas cápsidas virales son sensibles al pH; por ejemplo, las cápsidas del virus de la fiebre aftosa y de los mengovirus se disocian a pH bajo, y las cápsidas de los rinovirus y los poliovirus sufren cambios conformacionales a pH ácido. La forma de los ensamblados macromoleculares formados por CA (Gross et al., 2000; Ehrlich et al., 2001), y su velocidad de asociación, también son dependientes del pH (Lanman et al., 2002); en estos estudios se sugiere que, entre pH 6 y 7, se produce una transición hacia una conformación de CA más extendida a pHs básicos (Gross et al., 2000). Asimismo, en otros sistemas virales se han observado cambios morfológicos en CA en torno a pH neutro (Johnson, 1996). Los experimentos de cromatografía de exclusión por tamaño molecular realizados en este trabajo indican que a pHs neutros y básicos no se producen cambios significativos en la compacidad de CA-C (Fig. 2.4). Dado que el ensamblaje de la cápsida del VIH depende tanto de CA-C como de CA-N, sugerimos que, la mayoría de los cambios observados en CA a pHs neutros y básicos, son debidos a CA-N.

Finalmente, a partir de los resultados obtenidos en los experimentos de desnaturalización térmica, se puede pensar que, a temperaturas fisiológicas CA-C se encuentra en forma monomérica, ya que a pH neutro y 1 mM de CA-C, la temperatura de disociación fue de 307 K (Tabla 2.2). Sin embargo, se debe tener en cuenta que *in vivo*, CA-C es un dominio de CA, cuyas propiedades de asociación están muy influenciadas por el otro dominio, CA-N (Lanman et al., 2002). Quizás, la baja estabilidad térmica del dímero de CA-C aislado, junto con la baja afinidad de dimerización del dominio, son esenciales para permitir el desensamblaje de la cápsida (del Álamo y Mateu, 2005), y/o para mediar un amplio rango de interacciones macromoleculares durante el ensamblaje de la misma, como se ha sugerido para otras macromoléculas (Tang et al., 1999).

***DISEÑO DE UN PÉPTIDO
Y ESTUDIO DE SU
INTERACCIÓN
CON CA-C***



Para determinar si la α -hélice 2 (que constituye fundamentalmente la interfase de interacción de CA-C), por sí sola, es suficiente para inhibir la dimerización del dominio completo, se ha llevado a cabo el diseño de un péptido (CAC1) que comprende la secuencia de dicha hélice α , y se han estudiado, tanto su estructura como sus propiedades de asociación a CA-C, mediante diversas técnicas biofísicas: fluorescencia, DC, RMN, cromatografía de exclusión por tamaño molecular, cromatografía de afinidad analítica, e ITC.

Los experimentos de DC, cromatografía de exclusión por tamaño molecular, y RMN indican que CAC1 es capaz de interactuar con CA-C. Además, mediante fluorescencia, cromatografía de afinidad analítica, e ITC, se ha determinado la constante de disociación aparente del heterocomplejo formado entre CAC1 y CA-C, que es del orden de 50 μ M, sólo cinco veces superior que la del dominio completo (10 μ M) (Gamble et al., 1997).

3.1. MATERIALES Y MÉTODOS.

3.1.1. PRODUCCIÓN DE PROTEÍNA.

La expresión, purificación, determinación de la pureza y concentración de CA-C silvestre se realizó como se describe en el apartado 2.1.1.

La proteína EI silvestre de *S. coelicolor*, utilizada como un control de proteína dimerica en experimentos de unión, se expresó en la cepa *E. coli* BL21 (DE3), y se purificó utilizando cromatografía de afinidad Ni^{2+} -NTA (Invitrogen). Con el fin de eliminar cualquier proteína o ADN que hubiese coeluido con la proteína, se realizó una cromatografía de exclusión por tamaño molecular utilizando una columna Superdex 200 16/60 (Amersham Biosciences) acoplada a un aparato de cromatografía AKTA-FPLC con un detector de absorbancia a 280 nm (Amersham Biosciences).

3.1.2. DISEÑO DEL PÉPTIDO, SÍNTESIS Y PURIFICACIÓN.

Se diseñó un péptido de 20 aminoácidos al que se denominó CAC1, con la secuencia: Ac-EQASQEVKNWMTETLLVQNA-CONH₂, derivada de la secuencia de CA-C, que comprende los residuos 175-194. Estos aminoácidos forman la α -hélice 2 de CA-C (Ser178-Val191, es decir, la mayor parte de la interfase de dimerización de CA-C) más tres residuos a cada lado para evitar los efectos de deshilachamiento en los extremos de la posible hélice del péptido. Los extremos N y C terminal fueron acetilados y amidados, respectivamente, para evitar los efectos de carga.

Se obtuvo una versión fluorescente del péptido CAC1, al que se denominó CAC1Aib, marcando CAC1 con una sonda fluorescente derivada del ácido antranílico (Fig. 3.1), mediante incorporación directa durante la síntesis automática de un grupo antranilo en el grupo ϵ -amino de la cadena lateral del único residuo de lisina del péptido (Lys182). El péptido CAC1Aib mostró sendos máximos de absorción a 246 y 320 nm, como se espera del grupo antranilo (Stennicke et al., 1997).

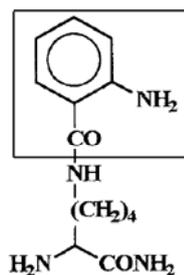


Figura 3.1 Estructura química de la sonda fluorescente incorporada al péptido CAC1. La sonda está formada por una molécula de ácido antranílico (delimitada por un cuadrado) que se acopla al grupo ϵ -amino de la cadena lateral de un residuo de lisina.

La pureza de ambos péptidos se comprobó por espectrometría de masas y HPLC, siendo ésta mayor del 95 %, y la concentración se determinó por peso de la correspondiente cantidad de péptido seco. Ambos péptidos fueron proporcionados por el Dr. Graham Bloomberg (Universidad de Bristol).

3.1.3. FLUORESCENCIA.

Los experimentos de fluorescencia se realizaron en un espectrofluorímetro SLM 8000 (Spectronics Instruments) conectado a un baño de agua Haake, o en un espectrofluorímetro Cary Eclipse (Varian) conectado a una celda Peltier. Se utilizó una cubeta de cuarzo de 0.5 cm paso óptico (Hellma) para las medidas en el espectrofluorímetro SLM, y una cubeta de cuarzo de 1 cm de paso óptico (Hellma) para el espectrofluorímetro Varian.

3.1.3.1. Determinación de la constante de autodisociación aparente de CAC1 y CAC1Aib mediante fluorescencia en el estado estacionario.

Se añadieron cantidades cada vez mayores de CAC1, en el rango de 0.5 a 190 μ M, a una solución de CAC1Aib de concentración fija (1.6 μ M) en tampón 25 mM fosfato pH 7.0, y se midió la fluorescencia después de dos horas de incubación a 298 K.

Los experimentos se realizaron en el espectrofluorímetro Cary Eclipse (Varian) excitando a 246 nm (o 320 nm), y recogiendo la emisión de fluorescencia entre 400 y 500 nm. Las rendijas de excitación y emisión fueron ambas de 5 nm, y la señal se adquirió cada 1 nm. Todos los espectros se corrigieron restando la línea de base adecuada. No se pudieron registrar concentraciones inferiores de CAC1Aib (0.9 μM) debido a la baja relación señal a ruido.

Se observó un cambio en la intensidad de fluorescencia del grupo antrnilo de CAC1Aib a 420 nm al unirse a CAC1. Así pues, la constante de autodisociación aparente se calculó ajustando ese cambio en la intensidad de fluorescencia con la concentración de CAC1 añadida a:

$$F_{meas} = F^0 + \Delta F_{max} \left(\frac{[CAC1Aib]}{([CAC1Aib] + [CAC1] + K_D) - \sqrt{([CAC1Aib] + [CAC1] + K_D)^2 - 4[CAC1][CAC1Aib]}} \right) \quad (3.1),$$

o a su forma simplificada:

$$F_{meas} = F^0 + \Delta F_{max} \frac{[CAC1]}{[CAC1] + K_D} \quad (3.2),$$

donde $[CAC1Aib]$ es la concentración del péptido marcado con antrnilo; $[CAC1]$ es la concentración de péptido sin marcar; F_{meas} es la intensidad de fluorescencia medida a cada concentración de CAC1; F^0 es la intensidad de fluorescencia del péptido marcado; ΔF_{max} es el cambio en la intensidad de fluorescencia medida cuando todo CAC1 se encuentra unido; y, K_D es la constante de autodisociación aparente de CAC1. En ambas ecuaciones, los parámetros derivados del ajuste: F^0 , ΔF_{max} y K_D , dependen del proceso iterativo y de sus valores iniciales, por lo que se hizo un análisis de la varianza; con las dos ecuaciones se obtuvieron los mismos parámetros. La constante de disociación aparente obtenida es la media de tres mediciones independientes. Los ajustes por análisis de mínimos cuadrados a las diferentes ecuaciones de unión, se llevaron a cabo empleando la opción de ajuste general de curvas del programa KaleidaGraph (*software* de Abelbeck).

El péptido marcado con antrnilo se disolvió en metanol (debido a su baja solubilidad en agua) alcanzando una concentración final de 410 μM . De esta disolución madre se añadieron alícuotas a las correspondientes soluciones (que contenían el péptido sin marcar o CA-C, ver el apartado 3.1.3.3.), alcanzando una concentración final de CAC1Aib de 1.6 μM . Así, la concentración residual final de metanol en la muestra fue inferior al 0.4 %. Estas pequeñas concentraciones de disolvente orgánico no alteraron la estructura ni las propiedades de autoasociación ni del péptido sin marcar ni

de CA-C, como se comprobó por fluorescencia y DC. La absorbancia de muestras de CAC1Aib y CAC1 (a concentraciones en el rango de 2 a 120 μM) a longitudes de onda mayores de 460 nm confirmó la ausencia de agregados de peso molecular elevado.

3.1.3.2. Determinación de la constante de autodisociación aparente de CAC1 y CAC1Aib mediante anisotropía en el estado estacionario.

Los experimentos de anisotropía fueron realizados por Francisco N. Barrera (IBMC-UMH, Elche). La anisotropía en el estado estacionario está definida por la ecuación (2.2) (apartado 2.1.2.2.3.). El valor del factor G fue de 0.84 para el grupo indol, y de 0.58 para el grupo antranilo. La anisotropía de fluorescencia se registró en una disposición en forma de T (tomando simultáneamente la luz emitida en paralelo y perpendicular) en el espectrofluorímetro SLM 8000. Para la determinación de la anisotropía del grupo indol se excitó a 280 nm y se recogió la emisión a 350 nm. Para la determinación de la anisotropía del grupo antranilo se excitó a 320 nm y se recogió la emisión a 420 nm. Se añadieron cantidades cada vez mayores de CAC1, en el rango de 2 a 250 μM , a una solución de CAC1Aib de concentración fija (1.6 μM) en tampón 25 mM fosfato pH 7.0, y se midió la anisotropía en el estado estacionario después de dos horas de incubación a 298 K. Cada valor de anisotropía obtenido es la media de cinco mediciones independientes.

La anisotropía de ambos grupos fluorescentes a cada concentración de péptido, se ajustó a:

$$\langle r \rangle_{meas} = \langle r \rangle^0 + \Delta \langle r \rangle_{max} \frac{[CAC1]}{[CAC1] + K_D} \quad (3.2'),$$

que es análoga a la ecuación (3.2).

3.1.3.3. Determinación de la constante de disociación aparente del complejo CAC1-CA-C mediante fluorescencia en el estado estacionario.

El procedimiento utilizado fue similar al descrito en el apartado 3.1.3.1. Se añadieron cantidades cada vez mayores de CA-C, en el rango de 0.5 a 120 μM , a una solución de CAC1Aib de concentración fija (1.6 μM) en tampón 25 mM fosfato pH 7.0, y se midió la fluorescencia después de dos horas de incubación a 298 K. Los experimentos se realizaron en el espectrofluorímetro Cary Eclipse (Varian) excitando a 246 nm (o 320 nm), y recogiendo la emisión de fluorescencia entre 400 y 500 nm. Las rendijas de excitación y emisión fueron ambas de 5 nm y la señal se adquirió cada 1 nm. Todos los espectros se corrigieron restando la línea de base adecuada. No se pudieron

registrar concentraciones inferiores de CAC1Aib (0.9 μM) debido a la baja relación señal a ruido.

Se observó un cambio en la intensidad de fluorescencia del grupo antrnilo de CAC1Aib a 420 nm al unirse a CA-C. Así pues, la constante de disociación aparente se calculó ajustando ese cambio en la intensidad de fluorescencia con la concentración de CA-C añadida a:

$$F_{meas} = F^0 + \Delta F_{max} \left(\frac{[CAC1Aib] + [CA-C] + K_D - \sqrt{([CAC1Aib] + [CA-C] + K_D)^2 - 4[CA-C][CAC1Aib]}}{2[CA-C]} \right) \quad (3.1')$$

(3.1'), que es análoga a la ecuación (3.1), pero aquí, la K_D es la constante de disociación aparente del heterocomplejo. De forma alternativa, también se utilizó su forma simplificada:

$$F_{meas} = F^0 + \Delta F_{max} \frac{[CA-C]}{[CA-C] + K_D} \quad (3.2'')$$

que es análoga a la ecuación (3.2).

3.1.3.4. *Quenching* de fluorescencia del grupo antrnilo.

Los experimentos de *quenching* de fluorescencia del grupo antrnilo por yoduro se realizaron excitando a 320 nm. Los espectros de emisión se recogieron entre 390 y 450 nm, a 298 K. La anchura de la rendija se fijó en 5 nm, tanto para la excitación como para la emisión, la señal se adquirió cada 1 nm, y el tiempo de integración fue de 1 s. El rango de concentración de yoduro utilizado fue de 0 a 0.7 M, en intervalos de 0.1 M. La fuerza iónica se mantuvo constante mediante la adición de KCl; además, se añadió $\text{Na}_2\text{S}_2\text{O}_3$ a una concentración final de 0.1 M con el fin de evitar la formación de I_3^- .

Los parámetros que definen el *quenching* dinámico por yoduro se obtuvieron mediante el ajuste de los datos de intensidad de fluorescencia a diferentes longitudes de onda (en el intervalo de 335 a 338 nm) a la ecuación de Stern-Volmer (Lakowicz, 1999):

$$\frac{F_0}{F} = 1 + K_{sv}[X] \quad (3.3),$$

donde K_{sv} es la constante de Stern-Volmer para el *quenching* dinámico; F_0 es la intensidad de fluorescencia en ausencia de *quencher*; y, F es la intensidad de fluorescencia medida a una concentración de *quencher* dada, $[X]$ (M). Cada valor obtenido para K_{sv} es la media de cuatro mediciones independientes. La concentración de CAC1Aib en todos los casos fue 2 μM , en tampón 25 mM fosfato pH 7.0; y para CA-C y EI fue 4 μM , en tampón 25 mM fosfato pH 7.0.

3.1.4. DICROÍSMO CIRCULAR.

El espectropolarímetro usado para la realización de los experimentos de DC ha sido descrito en el apartado 2.1.3.1. Se utilizó una cubeta de cuarzo de 0.1 cm de paso óptico (Hellma).

3.1.4.1. Determinación de la constante de autodisociación aparente de CAC1.

Se tomaron espectros de DC en el ultravioleta lejano, entre 190 y 250 nm, con una velocidad de barrido de 50 nm/min, una anchura de banda de 1 nm, y un tiempo de respuesta de 2 s, siendo promediados sobre cuatro experimentos a 298 K. Las concentraciones de CAC1 utilizadas estuvieron en el rango de 10 a 300 μM , en tampón 25 mM fosfato pH 7.0. No se pudieron realizar experimentos a concentraciones inferiores debido a la baja relación señal a ruido. Todos los espectros se corrigieron mediante la sustracción de la línea de base correspondiente.

La elipticidad molar por residuo se calculó utilizando la ecuación (2.3), definida en el apartado 2.1.3.1.

La constante de autodisociación aparente se calculó ajustando el cambio de $[\Theta]_{222}$ con la concentración de CAC1 añadida a:

$$[\Theta]_{222} = [\Theta]^0 + \Delta[\Theta]_{max} \frac{[CAC1]}{[CAC1] + K_D} \quad (3.2'''),$$

que es análoga a la ecuación (3.2).

3.1.4.2. Medidas de desnaturalización térmica.

Las desnaturalizaciones térmicas en el ultravioleta lejano se adquirieron siguiendo la elipticidad a 222 nm, desde 298 hasta 363 K, utilizando una velocidad de calentamiento de 60 K/h, una anchura de banda de 1 nm, y un tiempo de respuesta de 8 s. La concentración de CA-C utilizada fue 20 μM , en tampón 25 mM fosfato pH 7.0, añadiendo cantidades cada vez mayores de CAC1, en el rango de 20 a 60 μM . También se realizaron experimentos de desnaturalización térmica con CAC1 solo, en el rango de 40 a 192 μM . La reversibilidad de las transiciones térmicas se comprobó realizando un nuevo experimento tras enfriar hasta 298 K la muestra previamente desnaturalizada, y comparando éste con el espectro obtenido en el primer barrido. Cada experimento de desnaturalización térmica se repitió dos veces con muestras nuevas. La posibilidad de desviación del espectropolarímetro se descartó tomando espectros del tampón antes y después de cada experimento de temperatura, no encontrándose diferencia entre ambos.

3.1.5. CROMATOGRAFÍA DE EXCLUSIÓN POR TAMAÑO MOLECULAR.

Se usó una columna analítica de exclusión por tamaño molecular (Superdex 75 H/R 10/30, Amersham Biosciences) acoplada a un aparato de cromatografía AKTA-FPLC (Amersham Biosciences) a 298 K, y detectando la absorbancia a 280 nm. La columna se calibró como se indica en el apartado 2.1.4. Se utilizó un flujo de 1 ml/min, y se cargaron alícuotas de 100 μ l en la columna después de ser equilibrada con 4 volúmenes de columna en tampón de elución (25 mM fosfato pH 7.0, con 150 mM de NaCl para evitar interacciones no específicas entre la columna y la proteína). La concentración de CA-C utilizada fue 20 μ M en todos los casos, y se fueron añadiendo cantidades cada vez mayores de CAC1, en el rango de 60 a 128 μ M. La mezcla formada por la proteína y el péptido se incubó durante dos horas a 298 K. El V_e obtenido es el resultado de cuatro medidas independientes con muestras nuevas.

3.1.6. RESONANCIA MAGNÉTICA NUCLEAR.

Los experimentos monodimensionales se adquirieron como se indica en el apartado 2.1.6. Se tomaron 25000 espectros, durante 15 horas, y la anchura espectral fue 6000 Hz en todos los casos. Los experimentos se realizaron con 10 μ M de péptido, proteína, o la mezcla de ambos (10 μ M de CAC1 y 10 μ M de CA-C) en tampón 25 mM fosfato pH 7.0 en $^2\text{H}_2\text{O}$. También se realizaron experimentos con 100 μ M de CA-C, y se fueron añadiendo cantidades cada vez mayores de CAC1, en el rango de 100 a 200 μ M, en tampón 25 mM fosfato pH 7.0 en $^2\text{H}_2\text{O}$, y se tomaron 2048 espectros. No se hicieron correcciones de los valores de pH por los efectos isotópicos. Los experimentos de RMN se realizaron en $^2\text{H}_2\text{O}$, a pesar de que la mayoría de las señales de los protones amida desaparecían, con el fin de incrementar la relación señal a ruido a las concentraciones más bajas utilizadas. Los espectros se procesaron como se indica en el apartado 2.1.6.

3.1.7. CROMATOGRAFÍA DE AFINIDAD ANALÍTICA.

La constante de disociación aparente del heterocomplejo, formado entre CAC1 y CA-C, también se determinó mediante cromatografía de afinidad analítica (Dunn y Chaiken, 1974). Estos experimentos fueron realizados por el Dr. Mauricio G. Mateu (CSIC-UAM, Madrid). CA-C se unió covalentemente a una matriz de gel Affi-Prep10 (BioRad), y se empaquetaron 5 ml de este gel (con CA-C acoplada) en una columna de cromatografía FPLC (Amersham Biosciences). Esta columna se equilibró en tampón de elución (25 mM fosfato pH 7.0, 150 mM NaCl) a 298 K. Las muestras de

proteína o péptido se aplicaron en pequeñas cantidades (500 μ l) y se eluyeron a 1 ml/min, detectando la absorbancia a 280 nm. La columna se calibró aplicando CA-C silvestre a una concentración total de proteína (monómero), C_t , de 0.5 a 20 μ M. La constante de autodisociación de CA-C, K_D , y la cantidad efectiva de CA-C unida covalentemente a la matriz, M_t , se determinaron ajustando los datos experimentales a la siguiente ecuación (Rosé et al., 1992):

$$\frac{I}{(V_e - V_o)} = \frac{K_D}{M_t} + \frac{C_t}{M_t} \quad (3.4),$$

donde V_o es el volumen de elución para una proteína de tamaño similar a CA-C que no interacciona con ésta (anhidrasa carbónica). El valor de K_D así obtenido para CA-C fue $9.8 \pm 0.4 \mu$ M, muy similar al determinado mediante ultracentrifugación analítica (Gamble et al., 1997); y el valor hallado para M_t fue 9.0 ± 0.2 nmol. El uso de este método o de un método diferente basado en elución frontal para calibrar las columnas de afinidad CA-C-Affi-Prep10 resultó en valores similares para M_t .

Para determinar la constante de disociación del complejo CAC1-CA-C, se aplicaron alícuotas de 500 μ l de CAC1, a concentraciones comprendidas entre 0.5 y 20 μ M, en la misma columna descrita anteriormente. Los valores de $\frac{I}{(V_e - V_o)}$ obtenidos a las diferentes concentraciones se ajustaron a la ecuación anterior (3.4), y la constante de disociación aparente se obtuvo multiplicando directamente el valor de la ordenada en el origen (es decir, cuando $C_t = 0$) por el valor de M_t determinado como se ha descrito (Rosé et al., 1992).

3.1.8. CALORIMETRÍA DE TITULACIÓN ISOTÉRMICA.

Los experimentos de ITC fueron realizados por el Dr. Javier Gómez Pérez (IBMC-UMH, Elche), con un instrumento Omega (MicroCal Inc.) acoplado a un nanovoltímetro externo, que se utilizó para mejorar la relación señal a ruido (Cooper y Johnson, 1994). Todas las medidas se tomaron a 298 K, en tampón 25 mM fosfato pH 7.0. Previamente, CA-C se dializó exhaustivamente a 277 K frente al mismo tampón. CA-C se cargó en la celda calorimétrica a una concentración de 21.5 μ M, y CAC1 se cargó en la jeringa a una concentración de 483 μ M. Se fueron añadiendo, secuencialmente, microlitros de la disolución de ligando (CAC1) a la celda calorimétrica (1.37 ml), que contenía la disolución de proteína (CA-C). Se monitorizó la

cantidad de calor necesaria para mantener la temperatura de la celda constante después de cada inyección en función del tiempo. En un experimento independiente, los calores individuales de la dilución para CAC1 (principalmente debidos a la disociación de la forma asociada del péptido después de la dilución) se determinaron bajo las mismas condiciones experimentales, haciendo inyecciones idénticas de la disolución de CAC1 en la celda de titulación que sólo contenía tampón. El calor intercambiado debido a la reacción de unión entre CAC1 y CA-C se obtuvo como la diferencia entre el calor de la reacción CAC1/CA-C y el correspondiente calor de la dilución CAC1/tampón. La isoterma resultante (cambios de la entalpía individual aparente después de la unión frente a la relación molar [CAC1]/[CA-C]) se ajustó a un modelo de un único sitio de unión, asumiendo que el complejo CAC1-CA-C tenía una estequiometría 1:1. Los parámetros termodinámicos de la unión se obtuvieron mediante el *software* de MicroCal.

3.2. RESULTADOS.

3.2.1. CARACTERIZACIÓN ESPECTROSCÓPICA DE CAC1 Y CAC1Aib.

3.2.1.1. Dicroísmo circular en el ultravioleta lejano.

El espectro de DC de CAC1 a la mínima concentración estudiada (10 μM) mostró sendos mínimos intensos a 195 y 220 nm, y un máximo a 205 nm (Fig. 3.2 A). Este espectro no es característico ni de una estructura desordenada ni de una hélice α (Kelly y Price, 2000), y puede deberse a la presencia del único triptófano, como se ha descrito en proteínas (Freskgard et al., 1994). La intensidad de estas tres bandas y la forma de los espectros de DC cambiaron a medida que aumentaba la concentración de péptido. Todos los espectros presentaron un punto isodicroico a 210 nm (Fig. 3.2 A), lo que sugiere la existencia de un proceso de dos estados que se puede describir por una sola constante de equilibrio (Woody, 1995; Kelly y Price, 2000). El ajuste del cambio de $[\Theta]_{222}$ frente a la concentración de péptido, utilizando la ecuación (3.2''), dio un valor de $17 \pm 10 \mu\text{M}$ para la constante de autodisociación de CAC1 (Fig. 3.2 B). Este valor de K_D presenta una gran incertidumbre porque no se pudo obtener un valor de la elipticidad experimental para las especies monoméricas, probablemente debido a que el péptido se encuentra parcialmente autoasociado a la concentración de péptido más baja que se pudo estudiar.

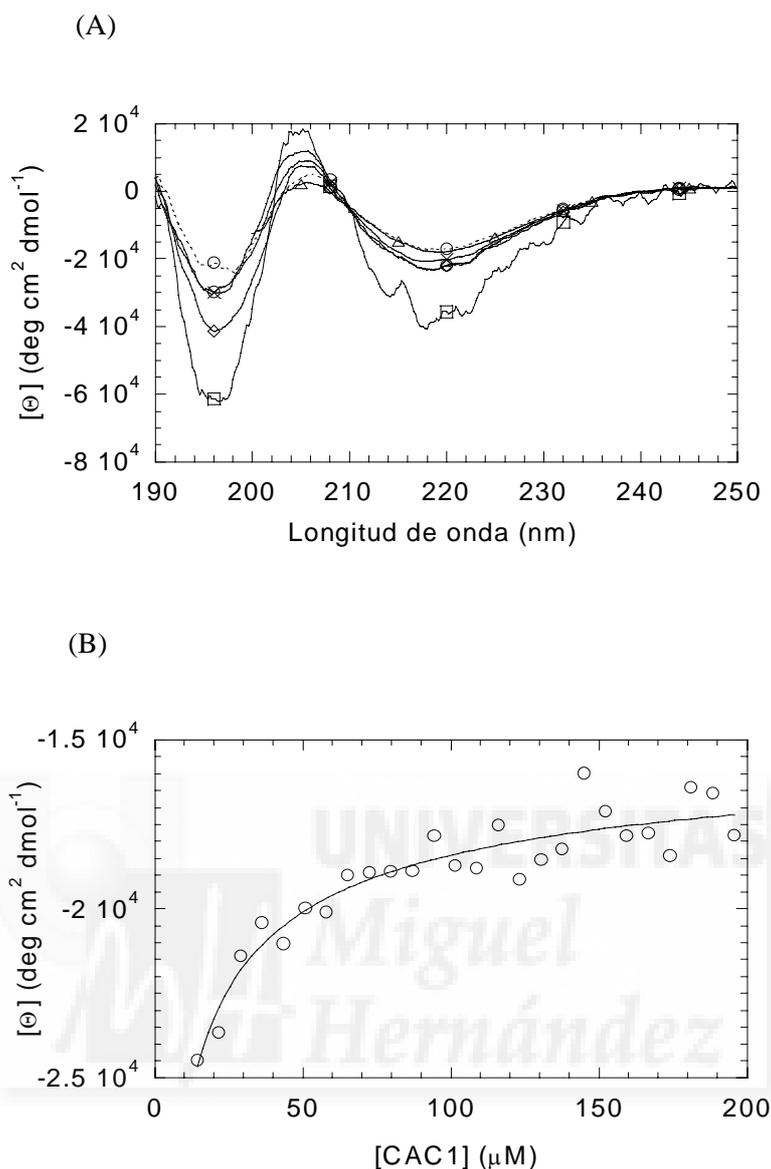


Figura 3.2 Determinación de la constante de autodisociación aparente de CAC1 mediante DC en el ultravioleta lejano. (A) Espectro de DC de CAC1 a distintas concentraciones: 10 μM (cuadrados), 20 μM (círculos), 30 μM (rombos), 50 μM (cruces), y 130 μM (triángulos), todos con línea continua; y, 100 μM (línea discontinua con círculos). (B) Elipticidad molar por residuo a 222 nm frente a la concentración de CAC1. La línea es el ajuste a la ecuación (3.2''). Los experimentos se realizaron a 298 K.

3.2.1.2. Fluorescencia en el estado estacionario.

El espectro de emisión de fluorescencia de CAC1 mostró un máximo en torno a 350 nm, correspondiente a la fluorescencia de un residuo de triptófano expuesto al disolvente (Pace y Scholtz, 1997). Al aumentar la concentración de péptido, se observó un desplazamiento del máximo desde 352 nm (a 2.3 μM) hasta 349 nm (a 120 μM).

La movilidad local y global del único grupo indol de CAC1 se puede determinar midiendo su anisotropía de fluorescencia (Lakowicz, 1999). La anisotropía aumentó de forma hiperbólica desde 0.05 (a 2 μM) hasta un valor de 0.12 (a 250 μM de CAC1) (Fig. 3.3). Este cambio en la anisotropía puede reflejar tanto un cambio del volumen de las especies moleculares (movimiento global), como un cambio en la movilidad del entorno cercano (movimiento local). Se obtuvo un valor para la constante de autodisociación de CAC1 de $16 \pm 5 \mu\text{M}$, que es similar, dentro del error experimental, al obtenido mediante DC, donde se está midiendo la formación total de estructura secundaria. Así pues, las medidas de anisotropía están indicando un cambio en el comportamiento global del péptido.

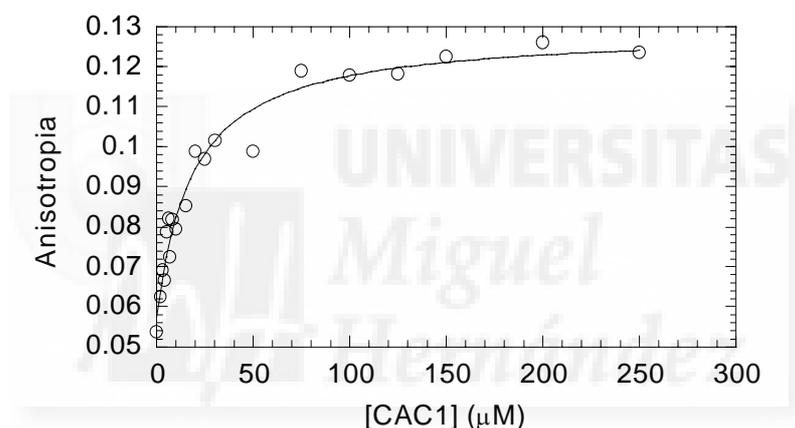


Figura 3.3 Determinación de la constante de autodisociación aparente de CAC1 mediante anisotropía. Anisotropía del Trp184 frente a la concentración de CAC1. La línea es el ajuste a la ecuación (3.2'). Los experimentos se realizaron a 298 K.

3.2.1.3. Resonancia magnética nuclear.

El espectro monodimensional de CAC1 a 10 μM en $^2\text{H}_2\text{O}$ (ver Fig. 3.6 A) muestra señales a campo alto en las mismas regiones que los modelos de ovillo estadístico (Wüthrich, 1986). Además, en la región de campo bajo sólo aparecieron las señales de protones del anillo aromático del triptófano, cuyos desplazamientos químicos también eran muy similares a los observados para los triptófanos en ovillo estadístico. Ninguno de los protones amida fue visible después de 5 minutos de estar expuesto en $^2\text{H}_2\text{O}$. Esto indica que el triptófano se encuentra expuesto al disolvente, confirmando el cambio observado en el máximo de emisión del espectro de fluorescencia de CAC1

(apartado 3.2.1.2.). Todo esto sugiere que, en la escala de tiempos de la RMN, el péptido se encuentra la mayor parte del tiempo desordenado. No se pudo determinar la estructura de CAC1 por RMN, ya que las muestras precipitaron a las altas concentraciones que se necesitan para la adquisición de espectros bidimensionales.

3.2.1.4. Diseño de un péptido CAC1 marcado con un grupo antranilo y estudio de sus propiedades de autoasociación.

Para el estudio de la unión de CAC1 a CA-C mediante fluorescencia, es necesaria la utilización de una sonda fluorescente alternativa que no interfiera con los fluoróforos intrínsecos del péptido (residuo de triptófano) o de la proteína (residuos de triptófano y tirosina). Con este fin, se creó una versión fluorescente de CAC1 (CAC1Aib), marcando este péptido con un grupo antranilo (Stennicke et al., 1997), cuya longitud de onda de máxima emisión no solapa con el máximo de emisión del triptófano ni de la tirosina.

Con la finalidad de comprobar si la introducción del grupo antranilo alteraba las tendencias conformacionales intrínsecas de CAC1, se estudiaron las propiedades de autoasociación de CAC1Aib mediante dos aproximaciones experimentales diferentes:

(A) Se siguió el cambio en la intensidad de fluorescencia a 420 nm del grupo antranilo con el aumento de la concentración de CAC1. A partir del ajuste de los datos a la ecuación (3.2), se obtuvo un valor para la constante de autodisociación de $8 \pm 3 \mu\text{M}$, que es similar, dentro del error experimental, a los valores obtenidos para el péptido sin marcar por DC y anisotropía.

(B) Se siguió el cambio en la anisotropía del grupo antranilo con el aumento de la concentración de CAC1. La constante de autodisociación, que se obtuvo a partir de la ecuación (3.2'), fue $9 \pm 3 \mu\text{M}$, similar a la determinada por fluorescencia.

Estos resultados indican que CAC1Aib y CAC1 se autoasocian de forma idéntica, y por tanto, ambos péptidos son equivalentes para los estudios de unión a CA-C.

3.2.2. UNIÓN DE CAC1 A CA-C.

Dos procedimientos experimentales fueron usados:

(A) En primer lugar, se estudió si CAC1 y CA-C interaccionaban mediante DC, cromatografía de exclusión por tamaño molecular y RMN.

(B) En segundo lugar, y una vez comprobado si había unión, se determinaron los valores de dicha interacción, mediante fluorescencia, cromatografía de afinidad analítica y calorimetría de titulación isotérmica.

3.2.2.1. Interacción CAC1-CA-C.

3.2.2.1.1. Dicroísmo circular.

Si se produce la unión entre CAC1 y CA-C, probablemente se alterará la estabilidad térmica de CA-C (Pantoliano et al., 2001). Si CAC1 se une al estado nativo de CA-C, y en ausencia de unión al estado desnaturalizado, necesariamente se producirá un aumento de la T_m ; por el contrario, la unión al estado desplegado de CA-C, en ausencia de unión al estado nativo, conllevará una disminución de la T_m respecto a CA-C aislada (Waldron y Murphy, 2003).

CA-C mostró una transición térmica única seguida por DC en el ultravioleta lejano (Fig. 2.9; Tabla 2.2); sin embargo, CAC1 no presentó transición térmica a ninguna de las concentraciones estudiadas (desde 40 hasta 192 μM). Cuando los experimentos de desnaturalización térmica se realizaron con una concentración de CA-C de 20 μM , y de CAC1 de 20 a 60 μM , la T_m disminuyó con el aumento de la concentración de CAC1 (Fig. 3.4). En todos los casos se produjo precipitación de la

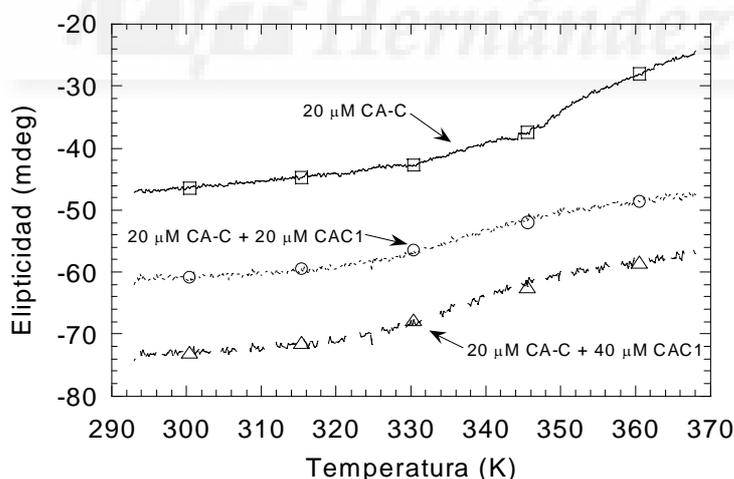


Figura 3.4 Estudio de la interacción CAC1-CA-C por desnaturalización térmica seguida mediante DC en el ultravioleta lejano. Elipticidad experimental a 222 nm frente a la temperatura para: CA-C sola, a 20 μM (línea continua con cuadrados); el complejo formado por 20 μM de CA-C y 20 μM de CAC1 (línea punteada con círculos); y, el complejo formado por 20 μM de CA-C y 40 μM de CAC1 (línea discontinua con triángulos).

muestra después del calentamiento. No se sabe si esta precipitación es el resultado de la interacción de CAC1 cuando CA-C está parcialmente desplegada (es decir, sólo especies de CA-C parcialmente plegadas serían competentes para la unión) o, si por el contrario, CAC1 y CA-C interaccionan y el complejo presenta agregación irreversible después de la desnaturalización inducida por calor (es decir, la unión ocurriría antes de que CA-C se despliegue). En cualquier caso, el decrecimiento del valor de T_m es una consecuencia de la interacción.

3.2.2.1.2. Cromatografía de exclusión por tamaño molecular.

Como se ha comentado en el apartado 2.2.1.5. del capítulo anterior, el V_e de CA-C a 20 μM y pH 7 fue 13.16 ml (Fig. 3.5). A medida que se fue incrementando la concentración de CAC1, el V_e del heterocomplejo aumentó: desde 13.22 ml (a 60 μM de CAC1) hasta 13.52 ml (a 128 μM de CAC1) (Fig. 3.5). Como se ha demostrado que la precisión inherente en la determinación del tamaño de proteínas globulares por métodos de partición en gel es ± 0.1 ml (Ackers, 1967), las diferencias encontradas en los V_e en nuestros experimentos son muy importantes. Estos resultados indican que el péptido interacciona con CA-C, y que esta interacción provoca la disociación del dímero de CA-C, ya que el mayor V_e indica una reducción en el tamaño de la macromolécula.

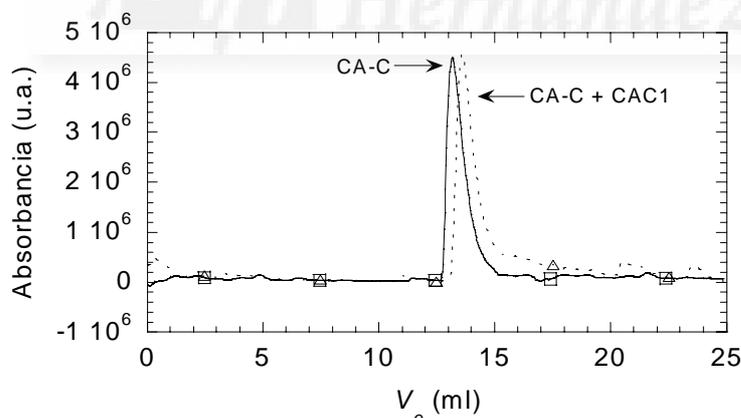


Figura 3.5 Estudio de la interacción CAC1-CA-C por cromatografía de exclusión por tamaño molecular. Perfiles de elución de: CA-C a 20 μM (línea continua con cuadrados), y el complejo formado por 20 μM de CA-C y 128 μM de CAC1 (línea punteada con triángulos). Los hombros pequeños que aparecen a la derecha del pico, en este segundo caso, probablemente se corresponden con CAC1 no unido. La escala en el eje de ordenadas, donde se representa la absorbancia a 280 nm, es arbitraria. Los experimentos se realizaron a 298 K.

3.2.2.1.3. Resonancia magnética nuclear.

Si CAC1 heterodimeriza con CA-C, y esto provoca la disociación del homodímero de CA-C (como se deduce de los experimentos de cromatografía de exclusión por tamaño molecular), las señales de RMN cambiarán (en los desplazamientos químicos, en la anchura de línea, y/o en la intensidad). Se realizaron experimentos con 10 μM de CAC1, 10 μM de CA-C, y la mezcla de ambos (10 μM de CAC1 y 10 μM de CA-C) (Fig. 3.6). El espectro de CAC1 ya se ha descrito en el apartado 3.2.1.3.

El espectro monodimensional de CA-C a 10 μM presentó señales de metilos a campo alto, a 0.2 ppm y 0.5 ppm (Fig. 3.6 D). Se observó una gran dispersión de la señal de los protones amida entre 8.0 y 9.0 ppm, y el espectro no presentaba protones entre 5.0 y 6.0 ppm. Además, varios protones amídicos fueron visibles después de 8 horas de disolver la proteína en $^2\text{H}_2\text{O}$. Estas son las características que se esperan para una proteína plegada en hélice α , de acuerdo con su estructura de rayos X (Gamble et al., 1997).

El espectro monodimensional del complejo formado entre CAC1 y CA-C, que se obtuvo mezclando 10 μM de CAC1 y 10 μM de CA-C (Fig. 3.6 C), fue distinto a la suma de espectros (Fig. 3.6 B) obtenida sumando el espectro de CAC1 (Fig. 3.6 A) y de CA-C (Fig. 3.6 D). Las principales diferencias se observaron en la región entre 1.5 y 2.0 ppm, donde están presentes los protones metileno correspondientes a las cadenas laterales de los aminoácidos con una larga cadena alquílica, tales como la lisina o la arginina (Wüthrich, 1986). También se observaron cambios menores en la región entre 0.8 y 1.0 ppm (Figs. 3.6 B, C). En la región de los protones H_α , donde normalmente se detectan los cambios en la estructura secundaria, no hubo cambios cuando se formó el complejo, probablemente debido a que CA-C es una proteína en hélice α , y hay un gran número de resonancias que solapan en la región entre 3.6 y 4.4 ppm. El hecho de que haya diferencias experimentales entre el espectro del complejo formado entre CAC1 y CA-C, y el espectro obtenido por la suma de los espectros de las especies individuales (CAC1 y CA-C), indica que hay unión entre CAC1 y CA-C.

También se realizaron experimentos con 100 μM de CA-C, añadiendo cantidades cada vez mayores de CAC1, en el rango de 100 a 200 μM . Sin embargo, en estas condiciones no se apreciaron cambios, ni en los desplazamientos químicos, ni en

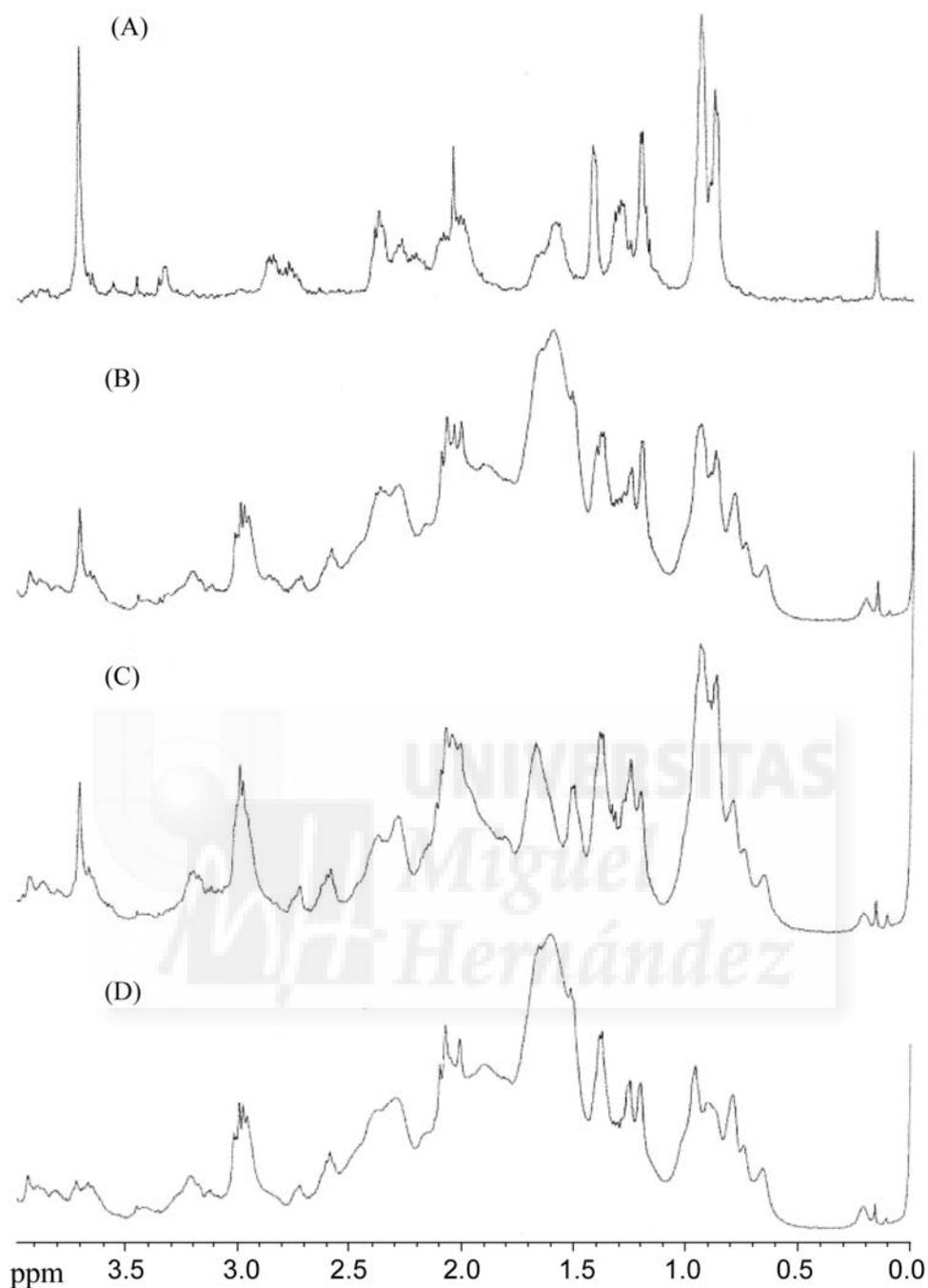


Figura 3.6 Estudio de la interacción CAC1-CA-C por RMN. (A) Región de campo alto del espectro de CAC1 a 10 μM . (B) Región de campo alto del espectro obtenido por la suma de los espectros de CAC1 (10 μM) y CA-C (10 μM). (C) Región de campo alto del espectro del complejo formado entre CAC1 y CA-C a la misma concentración (cada macromolécula a 10 μM). (D) Región de campo alto del espectro de CA-C a 10 μM . La señal aguda que se observa a la derecha en todos los espectros, excepto en el de CAC1, es la referencia interna de RMN (TSP). La intensidad en cada espectro de RMN es arbitraria. Los experimentos se realizaron a 298 K.

la anchura de línea, ni en la intensidad de la señal. Por lo tanto, a concentraciones altas no se obtuvo evidencia de la formación del complejo por RMN. En otras proteínas se ha observado un comportamiento similar al estudiar procesos de unión a altas concentraciones por RMN (Arac et al., 2003), y se ha interpretado como un límite para la constante de afinidad aparente de los complejos formados. Esto sugiere que la constante de disociación aparente del complejo CAC1-CA-C es menor de 100 μM .

3.2.2.2. Determinación de la constante de disociación aparente del complejo CAC1-CA-C.

Los experimentos de desnaturalización térmica seguidos por DC, cromatografía de exclusión por tamaño molecular y RMN, indican la existencia de interacción entre CAC1 y CA-C, pero ninguna de estas técnicas permitió obtener un valor cuantitativo para la constante de disociación. Con esta finalidad, se utilizaron tres aproximaciones experimentales adicionales:

(A) Se siguió el cambio en la intensidad de fluorescencia a 420 nm del grupo antrnilo de CAC1Aib, después de la unión a CA-C. No se observó precipitación en el rango de concentración de proteína estudiado (Fig. 3.7 A). La constante de disociación aparente del heterocomplejo, que se obtuvo a partir de la ecuación (3.2''), fue $30 \pm 10 \mu\text{M}$.

(B) Se realizaron experimentos de cromatografía de afinidad analítica, con una columna a cuya fase gel se había unido covalentemente CA-C. No se observó precipitación, y los datos se ajustaron bien a la ecuación (3.4) (Fig. 3.7 B). El valor obtenido para la constante de disociación aparente del heterocomplejo fue de $43 \pm 2 \mu\text{M}$.

(C) Se midió el calor intercambiado en la reacción de formación del complejo CAC1-CA-C por ITC (Fig. 3.7 C). A pesar de que se observó precipitación a altas concentraciones de CAC1, se pudo determinar la constante de disociación aparente, dando un valor de $79 \pm 7 \mu\text{M}$, que es algo mayor que los obtenidos por el resto de las técnicas, probablemente debido a la precipitación observada. Sin embargo, la precipitación se evitó en los experimentos de afinidad a causa de la unión a la fase estacionaria, y en los experimentos de fluorescencia por las bajas concentraciones utilizadas de CAC1Aib y CA-C.

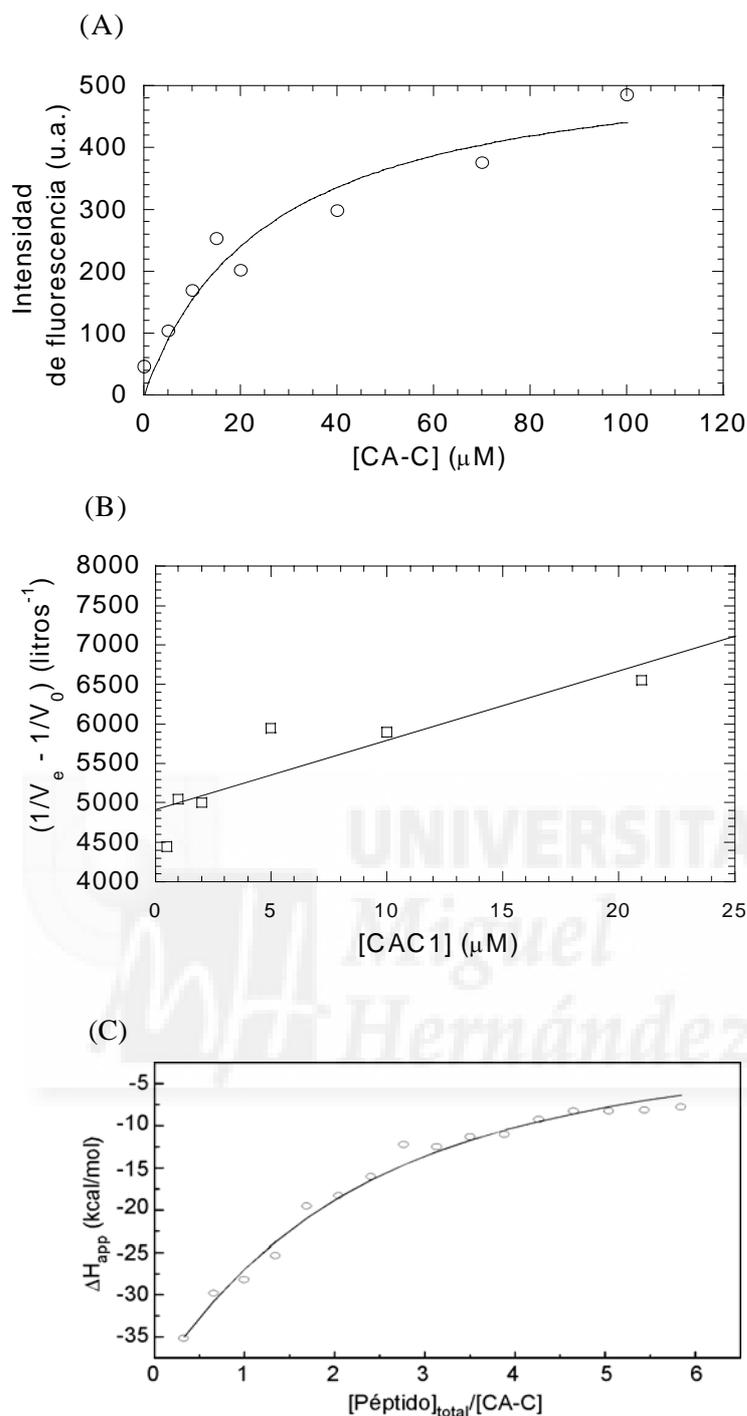


Figura 3.7 Determinación de la constante de disociación aparente del complejo CAC1-CA-C. (A) Fluorescencia en el estado estacionario: Intensidad de fluorescencia a 420 nm de CAC1Aib (en unidades arbitrarias) frente a la concentración de CA-C. Los espectros se tomaron con una longitud de onda de excitación de 320 nm (la curva fue similar cuando se excitó a 246 nm). La línea es el ajuste a la ecuación (3.2''). (B) Cromatografía de afinidad analítica. La línea es el ajuste a la ecuación (3.4). (C) Calorimetría de titulación isotérmica. La isoterma resultante se ajustó a un modelo de un único sitio de unión asumiendo que el complejo CAC1-CA-C tenía una estequiometría 1:1. Todos los experimentos se realizaron a 298 K.

3.3. DISCUSIÓN.

3.3.1. ESTRUCTURA DE CAC1.

CAC1 se encuentra mayoritariamente desordenado en disolución, como se observó por DC, fluorescencia y RMN. Sin embargo, el espectro de DC presentó bandas intensas que no son características de estructuras totalmente desordenadas, y que probablemente se deben a la presencia del único residuo de triptófano (Freskgard et al., 1994; Woody, 1995; Kelly y Price, 2000). Conforme fue aumentando la concentración de péptido, el valor absoluto de la elipticidad a 222 nm fue haciéndose progresivamente más pequeño (Fig. 3.2), sugiriendo la reducción del nivel de estructura helicoidal o un cambio en el entorno del único residuo de triptófano que posee el péptido (Freskgard et al., 1994; Woody, 1995; Kelly y Price, 2000); de hecho, se apreció un cambio en el entorno del triptófano con el aumento de la concentración de CAC1, ya que hubo un pequeño desplazamiento del espectro de emisión fluorescente del grupo indol hacia menores longitudes de onda, y un aumento en el valor de la anisotropía (Fig. 3.3). Por otra parte, CAC1 aislado en disolución presentaba poca tendencia a formar estructura helicoidal, como se vio por: (i) los espectros de DC (Fig. 3.2); y, (ii) predicciones teóricas realizadas utilizando el programa AGADIR (Muñoz y Serrano, 1994), que dieron un porcentaje muy bajo (2.2 %) de estructura helicoidal. Todos estos datos sugieren que, probablemente, el entorno del triptófano cambia con la autoasociación del péptido, pero no sabemos si para formar una estructura helicoidal como la nativa. En algunas proteínas o péptidos modelo, durante su plegamiento se induce la formación de estructura nativa alrededor de las cadenas laterales voluminosas de los residuos aromáticos (Simmerling y Elber, 1995; Kemming y Creighton, 1996). Tal y como se ha observado en algunos péptidos que pueden interactuar con proteínas de una gran variedad de formas (Stanfield y Wilson, 1995), probablemente en CAC1, el entorno del triptófano (y por tanto, la estructura nativa alrededor del mismo) se conserva después de su unión a CA-C.

El valor medio para la constante de autodisociación aparente del péptido (de los obtenidos por las diferentes técnicas biofísicas utilizadas) fue de $12 \pm 5 \mu\text{M}$, que es similar al hallado para la constante de disociación de CA-C ($10 \mu\text{M}$) (Gamble et al., 1997), sugiriendo que la homodimerización de CA-C está

principalmente dirigida por interacciones entre las dos α -hélices 2, una de cada monómero.

No pudimos determinar el estado de oligomerización de CAC1; no obstante, los estudios de cromatografía de exclusión por tamaño molecular sugieren que CAC1 se comporta como un dímero a altas concentraciones, ya que la determinación del peso molecular de CAC1 a partir del V_e observado, utilizando el peso molecular de los patrones conocidos, resultó en un valor de 4800 ± 500 Da, que es muy similar al peso molecular esperado para el dímero cuyo monómero tiene 2318.56 Da.

3.3.2. RECONOCIMIENTO MOLECULAR DE CA-C POR CAC1.

Hemos mostrado que CAC1 y CA-C interaccionan. Sin embargo, se podría pensar que la tendencia intrínseca de CAC1 a autoasociarse es la responsable de la unión a CA-C a través de asociación inespecífica. Para comprobar la ausencia de esta asociación inespecífica, se realizaron experimentos con EI, una proteína dimérica implicada en sistemas de fosforilación en bacterias, y sin homología con CA-C. Los experimentos control consistieron en determinar el valor de la constante de *quenching* para el grupo antrnilo de CAC1Aib a diferentes condiciones: (i) en presencia de EI; (ii) en presencia de CA-C; y, (iii) con CAC1Aib solo. La constante de Stern-Volmer, tanto en presencia de EI como en ausencia de CA-C, fue menor ($1.0 \pm 0.1 \text{ M}^{-1}$) que en presencia de CA-C ($1.7 \pm 0.1 \text{ M}^{-1}$). No sabemos por qué la constante de *quenching* del grupo antrnilo aumentó al unirse a CA-C, pero probablemente se debe a que con la unión, el péptido adquirió una conformación nativa, ya que en la estructura de rayos X de CA-C, la cadena principal del residuo de lisina (Lys182) se encuentra completamente enterrada, pero su cadena lateral está parcialmente expuesta (Gamble et al., 1997). Esto sugiere que en la unión entre péptido y proteína están implicadas las cadenas laterales.

Por otra parte, los experimentos de RMN indican igualmente que las regiones más importantes implicadas en la unión son las cadenas laterales de los aminoácidos de ambas especies (péptido y proteína), como se deduce de los cambios en la región de campo alto del espectro (Fig. 3.6). Asimismo, en otros estudios se ha visto que las regiones más afectadas con la unión de péptidos son las cadenas laterales de los aminoácidos de las macromoléculas (Stanfield y Wilson, 1995; Davis y Teague, 1999).

La similitud entre la constante de autodisociación aparente de CAC1 ($12 \pm 5 \mu\text{M}$) y la de CA-C ($10 \mu\text{M}$) (Gamble et al., 1997), y la diferencia relativamente pequeña entre la constante de disociación aparente del complejo CAC1-CA-C

($50 \pm 25 \mu\text{M}$) y la del homodímero de CA-C, indica que muchas interacciones importantes para el reconocimiento de CA-C están contenidas dentro de la α -hélice 2, de acuerdo con estudios previos del homodímero de CA-C, tanto estructurales (Gamble et al., 1997), como termodinámicos (del Álamo et al., 2003). Por lo tanto, el péptido mimetiza la mayor parte de las interacciones que forman la interfase de dimerización de CA-C. La diferencia entre los valores de afinidad obtenidos para el heterodímero péptido-proteína y el homodímero proteico puede deberse a consideraciones entrópicas, principalmente al gran gasto de entropía conformacional que se paga después de la unión de CAC1 a CA-C, debido a la mayor flexibilidad de CAC1 aislado cuando se compara con el monómero de CA-C. También, la diferencia debe depender de la contribución de otros residuos críticos para la dimerización presentes en el dominio intacto, pero ausentes en el péptido, como Ile150, Leu151, Arg154 y Leu172 (del Álamo et al., 2003). Las diferencias estructurales entre el péptido y el dominio intacto deben ser también en parte responsables de la diferencia entre ambas constantes. En particular, es necesario tener en cuenta que el péptido contiene los residuos interfásicos Ser178, Glu180, Glu187 y Gln192, que se ha visto que dificultan el proceso de dimerización de CA-C, ya que la mutación de cualquiera de ellos a alanina supuso un incremento en la afinidad de homodimerización entre 2 y 20 veces (del Álamo et al., 2003). El entorno estructural de estos residuos en CAC1 es, probablemente, diferente del que poseen en la forma dimerica de CA-C, y su efecto sobre la afinidad no se conoce. Cualesquiera que sean las razones exactas, es importante destacar que, a pesar de la carencia de algunos residuos críticos para la homodimerización, y de no estar completamente plegada en solución, la α -hélice 2 aislada de CA-C es capaz de unirse específicamente al dominio con gran afinidad.

Finalmente, se puede plantear la cuestión de cuántas moléculas de cada especie (péptido y proteína) forman el heterocomplejo. Hay varias líneas de evidencia que sugieren que el heterocomplejo se forma por una sola molécula de cada especie: (i) el V_e del complejo formado entre CAC1 y CA-C en los experimentos de cromatografía de exclusión por tamaño molecular (Fig. 3.5); (ii) el hecho de que en los experimentos de unión seguidos por la emisión de fluorescencia del grupo antrnilo no se observa fase estacionaria a bajas concentraciones de péptido antes de la curva sigmoideal (Fig. 3.7 A), cuya existencia indicaría la formación de dímeros de CAC1; y, (iii) por último, también hay que tener en cuenta que CA-C unida a la columna

permanece en forma monomérica, y por lo tanto, el heterocomplejo se forma con una sola molécula de CA-C, como se deduce de los buenos ajustes obtenidos de los datos experimentales a la ecuación (3.4), y a partir de estudios previos de cromatografía de afinidad analítica en CA-C (Rosé et al., 1992).



***DISEÑO, EXPRESIÓN Y
PURIFICACIÓN DE
MINIPROTEÍNAS***



Para obtener miniproteínas con la capacidad de heterodimerizar con CA-C silvestre e inhibir de este modo la dimerización natural de la cápsida, y en consecuencia, la infectividad del VIH, se ha llevado a cabo el diseño de miniproteínas (o fragmentos), basadas en la estructura tridimensional de CA-C, todas ellas conteniendo la α -hélice 2. Estas miniproteínas se han generado mediante técnicas de ingeniería de proteínas, y se ha intentado llevar a cabo su expresión y purificación.

Ninguna de las mutaciones diseñadas para la obtención de miniproteínas que sólo contuvieran hasta la segunda hélice α de CA-C permitió la expresión de proteína; no obstante, se ha conseguido la expresión y puesta a punto del protocolo de purificación del fragmento de CA-C que contiene los residuos 146-214 del dominio intacto.

4.1. MATERIALES Y MÉTODOS.

4.1.1. DISEÑO DE PROTEÍNAS MUTANTES.

Se siguieron dos aproximaciones diferentes, introduciendo:

(A) Un codon *STOP* en la posición deseada de la cadena. La ventaja principal de esta aproximación es que se obtiene de forma directa la expresión del fragmento proteico deseado, evitando así posteriores pasos de digestión y purificación que complican el proceso de obtención de la miniproteína.

(B) Un residuo de metionina en la posición deseada de la cadena para obtener los fragmentos posteriormente mediante ruptura química con BrCN (Gross, 1967; Liao et al., 1973; Doyen y LaPresle, 1979; Joppich-Kuhn et al., 1982).

Con el fin de obtener proteínas de menor tamaño que CA-C, cada una de ellas con un elemento de la estructura secundaria añadido, las mutaciones se introdujeron detrás de: (i) la α -hélice 2; o, (ii) la α -hélice 3 de CA-C (Fig. 4.1), teniendo en cuenta que los aminoácidos a mutar debían estar localizados en lugares accesibles al disolvente, y no estar implicados en interacciones importantes a nivel terciario (puentes de hidrógeno y/o interacciones electrostáticas).

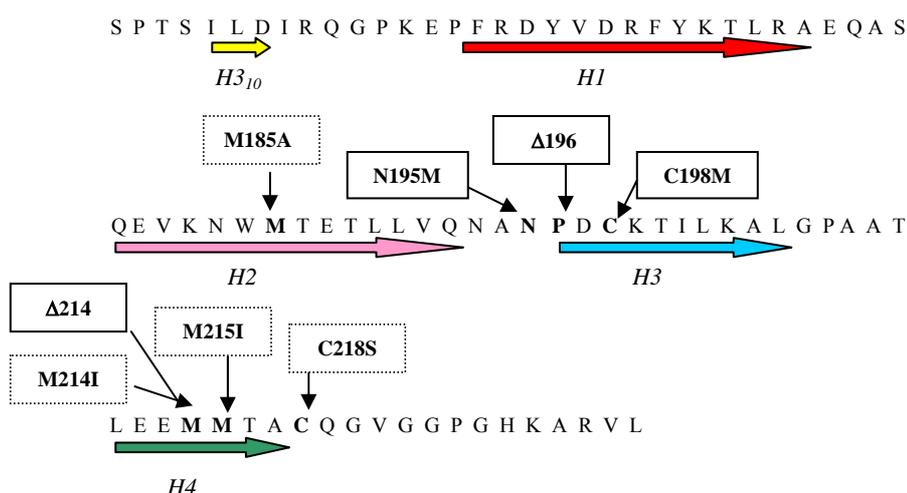


Figura 4.1 Mutantes diseñados sobre la secuencia de CA-C. Las distintas hélices α están representadas con flechas bajo los aminoácidos correspondientes en la secuencia y con colores diferentes: amarillo, hélice 3_{10} ($H3_{10}$); rojo, α -hélice 1 ($H1$); rosa, α -hélice-2 ($H2$); azul, α -hélice 3 ($H3$); y, verde, α -hélice 4 ($H4$). Las mutaciones introducidas se señalan con flechas sobre los aminoácidos seleccionados con la numeración de la proteína entera, y se representan en negrita. Las mutaciones rodeadas por un cuadrado con línea continua se corresponden con la introducción de un codon *STOP* (Δ) o de una metionina; y aquellas rodeadas por un cuadrado con línea discontinua se realizaron para conseguir las miniproteínas deseadas (ver apartados 4.2.1.2., 4.2.1.3., 4.2.2.2. y 4.2.2.3).

4.1.2. MUTAGÉNESIS DIRIGIDA.

4.1.2.1. Proceso de mutagénesis.

La obtención de las proteínas mutantes se realizó con la ayuda de Marco Caprini (IBMC-UMH, Elche), y se llevó a cabo mediante mutagénesis dirigida de CA-C silvestre, utilizando el método de la PCR inversa. Brevemente, el proceso de mutagénesis dirigida (Fig. 4.2) utiliza un plásmido de ADN de doble cadena y dos oligonucleótidos sintéticos complementarios (directo y reverso) que contienen la mutación deseada y que actúan como cebadores. En primer lugar, la cadena doble del ADN molde es separada en hebras sencillas por calentamiento a 95 °C. Entonces se produce la amplificación mediante ciclos sucesivos de hibridación (55 °C), extensión de la cadena por acción de la polimerasa (68 °C) y separación de las hebras producidas (95 °C). Así se obtiene un conjunto de nuevas hebras de ADN mutadas y melladas. A continuación se lleva a cabo la digestión de las hebras molde con la endonucleasa de

restricción *DpnI*, que es específica para ADN metilado. Finalmente, se introduce el plásmido mutado, mediante transformación, en células supercompetentes de la cepa *E. coli* XL1-Blue, donde se produce la reparación del ADN mellado. La pequeña cantidad de ADN molde necesaria para la PCR, y la gran fidelidad de la polimerasa, contribuyen a la alta eficiencia de la mutagénesis y a la baja posibilidad de mutaciones aleatorias durante la reacción. Aún así, los mutantes fueron secuenciados para comprobar la ausencia de cualquier otro tipo de mutación no deseada en el plásmido.

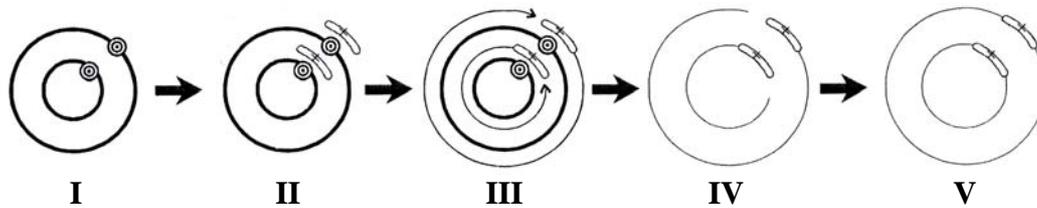


Figura 4.2 Esquema del proceso de mutagénesis dirigida. (I) Plásmido de doble cadena de ADN donde se muestra el sitio donde se va a introducir la mutación. (II) Separación de las dos hebras del ADN molde por calentamiento, e hibridación de los oligonucleótidos que contienen la mutación deseada. (III) Extensión de los cebadores por acción de la polimerasa dando lugar a hebras circulares mutadas y melladas. (IV) Tras la digestión con *DpnI* del ADN molde, sólo quedan los plásmidos mutados y mellados. (V) Después de la transformación en la cepa bacteriana *E. coli* XL1-Blue, las mellas de las dobles hebras de ADN mutantes son reparadas.

4.1.2.2. Diseño de oligonucleótidos.

Los oligonucleótidos de mutagénesis usados para la realización de este trabajo (Tabla 4.1), fueron diseñados teniendo en cuenta las siguientes consideraciones:

(A) Ambos oligonucleótidos (directo y reverso) deben contener la mutación deseada, y sus secuencias deben ser complementarias entre sí y a las de las hebras codificante y complementaria del gen de CA-C, respectivamente, en todas sus bases excepto en aquellas que codifican para el aminoácido a mutar.

(B) La longitud del oligonucleótido debe ser lo suficientemente grande para garantizar una hibridación eficiente, evitando oligonucleótidos demasiado largos, puesto que puede producirse autohibridación. En consecuencia, se suelen utilizar oligonucleótidos que contienen entre 25 y 45 bases.

(C) La temperatura de hibridación (T_m) debe ser igual o superior a 78 °C, calculada a partir de la siguiente fórmula:

$$T_m = 81.5 + 0.41 (\% GC) - 675/N - \% M$$

donde % GC es el porcentaje de guaninas y citosinas que posee el oligonucleótido; N es el número de bases del mismo; y, % M es el porcentaje de bases mutadas en el oligonucleótido.

(D) Los oligonucleótidos deben tener un contenido mínimo en guaninas y/o citosinas del 40 %.

(E) Las bases mutadas deben quedar aproximadamente en el centro del oligonucleótido, y se debe procurar que en los extremos del mismo aparezcan una o varias guaninas y/o citosinas, para promover uniones fuertes en ambos extremos.

Tabla 4.1 Oligonucleótidos diseñados^a.

Oligonucleótidos	Secuencia (5'→3') ^b	Mutación
Δ196D	CTTGTTGGTCCAAAATGCGAACT G AGATTGTAAGACTATTTTAAAAGC	
Δ196R	GCTTTTAAAATAGTCTTACAAT TCTC AGTTCGCATTTTGGACCAACAAG	Δ196
M214I/M215ID	GCGGCTACACTAGAAGAA ATCATC ACAGCATGTCAGGG	
M214I/M215IR	CCCTGACATGCTGT GATGAT TTTCTTCTAGTGTAGCCGC	M214I/M215I
N195MD	GAAACCTTGGTTGGTCCAAAATGCG ATG CCAGATTGTAAGAC	
N195MR	GTCTTACAATCTGG CAT CGCATTTTGGACCAACAAGGTTTC	N195M
C198MD	GCGAACCCAGAT ATGA AAGACTATTTTAAAAGCATTGGGACCAGCG	
C198MR	CGCTGGTCCCAATGCTTTTAAAATAGTCT TTCAT ATCTGGGTTTCGC	C198M
Δ214D	CAGCGGCTACACTAGAAGA ATGA ATGACAGCATGTCAGGGAG	
Δ214R	CTCCCTGACATGCTGTCAT TTCAT TCTTCTAGTGTAGCCGCTG	Δ214
C218SD	GAAATGATGACAGCAT CCC AGGGAGTAGGAGGACCC	
C218SR	GGGTCCTCCTACTCCCT GGG ATGCTGTCATCATTTTC	C218S

^aD, oligonucleótido directo; R, oligonucleótido reverso; Δ, codon *STOP* introducido en la posición indicada.

^bLa mutación se representa en negrita.

4.1.2.3. Método experimental.

Se utilizó el kit *QuickChange Site-Directed Mutagenesis* (Stratagene) sobre el plásmido recombinante pET21b(+)*CAC* (apartado 2.1.1.1.), siguiendo básicamente el

protocolo descrito por el fabricante. La mezcla de reacción, con un volumen final de 50 μ l, contenía en el tampón de PCR suministrado por el kit, unos 50 ng de ADN plasmídico, 125 ng de cada uno de los dos oligonucleótidos de mutagénesis complementarios (Tabla 4.1), 1 μ l de la mezcla de los cuatro desoxirribonucleósidos trifosfato (dNTPs) a la concentración suministrada por el kit, y 1 μ l de la ADN polimerasa *PfuTurbo* (con una actividad específica de 2.5 unidades/ μ l). La mezcla de reacción se introdujo en el termociclador (Eppendorf), y la reacción de PCR inversa se llevó a cabo mediante una incubación inicial de 30 s a 95 °C, seguida de 14 ciclos de amplificación (30 s a 95 °C, 1 min a 55 °C y 12 min a 68 °C). A continuación se añadió 1 μ l de la endonucleasa de restricción *DpnI* (10 unidades/ μ l) y la mezcla se incubó durante 1 hora a 37 °C. Inmediatamente se transformaron células supercompetentes de la cepa *E. coli* XL1-Blue con 1 μ l de la mezcla anterior, y se seleccionaron colonias de bacterias recombinantes (resistentes a ampicilina). Los plásmidos recombinantes obtenidos se purificaron utilizando el kit *HiSpeed Plasmid Midi* (Qiagen), y para comprobar la pureza, estos plásmidos fueron linearizados utilizando la enzima de restricción *BamHI* (Fermentas) siguiendo el protocolo recomendado por el fabricante, y posteriormente, visualizados mediante electroforesis en geles de agarosa al 1 % en un sistema electroforético Wide Mini SubCell GT (BioRad). La determinación de la concentración exacta de ADN plasmídico obtenido se realizó a partir de medidas de absorbancia a 260 nm, en un espectrofotómetro ultravioleta/visible Beckman DU 530 y con una cubeta de cuarzo de 1 cm de paso óptico (Hellma), considerando las referencias existentes para una concentración conocida: 1 unidad de absorbancia a 260 nm corresponde a 50 μ g/ml de ADN de doble cadena. Las mutaciones introducidas, así como el resto de la secuencia de la proteína, se confirmaron mediante secuenciación del ADN plasmídico correspondiente a la región de CA-C completa. Todas las reacciones de secuenciación fueron realizadas por Sistemas Genómicos S.L. (Parque Tecnológico de Valencia, Paterna, Valencia). Las secuencias se analizaron con el programa *Chromas*.

4.1.3. PRODUCCIÓN DE PROTEÍNAS MUTANTES.

4.1.3.1. Expresión y purificación de proteínas mutantes.

Los plásmidos pET21b(+)*CAC* correspondientes a CA-C mutantes secuenciados, se transformaron en la cepa *E. coli* BL21 (DE3). En primer lugar, se llevó a cabo la comprobación de que los mutantes expresaban la proteína en cantidad

suficiente, y sólo entonces, se procedió a la expresión del mutante de CA-C a gran escala para su posterior purificación.

4.1.3.1.1. Comprobación de la expresión.

Una colonia transformante recién obtenida se utilizó como inóculo para obtener un cultivo líquido en 30 ml de LB suplementado con ampicilina a una concentración final de 100 µg/ml. Este cultivo se incubó a 37 °C en agitación a 250 rpm hasta alcanzar una densidad óptica a 600 nm de 0.7-0.8, y en ese momento se tomó una alícuota de 200 µl, que se utilizó como control antes de la inducción. Entonces se indujo la expresión en el cultivo mediante la adición de IPTG a una concentración final de 1 mM. El cultivo se mantuvo en las mismas condiciones (temperatura y velocidad de agitación) durante 3 horas más, tomando cada hora una alícuota de 200 µl, utilizadas como control de la expresión proteica. Paralelamente, se realizó un cultivo en las mismas condiciones que el anterior, pero partiendo de una colonia transformada con el plásmido pET21b(+)CAC correspondiente a CA-C silvestre, para tener una referencia y poder comparar el nivel de expresión de CA-C mutado con el de CA-C no mutado.

Las alícuotas que se fueron tomando como control se centrifugaron a 12000 rpm durante 2 minutos en una centrífuga Minispin (Eppendorf), se descartó el sobrenadante, y las células se resuspendieron en 40 µl de tampón de carga (Tris 50 mM pH 6.8, glicerol 9.2 M, SDS 10 %, β-mercaptoetanol 5 % y azul de bromofenol 1 %). Estas alícuotas se guardaron congeladas a -20 °C hasta su utilización.

La presencia de proteína en la disolución obtenida se determinó mediante SDS-PAGE (gel concentrador: 4.5 %; gel separador: 15 %) (Laemmli, 1970), realizada en el sistema electroforético Mini Protean 3 (BioRad), conectado a una fuente PowerPac 300 (BioRad). Antes de cargar las alícuotas anteriores en el gel de poliacrilamida, éstas se calentaron a 100 °C durante 10 minutos. Como marcador molecular se utilizó un marcador de bajo peso molecular (Sigma). La electroforesis se realizó en tampón Tris-HCl 25 mM pH 8.3, glicina 200 mM, SDS 3.5 mM; y con un voltaje constante de 150 V. Para la tinción, el gel se mantuvo en agitación durante 15 minutos en una disolución formada por 0.1 % azul de coomassie, 40 % etanol, 20 % ácido acético. La destinción se realizó de la misma forma pero en la disolución 20 % etanol, 20 % ácido acético. La visualización de los geles se llevó a cabo en un transiluminador (Gelprinter Plus).

En los casos en los que no se observó expresión de CA-C mutante, se realizaron distintos ensayos modificando las condiciones utilizadas con anterioridad: (i) los plásmidos pET21b(+)*CAC* correspondientes a los mutantes de CA-C se transformaron en otras cepas celulares, *E. coli* C41 o *E. coli* C43; y, (ii) los distintos cultivos inoculados con las distintas cepas bacterianas se incubaron a dos temperaturas diferentes (30 o 37 °C).

4.1.3.1.2. Expresión a gran escala y purificación.

Sólo se expresaron a gran escala aquellos mutantes de CA-C que mostraron expresión en los miniensayos anteriores. Todo el proceso de expresión y purificación se llevó a cabo del mismo modo que para CA-C silvestre (apartados 2.1.1.1. y 2.1.1.2.).

4.1.3.2. Determinación de la pureza y concentración de proteínas mutantes.

Al igual que para CA-C silvestre, la pureza de las preparaciones de CA-C mutantes se determinó mediante SDS-PAGE y espectrometría de masas.

La concentración de proteína (monómero de CA-C mutante) se determinó a partir de medidas de absorbancia a 280 nm, utilizando un espectrofotómetro ultravioleta/visible Beckman DU 530 y una cubeta de cuarzo de 1 cm de paso óptico (Hellma). El coeficiente de extinción molar a 280 nm (ϵ_{280}) se calculó como se indica en el apartado 2.1.1.3. Para el mutante M185A/C218S de CA-C, que carece de puente disulfuro, el valor teórico fue: $\epsilon_{280} = 8400 \text{ M}^{-1} \text{ cm}^{-1}$.

4.1.4. DIGESTIÓN DE PROTEÍNAS MUTANTES CON BrCN Y PURIFICACIÓN DE LOS FRAGMENTOS OBTENIDOS.

Las proteínas mutantes de CA-C con metioninas, una vez purificadas fueron sometidas al proceso de ruptura química con BrCN, que hidroliza los enlaces peptídicos por el extremo carboxilo de los residuos de metionina (Fig. 4.3), quedando el fragmento proteico N-terminal con un residuo de homoserina-lactona en su extremo C-terminal, en lugar de la metionina original.

El protocolo utilizado para la digestión con BrCN es el siguiente: a la proteína mutante con una concentración entre 0.5 y 1.5 mg/ml, se le añadió ácido clorhídrico con la finalidad de disminuirle el pH, alcanzando un valor final entre pH 1 y 2. Este paso es necesario para que la proteína se despliegue totalmente, como se ha mostrado en el capítulo 2 de este trabajo, facilitando la acción posterior del BrCN. A continuación, a la disolución de proteína desplegada a pH ácido, se le añadió la cantidad correspondiente de BrCN (Fluka), utilizando una relación molar proteína: BrCN de 1:10. Esta mezcla de reacción se dejó incubando en ausencia de luz y a temperatura ambiente, entre 12 y 16

horas. Transcurrido ese tiempo, la reacción se detuvo enfriando a 4 °C, y se mantuvo en las mismas condiciones hasta su utilización.

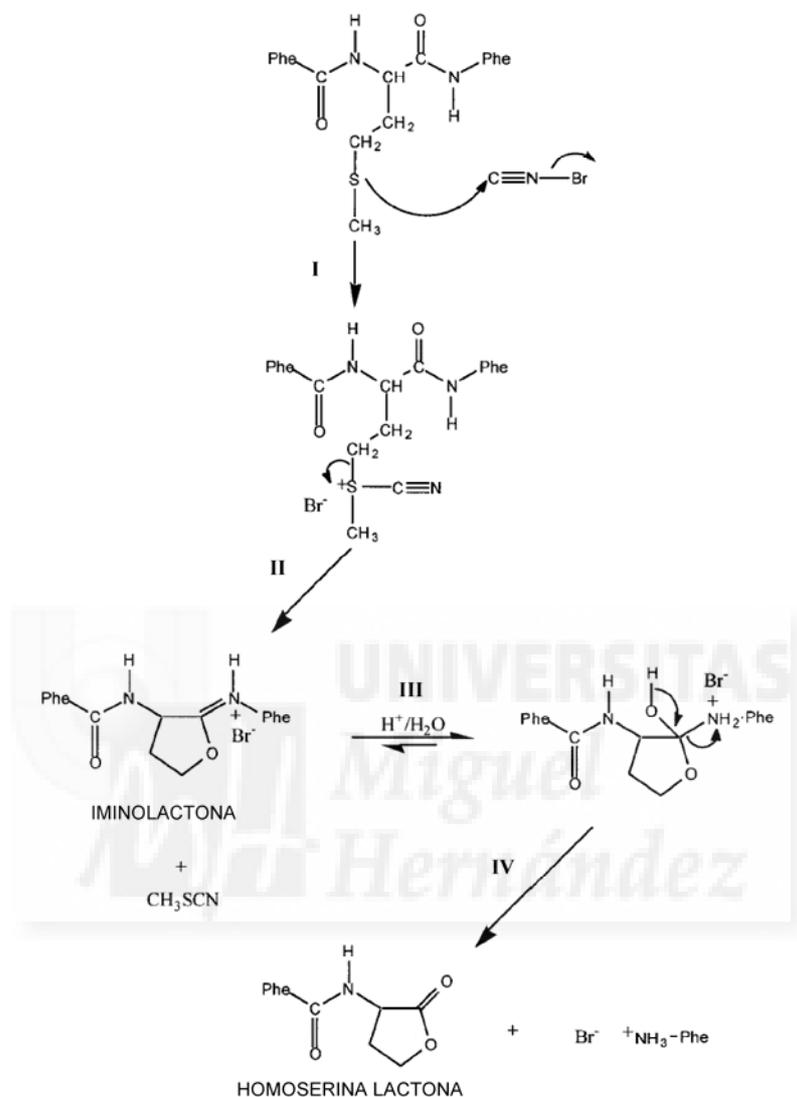


Figura 4.3 Reacción de ruptura química de un tripéptido teórico Phe-Met-Phe por BrCN.

(I) Ataque nucleofílico del átomo de azufre (S) del residuo de metionina por el átomo de carbono (C) del BrCN. (II) Reordenamiento intramolecular del complejo formado dando lugar a “iminolactona”, que es inmediatamente hidrolizada (III), produciéndose la ruptura de la cadena peptídica (IV), quedando el fragmento N-terminal con un residuo de “homoserina-lactona” en su extremo C-terminal.

Para la eliminación del BrCN, la disolución de proteína se cargó a temperatura ambiente en una columna de desalado de 53 ml (HiPrep 26/10 Desalting, Amersham Biosciences), previamente equilibrada en 25 mM fosfato potásico pH 6.0, acoplada a un

aparato de cromatografía AKTA-FPLC con un detector de absorbancia a 280 nm (Amersham Biosciences). La proteína se eluyó con un flujo de 5 ml/min. Las fracciones que contenían proteína se mezclaron, y sólo en los casos en los que fue necesario, esta mezcla se concentró a 4 °C mediante membranas Centricón YM-3 (Amicon, Millipore), hasta alcanzar la concentración apropiada para poder comprobar si la reacción de corte había funcionado correctamente, mediante espectrometría de masas.

A continuación, para la purificación de los fragmentos obtenidos, se siguieron dos procedimientos experimentales diferentes:

(A) Se usó una columna analítica de exclusión por tamaño molecular (Superdex 75 H/R 10/30, Amersham Biosciences), previamente equilibrada en 25 mM fosfato potásico, 150 mM NaCl, pH 6.0, aunque la composición de este tampón de elución se fue modificando, añadiendo: (i) un agente desnaturalizante (GdmHCl), a una concentración entre 1 y 2.5 M; y, (ii) un agente reductor (DTT o B-mercaptoetanol), a una concentración entre 1 y 5 mM. La proteína se eluyó con un flujo de 1 ml/min

(B) Se utilizó una columna de intercambio aniónico de 5 ml (Q-Sepharose, Amersham Biosciences), previamente equilibrada en 25 mM fosfato potásico pH 6.0. La proteína retenida se eluyó con un gradiente lineal de 0 a 0.5 M de NaCl, con un flujo de 1 ml/min.

En cualquier caso, las fracciones recogidas de las columnas y que contenían proteína, se dializaron frente a H₂O (desionizada y purificada utilizando un sistema Millipore) a 4 °C, y se comprobó que se correspondían con los fragmentos esperados mediante espectrometría de masas. Tras la comprobación de la pureza de los fragmentos obtenidos mediante SDS-PAGE, las alícuotas correspondientes se congelaron inmediatamente en nitrógeno líquido, y se almacenaron a -20 °C.

4.2. RESULTADOS.

4.2.1. OBTENCIÓN DE PROTEÍNAS MUTANTES DE CA-C QUE SÓLO CONTIENEN LOS ELEMENTOS DE ESTRUCTURA SECUNDARIA HASTA LA α -HÉLICE 2.

En primer lugar, y con el fin de obtener una miniproteína basada en la estructura de CA-C, que sólo contuviera hasta la α -hélice 2, se llevó a cabo la introducción de mutaciones en la región localizada detrás de dicha hélice α de CA-C (Fig. 4.1).

4.2.1.1. Diseño y análisis de la expresión del mutante $\Delta 196$.

El mutante $\Delta 196$ introdujo un codon *STOP* en la posición 196 (Fig. 4.1). En las mismas condiciones de expresión que CA-C silvestre (Fig. 4.4 A), el mutante $\Delta 196$ no se sobreexpresó (Fig. 4.4 B).

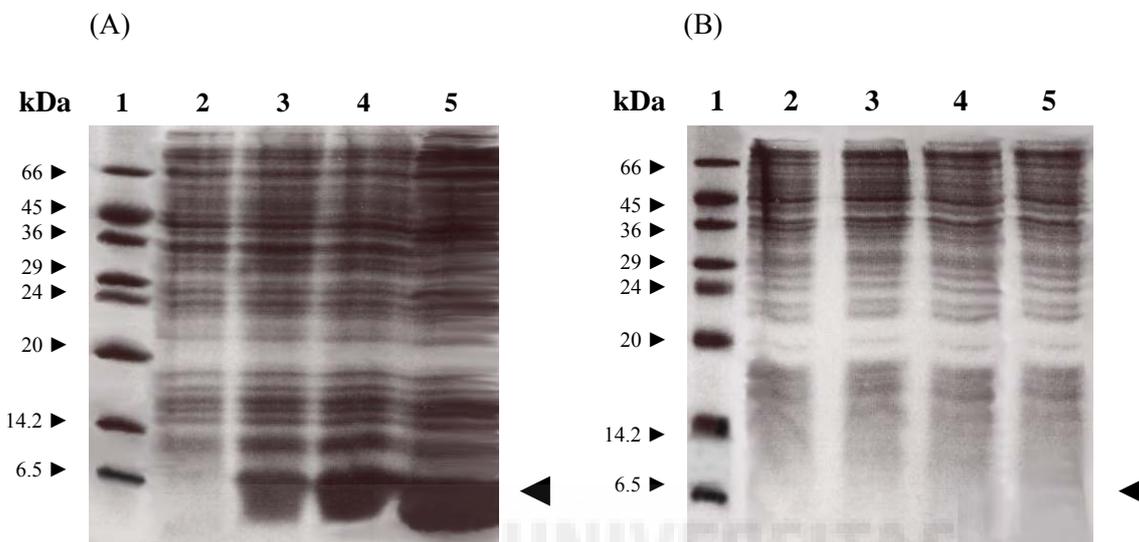


Figura 4.4 SDS-PAGE de extractos de bacterias recombinantes que contienen el ADN plasmídico correspondiente a CA-C silvestre (A) y al mutante $\Delta 196$ (B). Calle 1: marcadores de peso molecular. Calle 2: extracto celular control, antes de la inducción con IPTG. Las calles 3, 4 y 5 se corresponden con el extracto celular tras 1, 2 y 3 horas de inducción con IPTG, respectivamente. La flecha a la derecha indica la banda correspondiente a CA-C silvestre (A), o dónde cabría esperar que apareciera la banda del mutante $\Delta 196$ (B).

Entonces, se realizaron distintos ensayos, modificando las condiciones de cultivo iniciales: (i) utilizando diferentes cepas de *E. coli* (C41 o C43); y (ii) a distintas temperaturas de incubación (30 o 37 °C). En ningún caso se obtuvo sobreexpresión de proteína. No se conocen las razones exactas por las que no se consiguió expresión de este mutante de CA-C, pero probablemente se deba a que esta proteína no se pliegue correctamente, y en consecuencia, sea fácilmente degradada por la maquinaria celular. Por lo tanto, el mutante $\Delta 196$ fue descartado.

4.2.1.2. Diseño y análisis de la expresión del mutante M185A/M214I/M215I/N195M.

El siguiente mutante introdujo una metionina detrás de la α -hélice 2 de CA-C. Como en la secuencia de CA-C aparecen tres residuos de metionina (Met185, Met214 y

Met215), ninguno de ellos localizado en la posición de interés (Fig. 4.1), el primer paso consistió en la sustitución de estos aminoácidos.

A partir del plásmido pET21b(+)CAC correspondiente al mutante M185A, cedido por el Dr. Mauricio G. Mateu (CSIC-UAM, Madrid), se introdujeron dos nuevas mutaciones: M214I y M215I. La sustitución del residuo de metionina por isoleucina se eligió para preservar, en la medida de lo posible, el carácter hidrofóbico de la cadena lateral de la metionina original. Así, se obtuvo el triple mutante M185A/M214I/M215I, que carecía totalmente de metioninas. El siguiente paso consistió en la introducción de una nueva metionina en la posición 195 sobre el triple mutante anterior, dando lugar al mutante cuádruple M185A/M214I/M215I/N195M.

No se observó sobreexpresión de este mutante cuádruple (Fig. 4.5), ni utilizando diferentes cepas de *E. coli* (BL21 (DE3), C41 o C43), ni a distintas temperaturas de incubación (30 o 37 °C). Por lo tanto, este mutante de CA-C también se descartó.

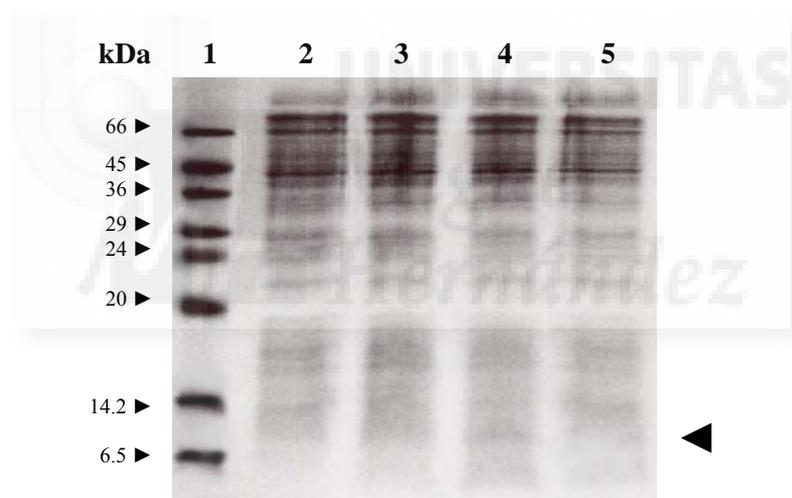


Figura 4.5 SDS-PAGE de extractos de bacterias recombinantes que contienen el ADN plasmídico correspondiente al mutante M185A/M214I/M215I/N195M. Calle 1: marcadores de peso molecular. Calle 2: extracto celular control, antes de la inducción con IPTG. Las calles 3, 4 y 5 se corresponden con el extracto celular tras 1, 2 y 3 horas de inducción con IPTG, respectivamente. La flecha a la derecha indica dónde cabría esperar que apareciera la banda correspondiente al mutante.

La ausencia de expresión del mutante cuádruple de CA-C fue a consecuencia de la última mutación introducida (N195M), ya que el triple mutante

M185A/M214I/M215I sí que mostró sobreexpresión (Fig. 4.6). No se conoce la razón por la que la mutación N195M impide la expresión de la proteína.

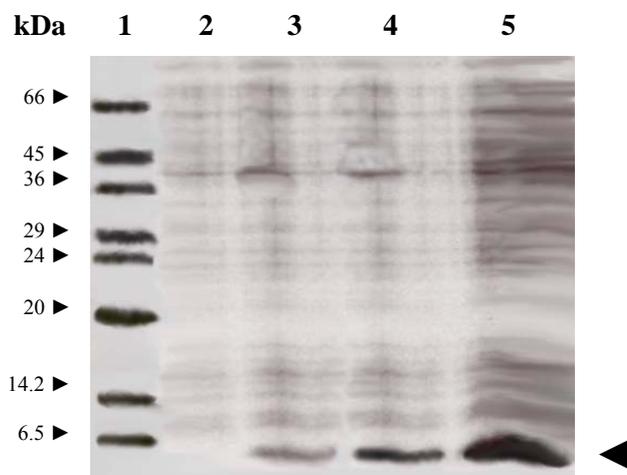


Figura 4.6 SDS-PAGE de extractos de bacterias recombinantes que contienen el ADN plasmídico correspondiente al mutante M185A/M214I/M215I. Calle 1: marcadores de peso molecular. Calle 2: extracto celular control, antes de la inducción con IPTG. Las calles 3, 4 y 5 se corresponden con el extracto celular tras 1, 2 y 3 horas de inducción con IPTG, respectivamente. La flecha a la derecha indica la banda correspondiente al mutante.

4.2.1.3. Diseño y análisis de la expresión del mutante M185A/M214I/M215I/C198M.

Debido a la ausencia de expresión obtenida en los mutantes anteriores, se diseñó otro mutante en el que el residuo de metionina se introdujo en la posición 198 (Fig. 4.1) sobre el triple mutante M185A/M214I/M215I, dando lugar al mutante M185A/M214I/M215I/C198M. Se eligió mutar el aminoácido localizado en la posición 198 porque con la eliminación del residuo de cisteína se consigue evitar la formación del puente disulfuro intramolecular de CA-C, que podría dificultar la posterior purificación de los fragmentos una vez se hubiera cortado la proteína con BrCN.

Tras varios intentos en diferentes combinaciones de cepas de *E. coli* (BL21 (DE3), C41 o C43), y a temperaturas de incubación (30 o 37 °C), no se observó sobreexpresión del mutante M185A/M214I/M215I/C198M (Fig. 4.7), y en consecuencia, también fue descartado.

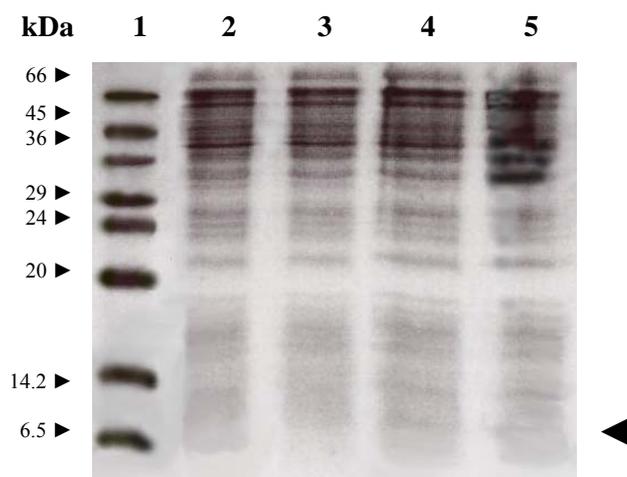


Figura 4.7 SDS-PAGE de extractos de bacterias recombinantes que contienen el ADN plasmídico correspondiente al mutante M185A/M214I/M215I/C198M. Calle 1: marcadores de peso molecular. Calle 2: extracto celular control, antes de la inducción con IPTG. Las calles 3, 4 y 5 se corresponden con el extracto celular tras 1, 2 y 3 horas de inducción con IPTG, respectivamente. La flecha a la derecha indica dónde cabría esperar que apareciera la banda correspondiente al mutante.

Del mismo modo que se observó en el apartado 4.2.1.2., la mutación adicional introducida (C198M) sobre el triple mutante M185A/M214I/M215I, fue la causante de la ausencia de expresión de este mutante cuádruple de CA-C (Fig. 4.6).

4.2.2. OBTENCIÓN DE PROTEÍNAS MUTANTES DE CA-C QUE SÓLO CONTIENEN LOS ELEMENTOS DE ESTRUCTURA SECUNDARIA HASTA LA α -HÉLICE 3.

Dado que la introducción de mutaciones en posiciones cercanas a la α -hélice 2 parecía afectar a la estabilidad y plegamiento de la proteína, se decidió abordar el problema del diseño de miniproteínas mediante la introducción de las mutaciones en la región localizada detrás de α -hélice 3 de CA-C (Fig. 4.1).

4.2.2.1. Diseño y análisis de la expresión del mutante Δ 214.

Se decidió introducir un codon *STOP* en la posición 214, porque previamente se vio que una mutación en esta posición (M214I) no afectaba a la expresión de la proteína (Fig. 4.6), dando lugar al mutante Δ 214.

No se obtuvo sobreexpresión del mutante $\Delta 214$ (Fig. 4.8), ni con diferentes cepas de *E. coli* (BL21 (DE3), C41 o C43), ni a distintas temperaturas de incubación (30 o 37 °C). Por lo tanto, este mutante se descartó.

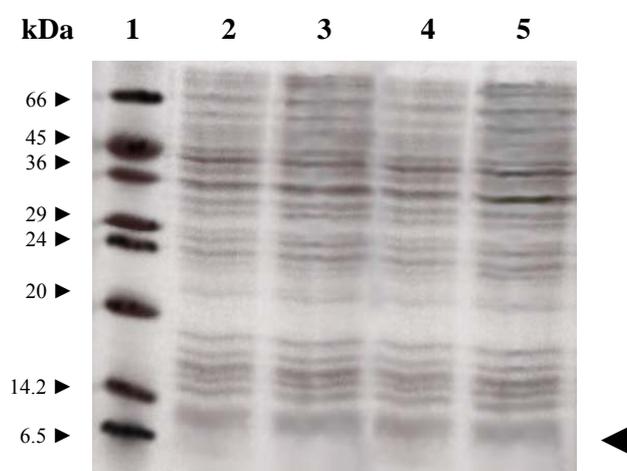


Figura 4.8 SDS-PAGE de extractos de bacterias recombinantes que contienen el ADN plasmídico correspondiente al mutante $\Delta 214$. Calle 1: marcadores de peso molecular. Calle 2: extracto celular control, antes de la inducción con IPTG. Las calles 3, 4 y 5 se corresponden con el extracto celular tras 1, 2 y 3 horas de inducción con IPTG, respectivamente. La flecha a la derecha indica dónde cabría esperar que apareciera la banda correspondiente al mutante.

4.2.2.2. Obtención y digestión del mutante M185A.

4.2.2.2.1. Diseño, análisis de la expresión, expresión a gran escala y purificación.

Como segunda aproximación, se pensó en la obtención de la miniproteína a partir de las metioninas localizadas en las posiciones 214 y 215, que están situadas detrás de la α -hélice 3 (Fig. 4.1). Sin embargo, debido a que en la secuencia de CA-C aparece otro residuo de metionina en la posición 185 (Met185), el primer paso consistió en la sustitución del mismo. Para ello, se utilizó el plásmido pET21b(+)CAC correspondiente al mutante M185A, cedido por el Dr. Mauricio G. Mateu (CSIC-UAM, Madrid). Este mutante es esencialmente monomérico (Gamble et al., 1997; del Álamo et al., 2003), ya que la cadena lateral del residuo Met185, localizado en el centro de la α -hélice 2 de CA-C (Fig. 4.1), se encuentra enterrada en la interfase dimerica participando en contactos hidrofóbicos entre los monómeros.

El mutante M185A se expresaba bien en *E. coli* BL21 (DE3) (Fig. 4.9), por lo tanto, se llevó a cabo la expresión y posterior purificación a gran escala (apartado 4.1.3.1.2.). El rendimiento de la purificación osciló entre 10 y 30 mg de proteína por litro de cultivo, comparable al obtenido para CA-C silvestre.

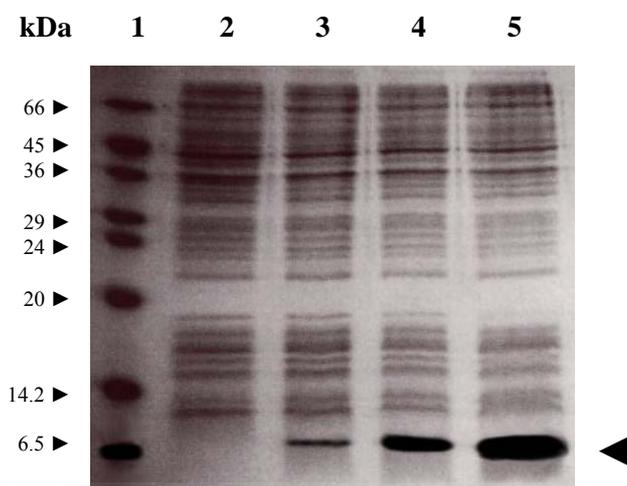


Figura 4.9 SDS-PAGE de extractos de bacterias recombinantes que contienen el ADN plasmídico correspondiente al mutante M185A. Calle 1: marcadores de peso molecular. Calle 2: extracto celular control, antes de la inducción con IPTG. Las calles 3, 4 y 5 se corresponden con el extracto celular tras 1, 2 y 3 horas de inducción con IPTG, respectivamente. La flecha a la derecha indica la banda correspondiente al mutante.

4.2.2.2.2. Digestión con BrCN y purificación de los fragmentos obtenidos.

El mutante M185A sólo presenta dos metioninas en su secuencia (Met214 y Met215), situadas detrás de la α -hélice 3 de CA-C (Fig. 4.1). En consecuencia, la acción del BrCN sobre este mutante, de acuerdo al apartado 4.1.4., da lugar a dos fragmentos de tamaño muy distinto:

(A) El fragmento de interés, que comprende los residuos 146-214 de la secuencia de CA-C, y por lo tanto, presenta la mayor parte de su estructura, con un peso molecular de 7940.0 Da.

(B) El fragmento de menor tamaño, con los residuos 215-231 de CA-C, y un peso molecular de 1681.9 Da.

Para llevar a cabo la purificación del fragmento 146-214 de CA-C, se optó por realizar una cromatografía de exclusión por tamaño molecular. Sin embargo, no se consiguió la purificación del fragmento deseado, probablemente debido a que ambos

fragmentos (el de mayor y el de menor tamaño), después de la digestión con BrCN, permanecían unidos por un puente disulfuro entre los residuos Cys198 y Cys218, impidiendo la óptima separación de los mismos. Por ello, se decidió diseñar otro mutante en el que dicho puente disulfuro no existiera.

4.2.2.3. Obtención y digestión del mutante M185A/C218S.

4.2.2.3.1. Diseño, análisis de la expresión, expresión a gran escala y purificación.

Como previamente se vio que la mutación del residuo Cys198 impedía la expresión de la proteína (apartado 4.2.1.3.), se optó por mutar la Cys218 (Fig. 4.1). Así pues, sobre el mutante M185A se introdujo una nueva mutación: C218S. La sustitución del residuo de cisteína por serina se eligió para evitar, en la medida de lo posible, efectos estéricos; además, en estudios previos se había visto que la mutación C218S no había impedido la expresión de proteína (McDermott et al., 1996; Mateu, 2002).

El mutante M185A/C218S se expresaba en *E. coli* BL21 (DE3) (Fig. 4.10), por lo tanto, se llevó a cabo la expresión a gran escala y purificación (apartado 4.1.3.1.2.), aunque el rendimiento de la purificación (1-10 mg/litro) fue inferior al obtenido para CA-C silvestre (10-30 mg/litro).

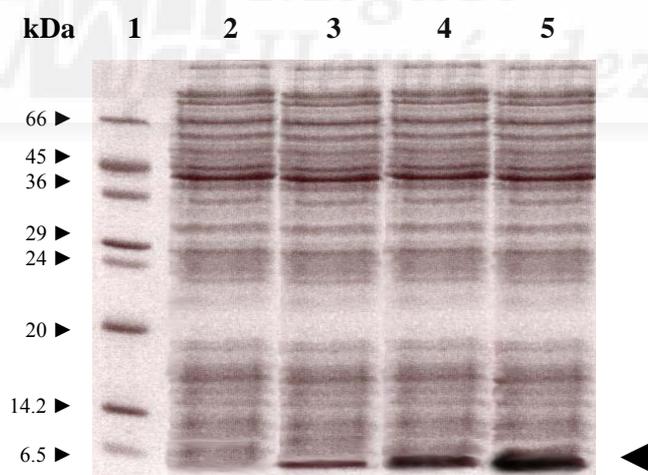


Figura 4.10 SDS-PAGE de extractos de bacterias recombinantes que contienen el ADN plasmídico correspondiente al mutante M185A/C218S. Calle 1: marcadores de peso molecular. Calle 2: extracto celular control, antes de la inducción con IPTG. Las calles 3, 4 y 5 se corresponden con el extracto celular tras 1, 2 y 3 horas de inducción con IPTG, respectivamente. La flecha a la derecha indica la banda correspondiente al mutante.

4.2.2.3.2. Digestión con BrCN y purificación de los fragmentos obtenidos.

La acción del BrCN sobre el mutante M185A/C218S, siguiendo el protocolo del apartado 4.1.4., produce los dos fragmentos descritos en el apartado 4.2.2.2. Sin embargo, en este caso, ambos fragmentos no permanecen unidos tras el corte.

Para purificar el fragmento 146-214 de CA-C, en primer lugar se realizó una cromatografía de exclusión por tamaño molecular. No obstante, tras varios intentos, y en vista de que no se lograron optimizar las condiciones de la purificación, se optó por utilizar otro tipo de cromatografía. Debido a que, tanto la proteína intacta como los dos fragmentos resultantes de la digestión, presentan distintos pIs (Tabla 4.2), se decidió realizar una cromatografía de intercambio aniónico (apartado 4.1.4.).

Tabla 4.2 pIs teóricos de los distintos fragmentos del mutante M185A/C218S.

	Residuos	pI
Proteína intacta	146-231	7.88
Fragmento mayor	146-214	5.28
Fragmento menor	215-231	11

Al aplicar la mezcla de reacción a temperatura ambiente a una columna de intercambio aniónico, previamente equilibrada en tampón a pH 6.0, tanto la proteína que había quedado sin digerir (proteína intacta) como el fragmento 215-231 de CA-C resultante de la digestión, eluyeron en la fracción no retenida; mientras que el fragmento de interés, que comprende los residuos 146-214, quedó retenido en la columna y eluyó posteriormente, entre 130 y 160 mM de NaCl (Fig. 4.11). La muestra fue pura, como indica su espectro de masas (Fig 4.12).

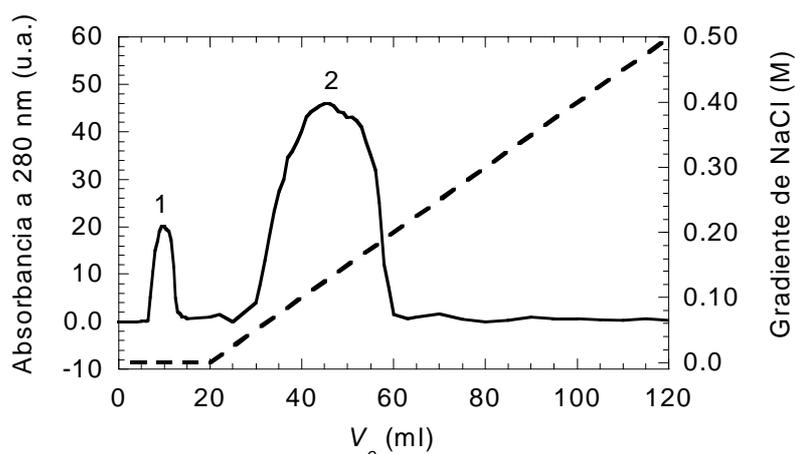


Figura 4.11 Cromatografía de intercambio aniónico. Cromatograma de la elución resultante al aplicar la reacción de corte con BrCN del mutante M185A/C218S (línea continua, eje izquierdo; la escala es arbitraria). El pico 1 se corresponde con la fracción no retenida, que contiene tanto la proteína intacta como el fragmento 215-231 de CA-C. El pico 2 es la fracción retenida en la columna, y contiene el fragmento de interés (residuos 146-214 de CA-C). La línea discontinua representa el gradiente de NaCl (de 0 a 0.5 M, eje derecho).

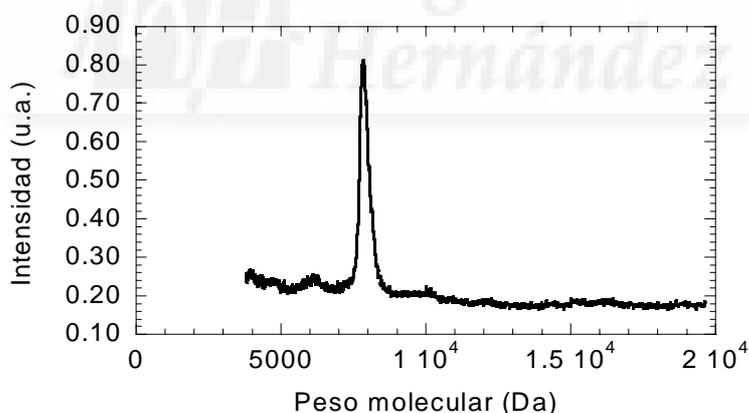


Figura 4.12 Espectro de masas del fragmento 146-214 de CA-C purificado. Se observa un pico único, con un peso molecular de 7940 Da. La escala en el eje de ordenadas es arbitraria.

4.3. DISCUSIÓN.

Como ya se mencionó en la Introducción, la dimerización de CA a través de CA-C es crítica para el ensamblaje de la cápsida del VIH-1. Por lo tanto, el principal objetivo de este trabajo ha sido, en primer lugar, caracterizar termodinámicamente el dominio C-terminal de CA (capítulo 2), para que a partir del conocimiento estructural y

energético disponible del mismo, poder diseñar un inhibidor de la interfase de reconocimiento entre ambos monómeros de CA-C, e impedir de esta manera el ensamblaje de la cápsida, y por tanto la infectividad del VIH.

Como primera aproximación, se llevó a cabo el diseño de un péptido que reproduce esencialmente la α -hélice 2 de CA-C, y se vio que es capaz de reconocer e interactuar específicamente con CA-C, con una afinidad tan sólo cinco veces inferior a la del dominio completo, inhibiendo a la vez su propia dimerización. De esta manera, este péptido resulta útil como cabeza de serie para el diseño racional de compuestos antivirales con la capacidad de inhibir el ensamblaje de la cápsida del VIH-1 (capítulo 3).

Sin embargo, también se ha abordado como aproximación, el diseño de proteínas que presenten menor tamaño que el dominio completo. Las ventajas de estas proteínas minimizadas respecto al dominio intacto son que poseen una mayor posibilidad de aplicación clínica directa (mediante métodos de terapia génica que actualmente están en desarrollo), y además, una mayor simplicidad, por su menor tamaño, para ser utilizadas como modelo en el diseño de inhibidores. Para que todas estas miniproteínas, de forma aislada, tengan la capacidad de asociarse con CA-C formando heterodímeros inactivos, impidiendo de esta manera la dimerización natural de la cápsida, requieren la presencia de la interfase de interacción, formada fundamentalmente por la α -hélice 2 de CA-C. Este estudio también permitirá determinar si un fragmento de CA-C que contenga la α -hélice 2, a pesar de estar desplegado, presenta la capacidad de dimerizar con el dominio completo, y por lo tanto, establecer si plegamiento y heterodimerización van necesariamente acoplados, o son procesos separados y consecutivos.

La estrategia que se ha seguido para la obtención de estas miniproteínas ha consistido en la introducción de: (i) un codón *STOP* para conseguir la expresión directa del fragmento proteico deseado; o, (ii) un residuo de metionina para obtener el fragmento posteriormente mediante ruptura química con BrCN. Las posiciones que se seleccionaron para la introducción de estas mutaciones fueron aquellas situadas: (i) detrás de la α -hélice 2 de CA-C, con el fin de obtener miniproteínas que comprendieran desde el extremo N-terminal de CA-C hasta su α -hélice 2; o, (ii) detrás de la α -hélice 3, para conseguir otras miniproteínas de mayor tamaño que las anteriores (Fig. 4.1).

Ninguna de las mutaciones que se diseñaron (ni mediante la introducción de un codon *STOP*, ni de un residuo de metionina) para la obtención de miniproteínas que sólo contuvieran hasta la segunda hélice α de CA-C ($\Delta 196$, M185A/M214I/M215I/N195M y M185A/M214I/M215I/C198M) permitieron la expresión de proteína (Figs. 4.4 B, 4.5 y 4.7). No obstante, sí se observó sobreexpresión proteica para el mutante M185A/M214I/M215I (Fig. 4.6). Este resultado permite afirmar que las mutaciones adicionales introducidas (N195M o C198M) sobre el triple mutante M185A/M214I/M215I, son las responsables de la ausencia de expresión observada en el cuádruple mutante. Por lo tanto, una posible explicación tentativa es que las mutaciones en la región próxima al extremo C-terminal de la α -hélice 2 de CA-C ($\Delta 196$, N195M o C198M) den lugar a proteínas mutantes que son muy inestables, de modo que no se pliegan correctamente, y por tanto, son degradadas proteolíticamente por la maquinaria celular (Hartl y Hayer-Hartl, 2002).

Por otra parte, los mutantes diseñados para la obtención de miniproteínas que contuvieran hasta la tercera hélice α de CA-C (M185A y M185A/C218S) sí que se expresaron bien (Figs. 4.9 y 4.10). Sin embargo, el mutante $\Delta 214$ no se expresó (Fig. 4.8).

El mutante M185A fue sometido a tratamiento con BrCN, pero en ninguno de los ensayos realizados se logró la purificación del fragmento de interés (residuos 146-214 de CA-C), probablemente debido a que permanecía unido al fragmento de menor tamaño (residuos 215-231) mediante un puente disulfuro entre los residuos Cys198 y Cys218 de CA-C. Para evitar la formación del puente disulfuro, se diseñó otra mutación, dando lugar al doble mutante M185A/C218S. A partir de este mutante, se puso a punto el protocolo de purificación del fragmento 146-214 de CA-C resultante de la digestión con BrCN (Figs. 4.11 y 4.12). Por lo tanto, se abre la posibilidad de continuar el estudio de esta miniproteína, que contiene los residuos 146-214 de CA-C, comprendiendo la mayor parte de su estructura (Fig. 4.1). El siguiente paso consistiría en estudiar la estabilidad y tendencias conformacionales de este fragmento, así como determinar la capacidad de dimerizar consigo mismo, y de formar heterodímeros con CA-C silvestre.

CONCLUSIONES



1. CA-C, tanto en forma monomérica como dimérica, experimenta una transición conformacional a pH ácido, con un valor medio de pK_a de 4.3 ± 0.2 , y otra a pH básico. En la transición ácida, CA-C forma un glóbulo fundido.
2. El desplegamiento térmico del dímero de CA-C sigue un modelo de cuatro estados, ya que el equilibrio de desnaturalización térmica de la proteína se compone de tres transiciones térmicas independientes:
 - La primera transición, con un valor de T_m en torno a 307 K, que sólo se observa por FTIR y RMN, supone la disociación del dímero de CA-C para dar lugar a un intermediario monomérico plegado.
 - La segunda transición, que posee un valor de T_m en torno a 325 K, observada por fluorescencia del ANS, anisotropía y DC en el ultravioleta cercano, ocasiona la reorganización conformacional del intermediario monomérico plegado, dando lugar a una especie monomérica parcialmente plegada, cuyas características son similares a las del glóbulo fundido de CA-C detectado a pH ácido.
 - La tercera transición, con un valor de T_m en torno a 332 K, que se observa por DC en el ultravioleta lejano, absorbancia, FTIR y RMN, da lugar al estado monomérico completamente desnaturalizado.
3. Se han determinado los parámetros termodinámicos de la disociación y desnaturalización de las formas dimérica y monomérica de CA-C a pH fisiológico. Mientras que los valores de ΔH_m y ΔC_p de la desnaturalización de la forma monomérica de CA-C son similares a los encontrados en otras proteínas monoméricas, los valores de ΔH y ΔS de la disociación del dímero de CA-C son pequeños en comparación con los correspondientes para otras proteínas oligoméricas de tamaño similar. Esto se debe a que el producto directo de la disociación de CA-C está completamente plegado, de modo que, la disociación del dímero y la desnaturalización del monómero de CA-C no están acoplados.
4. CAC1, un péptido sintético de 20 aminoácidos, que comprende la secuencia de la hélice principal de dimerización de CA-C (α -hélice 2), se encuentra mayoritariamente desordenado en disolución, y presenta una alta tendencia a autoasociarse, con una constante de autodisociación aparente de $12 \pm 5 \mu\text{M}$.

5. CAC1 es capaz de interactuar con CA-C de forma específica, inhibiendo su propia dimerización. La constante de disociación aparente del heterocomplejo formado entre CAC1 y CA-C es del orden de 50 μM , sólo cinco veces superior que la del dominio completo. La similitud de la constante de disociación de CA-C con la constante de autodisociación aparente de CAC1 ($12 \pm 5 \mu\text{M}$), y su diferencia relativamente pequeña con la constante de disociación aparente del heterocomplejo ($50 \pm 25 \mu\text{M}$), indica que el péptido mimetiza la mayor parte de la interfase de dimerización de CA-C. Por todo ello, este péptido puede ser útil como modelo para el diseño de un agente antiviral capaz de inhibir el ensamblaje de la cápsida.
6. Los aminoácidos situados en la región próxima al extremo C-terminal de la α -hélice 2 de CA-C probablemente desempeñan un papel importante para el correcto plegamiento de la proteína, ya que las mutaciones introducidas en estas posiciones impidieron la expresión.
7. Se ha conseguido la puesta a punto del protocolo de digestión con BrCN y posterior purificación de un fragmento de CA-C que contiene los residuos 146-214 del dominio intacto, para la posterior realización de estudios biofísicos.

BIBLIOGRAFÍA



- Ackers, G.K. (1967). Molecular sieve studies of interacting protein systems. I. Equations for transport of associating systems. *J. Biol. Chem.* **242**, 3026-3034.
- Ackers, G.K. (1970). Analytical gel chromatography of proteins. *Adv. Protein Chem.* **24**, 343-446.
- Alin, K. y Goff, S.P. (1996). Amino acid substitutions in the CA protein of Moloney murine leukemia virus that block early events in infection. *Virology* **222**, 339-351.
- Anderson, D.E., Becktel, W.J. y Dahlquist, F.W. (1990). pH-induced denaturation of proteins: a single salt bridge contributes 3-5 kcal/mol to the free energy of folding of T4 lysozyme. *Biochemistry* **29**, 2403-2408.
- Anfinsen, C. (1973). Principles that govern the folding of protein chains. *Science* **181**, 223-230.
- Arac, D., Murphy, T. y Rizo, J. (2003). Facile detection of protein-protein interactions by one-dimensional NMR spectroscopy. *Biochemistry* **42**, 2774-2780.
- Arthur, L.O., Bess, J.W., Jr, Sowder, R.C.I., Benveniste, R.E., Mann, D.L., Chermann, J.C. y Henderson, L.E. (1992). Cellular proteins bound to immunodeficiency virus: implications for pathogenesis and vaccines. *Science* **258**, 1935-1938.
- Babé, L.M., Pichuantes, S. y Craik, C.S. (1991). Inhibition of HIV protease activity by heterodimer formation. *Biochemistry* **30**, 106-111.
- Backmann, J., Schäfer, G., Wyns, L. y Bönisch, H. (1998). Thermodynamics and kinetics of unfolding of the thermostable trimeric adenylate kinase from the archeon *Sulfolobus acidocaldarius*. *J. Mol. Biol.* **284**, 817-833.
- Barré-Sinoussi, F., Chermann, J.C., Rey, F., Nugeyre, M.T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vézinet-Brun, F., Rouzioux, C., Rozenbaum, W. y Montagnier, L. (1983). Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immunodeficiency syndrome (AIDS). *Science* **220**, 868-871.
- Barth, A. y Zscherp, C. (2002). What vibrations tell us about proteins. *Quart. Rev. of Biophys.* **35**, 369-430.
- Berthet-Colominas, C., Monaco, S., Novelli, A., Sibai, G., Mallet, F. y Cusack, S. (1999). Head-to-tail dimers and interdomain flexibility revealed by the crystal structure of HIV-1 capsid protein (p24) complexed with a monoclonal antibody Fab. *EMBO J.* **18**, 1124-1136.
- Blauer, G. (1974). Probing protein binding sites by circular dichroism spectroscopy. *Struct. Bonding (Berlin)* **18**, 69-129.

- Borsetti, A., Öhagen, A. y Göttlinger, H.G. (1998). The C-terminal half of the human immunodeficiency virus type 1 Gag precursor is sufficient for efficient particle assembly. *J. Virol.* **72**, 9313-9317.
- Branden, C. y Tooze, J. (1991). *Introduction to Protein Structure*. Garland Press, Nueva York.
- Briggs, J.A., Wilk, T., Welker, R., Krausslich, J.G. y Fuller, S.D. (2003). Structural organization of authentic, mature HIV-1 virions and cores. *EMBO J.* **22**, 1707-1715.
- Briggs, J.A., Simon, M.N., Gross, I., Krausslich, H.G., Fuller, S.D., Vogt, V.M. y Johnson, M.C. (2004). The stoichiometry of Gag protein in HIV-1. *Nature Struct. Mol. Biol.* **11**, 672-675.
- Cairns, T.M. y Craven, R.C. (2001). Viral DNA synthesis defects in assembly-competent Rous sarcoma virus CA mutants. *J. Virol.* **75**, 242-250.
- Campos, L.A., Bueno, M., López-Llano, J., Jiménez, M.A. y Sancho, J. (2004). Structure of stable protein folding intermediates by equilibrium ϕ -analysis: the apoflavodoxin thermal intermediate. *J. Mol. Biol.* **344**, 239-255.
- Cantor, C.R. y Schimmel, P.R. (1980). *Biophysical Chemistry*. W.H. Freeman, Nueva York.
- Chen, Y. y Barkley, M.D. (1998). Toward understanding tryptophan fluorescence in proteins. *Biochemistry* **37**, 9976-9982.
- Chiti, F., van Nuland, N.A., Taddei, N., Magherini, F., Stefani, M., Ramponi, G. y Dobson, C.M. (1998 a). Conformational stability of muscle acylphosphatase: the role of temperature, denaturant concentration, and pH. *Biochemistry* **37**, 1447-1455.
- Chiti, F., Taddei, N., van Nuland, N.A., Magherini, F., Stefani, M., Ramponi, G. y Dobson, C.M. (1998 b). Structural characterization of the transition state for folding of muscle acylphosphatase. *J. Mol. Biol.* **283**, 893-903.
- Chun, T.W., Carruth, L., Finzi, D., Shen, X., DiGiuseppe, J.A., Taylor, H., Hermankova, M., Chadwick, K., Margolick, J., Quinn, T.C., Kuo, Y.H., Brookmeyer, R., Zeiger, M.A., Bardich-Crovo, P. y Siliciano, R.F. (1997). Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection. *Nature* **387**, 183-188.
- Clarke, J. y Fersht, A.R. (1993). Engineered disulfide bonds as probes of the folding pathway of barnase: increasing the stability of proteins against the rate of denaturation. *Biochemistry* **32**, 4322-4329.

- Coffin, J.M. (1995). HIV population dynamics *in vivo*: implications for genetic variation, pathogenesis, and therapy. *Science* **267**, 383-386.
- Colucci, W.J., Tilstra, L., Sattler, M.C., Fronczek, F.R. y Barkley, M.D. (1990). Conformational studies of a constrained tryptophan derivative: implications for the fluorescence quenching mechanism. *J. Am. Chem. Soc.* **112**, 9182-9190.
- Cooper, A. y Johnson, C.M. (1994). Isothermal titration calorimetry. *Methods Mol. Biol.* **22**, 137-150.
- Corbett, R.J.T. y Roche, R.S. (1984). Use of high-speed size-exclusion chromatography for the study of protein folding and stability. *Biochemistry* **23**, 1888-1894.
- Craven, R.C., Leure-duPree, A.E., Weldon, R.A., Jr. y Wills, J.W. (1995). Genetic analysis of the major homology region of the Rous sarcoma virus Gag protein. *J. Virol.* **69**, 4213-4227.
- Davis, A.M. y Teague, S.J. (1999). Hydrogen bonding, hydrophobic interactions, and failure of the rigid receptor hypothesis. *Angew. Chem. Int. Ed.* **38**, 736-749.
- De Clercq, E. (1998). The role of non-nucleoside reverse transcriptase inhibitors (NNRTIs) in the therapy of HIV-1 infection. *Antiviral Res.* **38**, 153-179.
- De Clercq, E. (2002). New developments in anti-HIV chemotherapy. *Biochem. Biophys. Acta.* **1587**, 258-275.
- de Jongh, H.H., Goormaghtigh, E. y Ruyschaert, J.M. (1996). The different molar absorptivities of the secondary structure types in the amide I region: an attenuated total reflection infrared study on globular proteins. *Anal. Biochem.* **242**, 95-103.
- del Álamo, M., Neira, J.L. y Mateu, M.G. (2003). Thermodynamic dissection of a low-affinity protein-protein interface involved in human immunodeficiency virus. *J. Biol. Chem.* **278**, 27923-27929.
- del Álamo, M. y Mateu, M.G. (2005). Electrostatic repulsion, compensatory mutations, and long range non-additive effects at the dimerization interface of the HIV capsid protein. *J. Mol. Biol.* **345**, 893-906.
- Dill, D.A. y Shortle, D. (1991). Denatured states of proteins. *Annu. Rev. Biochem.* **60**, 795-825.
- Ding, K., Louis, J.M. y Gronenberg, A.M. (2004). Insights into conformation and dynamics of protein GB1 during folding and unfolding by NMR. *J. Mol. Biol.* **335**, 1299-1307.

- Divita, G., Baillon, J.G., Rittinger, K., Chermann, J.C. y Goody, R.S. (1995). Interface peptides as structure-based human immunodeficiency virus reverse transcriptase inhibitors. *J. Biol. Chem.* **270**, 28642-28646.
- Domingo, E. y Holland, J.J. (1997). RNA virus mutations and fitness for survival. *Annu Rev. Microbiol.* **51**, 151-178.
- Domingo, E., Biebricher, C.K., Holland, J.J. y Eigen, M. (2001 a). Quasispecies and RNA virus evolution: principles and consequences. Landes Bioscience, Austin, Texas.
- Domingo, E., Mas, A., Yuste, E., Pariente, N., Sierra, S., Gutiérrez-Rivas, M. y Menéndez-Arias, L. (2001 b) Virus population dynamics, fitness variations and the control of viral disease: an update. *Prog. Drug Res.* **57**, 77-115.
- Dorfman, T., Bukovsky, A., Öhagen, A., Hoglund, S. y Göttlinger, H.G. (1994). Functional domains of the capsid protein of human immunodeficiency virus. *J. Virol.* **68**, 8180-8187.
- Doyen, N. y LaPresle, C. (1979). Partial non-cleavage by cyanogen bromide of a methionine-cystine bond from human serum albumin and bovine α -lactalbumin. *Biochem. J.* **177**, 251-254.
- Doyle, M.L. (1997). Characterization of binding interactions by isothermal titration calorimetry. *Curr. Op. Biotech.* **8**, 31-35.
- Dunn, B.M. y Chaiken, I.M. (1974). Quantitative affinity chromatography. Determination of binding constants by elution with competitive inhibitors. *Proc. Natl. Acad. Sci. USA* **71**, 2382-2385.
- Eftink, M.R. (1994). The use of fluorescence methods to monitor unfolding transitions in proteins. *Biophys. J.* **66**, 482-501.
- Eftink, M.R. (1997). Fluorescence methods for studying equilibrium macromolecule-ligand interactions. *Methods Enzymol.* **278**, 221-257.
- Ehrlich, L.S., Liu, T., Scarlata, S., Chu, B. y Carter, C.A. (2001). HIV-1 capsid proteins forms spherical (immature-like) and tubular (mature-like) particles *in vitro*: structure switching by pH-induced conformational changes. *Biophys. J.* **81**, 586-594.
- Epstein, C.J., Goldberger, R.F. y Anfinsen, C.B. (1963). The genetic control of tertiary protein structure: studies with model systems. *Cold Spr. Harb. Symp. Quant. Biol.* **28**, 439-449.
- Fasman, G.D. (1990). *Circular Dichroism and the conformational analysis of biomolecules*. Plenum Press, Nueva York.

- Fersht, A.R. (1999). *Structure and Mechanism in Protein Science*. W.H. Freeman, Nueva York.
- Finzi, D., Hermankova, M., Pierson, T., Carruth, L.M., Buck, C., Chaisson, R.E., Quinn, T.C., Chadwick, K., Margolick, J., Brookmeyer, R., Gallant, J., Markowitz, M., Ho, D.D., Richman, D.D. y Siliciano, R.F. (1997). Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy. *Science* **278**, 1295-1300.
- Formosa, T., Barry, J., Alberts, B.M. y Greenblatt, J. (1991). Using protein affinity chromatography to probe structure of protein machines. *Methods Enzymol.* **208**, 24-45.
- Forshey, B.M., von Schwedler, U.K., Sundquist, W.I. y Aiken, C. (2002). Formation of a human immunodeficiency virus type 1 core of optimal stability is crucial for viral replication. *J. Virol.* **70**, 8645-8652.
- Forsyth, W.R. y Robertson, A.D. (2000). Insensitivity of perturbed pKa values in ovomucoid third domain to charge replacement at a neighboring residue. *Biochemistry* **39**, 8067-8072.
- Franke, E.K., Yuan, H.E. y Luban, J. (1994). Specific incorporation of cyclophilin A into HIV-1 virions. *Nature* **372**, 359-362.
- Frankel, A.D. y Young, J.A. (1998). HIV-1: fifteen proteins and an RNA. *Annu. Rev. Biochem.* **67**, 1-25.
- Freed, E.O. (1998). HIV-1 gag proteins: diverse functions in the virus life cycle. *Virology* **251**, 1-15.
- Freire, E. (1995). Differential scanning calorimetry, en *Protein Stability and Folding: Methods in Molecular Biology*, Vol. 40 (Shirley, B. A. Ed.), pp. 191-218, Humana Press Inc., Towota, Nueva Jersey.
- Frenkel, L.M. y Tobin, N.H. (2004). Understanding HIV-1 drug resistance. *Ther. Drug Monit.* **26**, 116-121.
- Freskgard, P.O., Martesson, L.G., Jonasson, P., Jonsson, B.H. y Carlsson, U. (1994). Assignment of the contribution of the tryptophan residues to the circular dichroism spectrum of human carbonic anhydrase II. *Biochemistry* **33**, 14281-14288.
- Gallo, R.C., Salahuddin, S.Z., Popovic, M., Shearer, G.M., Kaplan, M., Haynes, B.F., Palker, T.J., Redfield, R., Oleske, J., Safai, B., White, G., Foster, P. y Markham, P.D. (1984). Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science* **224**, 500-503.

- Gamble, T.R., Vajdos, F.F., Yoo, S., Worthylake, D.K., Houseweart, M., Sundquist, W.I. y Hill, C.P. (1996). Crystal structure of human cyclophilin A bound to the amino terminal domain of HIV-1 capsid. *Cell* **87**, 1285-1294.
- Gamble, T.R., Yoo, S., Vajdos, F.F., von Schwedler, U.K., Worthylake, D.K., Wang, H., McCutcheon, J.P., Sundquist, W.I. y Hill, C.P. (1997). Structure of the carboxyl-terminal dimerization domain of the HIV-1 capsid protein. *Science* **278**, 849-853.
- Ganser, B.K., Li, S., Kliskho, V.Y., Finch, J.T. y Sundquist, W.I. (1999). Assembly and analysis of conical models for the HIV-1 core. *Science* **283**, 80-83.
- Gelderblom, H.R., Marx, P.A., Ozel, M., Gheysen, D., Munn, R.J., Joy, K.I. y Pauli, G. (1990). *Morphogenesis, maturation and fine structure of lentiviruses*. Stockton Press, Nueva York.
- Gelderblom, H.R. (1991). Assembly and morphology of HIV: potential effect of structure on viral function. *AIDS* **5**, 617-637.
- Gitti, R.K., Lee, B.M., Walker, J., Summers, M.F., Yoo, S. y Sundquist, W.I. (1996). Structure of the amino-terminal core domain of the HIV-1 capsid protein. *Science* **273**, 231-235.
- Goldberg, M.E., Semisotnov, G.V., Friguier, B., Kuwajima, K., Ptitsyn, O.B. y Sugai, S. (1990). An early immunoreactive folding intermediate of the tryptophan synthase β 2 subunit is a molten globule. *FEBS. Lett.* **263**, 51-56.
- Göttlinger, H.G. (2001 a). The HIV-1 assembly machine. *AIDS* **15**, 13-20.
- Göttlinger, H.G. (2001 b). HIV-Gag: a molecular machine driving viral particle assembly and release, en *HIV Sequence Compendium* (2001), pp. 2-28. Theoretical Biology and Biophysics Group, Los Alamos National Laboratori.
- Greenfield, N.J. (1975). Circular dichroism studies of dihydrofolate reductase from a methotrexate-resistant strain of Escherichia coli B, MB 1428: ternary complexes. *Biochem. Biophys. Acta.* **403**, 32-46.
- Gross, E. (1967). The cyanogen bromide reaction. *Methods Enzymol.* **11**, 238-255.
- Gross, I., Hohenberg, H. y Kräusslich, H.G. (1997). *In vitro* assembly properties of purified bacterially expressed capsid proteins of human immunodeficiency virus. *Eur. J. Biochem.* **249**, 592-600.
- Gross, I., Hohenberg, H., Huckangel, C. y Kräusslich, H.G. (1998). N-terminal extension of human immunodeficiency virus capsid protein converts the *in vitro* assembly phenotype from tubular to spherical particles. *J. Virol.* **72**, 4798-4810.

- Gross, I., Hohenberg, H., Wilk, T., Wiegers, K., Grättinger, M., Müller, B., Fuller, S. y Kräulich, H.G. (2000). A conformational switch controlling HIV-1 morphogenesis. *EMBO J.* **19**, 103-113.
- Gulick, R.M., Mellors, J.W., Havlir, D., Eron, J.J., Gonzalez, C., McMahon, D., Richman, D.D., Valentine, F.T., Jonas, L., Meibohm, A., Emini, E.A. y Chodakewitz, J.A. (1997). Treatment with indinavir, zidovudine, and lamivudine in adults with human immunodeficiency virus infection and prior antiretroviral therapy. *New Eng. J. Med.* **337**, 734-739.
- Hartl, F.U. y Hayer-Hartl, M. (2002). Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science* **295**, 1852-1858.
- Henderson, L.E., Bowers, M.A., Sowder, R.C., Serabyn, S.A., Johnson, D.G., Bess, J.W., Jr., Arthur, L.O., Bryant, D.K. y Fenselau, C. (1992). Gag proteins of the highly replicative MN strain of human immunodeficiency virus type 1: posttranslational modifications, proteolytic processings, and complete amino acid sequences. *J. Virol.* **66**, 1856-1865.
- Hinkle, A., Goranson, A., Butters, C.A. y Tobacman, L.S. (1999). Roles for the troponin tail domain in thin filament assembly and regulation. A deletional study of cardiac troponin T. *J. Biol. Chem.* **274**, 7157-7164.
- Ho, D.D., Neumann, A.U., Perelson, A.A., Chen, W., Leonard, J.M. y Markowitz, M. (1995). Rapid turn-over of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature* **373**, 123-126.
- Irún, M.P., García-Mira, M.M., Sánchez-Ruiz, J.M. y Sancho, J. (2001). Native hydrogen bonds in a molten globule: the apoflavodoxin thermal intermediate. *J. Mol. Biol.* **306**, 877-888.
- Jaenicke, R. y Lillie, H. (2000). Folding and association of oligomeric and multimeric proteins. *Adv. Protein Chem.* **53**, 329-401.
- Jameson, D.M. y Seifried, S.E. (1999). Quantification of protein-protein interactions using fluorescence polarization. *Methods* **19**, 222-233.
- Johnson, J.E. (1996). Functional implications of protein-protein interaction in icosahedral viruses. *Proc. Natl. Acad. Sci. USA* **93**, 27-33.
- Jones, S. y Thornton, J.M. (1996). Principles of protein-protein interactions. *Proc. Natl. Acad. Sci. USA* **93**, 13-20.

- Joppich-Kuhn, R., Corkill, J.A. y Giese, R.W. (1982). Oxidation of a methionine to methionine sulfoxide as a side reaction of cyanogen bromide cleavage. *Anal. Biochem.* **119**, 73-77.
- Jowett, J., Hockley, D., Nermut, M.V. y Jones, I.M. (1992). Distinct signals in human immunodeficiency virus type 1 Pr55 necessary for RNA binding and particle formation. *J. Gen. Virol.* **73**, 3079-3086.
- Kelly, S.M. y Price, N.C. (2000). The use of circular dichroism in the investigation of protein structure and function. *Cur. Prot. and Peptide Sci.* **1**, 349-384.
- Kemming, J. y Creighton, T.E. (1996). The physical properties of local interactions of tyrosine residues in peptides and unfolded proteins. *J. Mol. Biol.* **245**, 251-260.
- Khorasanizadeh, S., Campos-Olivas, R. y Summers, M.F. (1999). Solution structure of the capsid protein from the human T-cell leukemia virus type 1. *J. Mol. Biol.* **291**, 491-505.
- Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.
- Lakowicz, J.R. (1999). *Principles of fluorescence spectroscopy*, 2nd Ed. Plenum Press, Nueva York.
- Lanman, J., Sexton, J., Sakalian, M. y Prevelige, P.E., Jr. (2002). Kinetic analysis of the role of intersubunit interactions in human immunodeficiency virus type 1 capsid protein assembly *in vitro*. *J. Virol.* **76**, 6900-6908.
- Levy, J.A., Hoffmann, A.D., Kramer, S.M., Landis, J.A., Shimabukuro, J.M. y Oshiro, L.S. (1984). Recovery of AIDS-associated retroviruses from patients with AIDS or AIDS-related conditions and from clinically healthy individuals. *J. Infect. Dis.* **152**, 734-738.
- Li, S., Hill, C.P., Sundquist, W.I. y Finch, J.T. (2000). Image reconstructions of helical assemblies of the HIV-1 CA protein. *Nature* **407**, 409-413.
- Liao, T.H., Salnikow, J., Moore, S. y Stein, W.H. (1973). Bovine pancreatic deoxyribonuclease A. Isolation of cyanogen bromide peptides; complete covalent structure of the polypeptide chain. *J. Biol. Chem.* **248**, 1489-1495.
- Luo, J., Iwakura, M. y Matthews, C.R. (1995). Detection of an intermediate in the thermal unfolding of a cysteine-free form of dihydrofolate from *Escherichia coli*. *Biochemistry* **34**, 10669-10675.

- Luo, Y., Kay, M. y Baldwin, R.L. (1997). Cooperativity of folding of apomyoglobin pH 4 intermediate studied by glycine and proline mutations. *Nat. Struct. Biol.* **4**, 925-930.
- Maldonado, S., Jiménez, M.A., Langdon, G.M. y Sancho, J. (1998). Cooperative stabilization of a molten-globule apoflavodoxin fragment. *Biochemistry* **37**, 10589-10596.
- Mammano, F., Ohagen, A., Hoglund, S. y Gottlinger, H. G. (1994). Role of the major homology region of human immunodeficiency virus type 1 in virion morphogenesis. *J. Virol.* **68**, 4927-4936.
- 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**, 5282-5290.
- Mansky, M. (2002). HIV mutagenesis and the evolution of antiretroviral drug resistance. *Drug Resis Updat* **5**, 219-223.
- Marky, L.A. y Breslauer, K.J. (1987). Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves. *Biopolymers* **26**, 1601-1620.
- Maroun, R.G., Gayet, S., Benleulmi, M.S., Porumb, H., Zaragarian, L., Merad, H., Leh, H., Mouscadet, J.F., Troalen, F. y Fermadjina, S. (2001). Peptide inhibitors of HIV-1 integrase dissociate the enzyme oligomers. *Biochemistry* **40**, 13840-13848.
- Match, H., Volkin, D.B., Burke, C.J. y Middaugh, C.R. (1995). Ultraviolet absorption spectroscopy, en *Protein Folding and Stability: Theory and Practice* (Shirley, B.A., Ed.), pp. 91-114, Humana Press Inc., Totowa, Nueva Jersey.
- Mateu, M.G. (2002). Conformational stability of dimeric and monomeric forms of the C-terminal domain of human immunodeficiency virus. *J. Mol. Biol.* **318**, 519-531.
- McDermott, J., Farrell, L., Ross, R. y Barklis, E. (1996). Structural analysis of human immunodeficiency virus type 1 Gag protein interactions, using cysteine-specific reagents. *J. Virol.* **70**, 5106-5114.
- Mendelsohn, R. y Mantsch, H.H. (1986). Fourier transform infrared studies of lipid-protein interaction, en *Progress in Lipid-Protein Interactions* (Watts, A. y De Pont, J. J. H. H. M., Eds), Elsevier, Amsterdam.

- Mervis, R.J., Ahmad, N., Lillehoj, E.P., Raum, M.G., Salazar, F.H., Chan, H.W. y Venkatesan, S. (1988). The gag gene products of human immunodeficiency virus type 1: alignment within the gag open reading frame, identification of posttranslational modifications, and evidence for alternative gag precursors. *J. Virol.* **62**, 3993-4002.
- Misselwitz, R., Hausdorf, G., Welfe, K., Höhne, W.E. y Welfe, H. (1995). Conformation and stability of recombinant HIV-1 capsid protein p24 (rp24). *Biochem. Biophys. Acta.* **1250**, 9-18.
- Momany, C., Kovari, L.C., Prongay, A.J., Keller, W., Gitti, R.K., Lee, B.M., Gorbalenya, A.E., Tong, L., McClure, J., Ehrlich, L.S., Summers, M.F., Carter, C. y Rossmann, M.G. (1996). Crystal structure of dimeric HIV-1 capsid protein. *Nature Struct. Biol.* **3**, 763-770.
- Muñoz, V. y Serrano, L. (1994). Elucidating the folding problem of helical peptides using empirical parameters. *Nature Struct. Biol.* **1**, 399-409.
- 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 Sci.* **4**, 2138-2148.
- Naesens, L., Snoeck, R., Andrei, G., Balzarini, J., Neyts, J. y De Clercq, E. (1997). HPMPC (cidofovir), PMEAs (adefovir) and related acyclic nucleoside phosphonate analogues: A review of their pharmacology and clinical potential in the treatment of viral infections. *Antiviral Chem. Chemother.* **8**, 1-23.
- Neet, D.E. y Timm, D.E. (1994). Conformational stability of dimeric proteins: quantitative studies by equilibrium denaturation. *Protein Sci.* **3**, 2167-2174.
- Neira, J.L. y Gómez, F.J. (2004). The conformational stability of the *Streptomyces coelicolor* histidine-phosphocarrier protein: characterization of cold-denaturation and urea-protein interactions. *Eur. J. Biochem.* **271**, 2165-2181.
- Pace, C.N. (1986). Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol.* **131**, 266-280.
- Pace, C.N. y Laurents, D.V. (1989). A new method for determining the heat capacity change for protein folding. *Biochemistry* **28**, 2520-2525.
- Pace, C.N. y Scholtz, J.M. (1997). Measuring the conformational stability of a protein, en *Protein Structure* 2nd Ed. (Creighton, T. E., Ed.), pp. 253-259, Oxford University Press, Oxford.

- Padmanabhan, S., Jimenez, M.A., Gonzalez, C., Sanz, J.M., Gimenez-Gallego, G. y Rico, M. (1997). Three-dimensional solution structure and stability of phage 434 Cro protein. *Biochemistry* **36**, 6424-6436.
- Pantoliano, M.W., Petrella, E.C., Kwanoski, J.D., Lobanov, V.S., Myslik, J., Graf, E., Carver, T., Asel, E., Springer, A.R., Lane, P. y Salemme, F.R. (2001). High-density miniaturized thermal shift assays as a general strategy for drug discovery. *J. Biomol. Screening* **6**, 429-440.
- Perelson, A.A., Neumann, A.U., Markowitz, M., Leonard, J.M. y Ho, D.D. (1996). HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science* **271**, 1582-1586.
- Pfeil, W. y Privalov, P.L. (1976). Thermodynamic investigations of proteins. II. Calorimetric study of lysozyme denaturation by guanidine hydrochloride. *Biophys. Chem.* **4**, 33-40.
- Pierce, M.M., Raman, C.S. y Nall, B.T. (1999). Isothermal titration calorimetry of protein-protein interactions. *Methods* **19**, 213-221.
- Pillay, D. (2000). Combination antiviral therapy: but which combination, and why?. *Curr. Opin. Infect. Dis.* **13**, 621-623.
- Piotto, M., Saudek, V. y Sklenar, V. (1992). Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR.* **2**, 661-665.
- Privalov, P.L., Tiktopulo, E.I., Venyaminov, S.Yu., Griko, Yu.V., Makhataдзе, G.I. y Khechinashvili, N.N. (1989). Heat capacity and conformation of proteins in the denatured state. *J. Mol. Biol.* **205**, 737-750.
- Privalov, P.L. (1992). Physical basis of the stability of the folded conformations of proteins, en *Protein Folding* (Creighton, T.E., Ed.), pp. 83-126, Ed. W.H. Freeman, Nueva York.
- Ptitsyn, O.B. (1995). Molten globule and protein folding. *Adv. Protein Chem.* **47**, 83-229.
- Reicin, A.S., Paik, S., Berkowitz, R.D., Luban, J., Lowy, I. y Goff, S.P. (1995). Linker insertion mutations in the human immunodeficiency virus type 1 gag gene: effects on virion particle assembly, release, and infectivity. *J. Virol.* **69**, 642-650.
- Richman, D.D. (2001). HIV chemotherapy. *Nature* **410**, 995-1001.
- Robertson, A.D. y Murphy, K.P. (1997). Protein structure and the energetics of protein stability. *Chem. Rev.* **97**, 1251-1267.

- Rosé, S., Hensley, P., O`Shannessy, D.J., Culp, J., Debouck, C. y Chaiken, I. (1992). Characterization of HIV-1 p24 self-association using analytical chromatography. *Proteins* **13**, 112-119.
- Royer, C.A. (1995). Fluorescence spectroscopy, en *Protein Stability and Folding*, Methods in Molecular Biology, Vol. 40 (Shirley, B.A., Ed.), pp. 65-89, Humana Press Inc., Towota, Nueva Jersey.
- Rubinson, J.F. y Rubinson, K.A. (1998). *Contemporary Chemical Analysis*, 5th Ed. Prentice Hall, Nueva Jersey.
- Schmid, F.X. (1997). Optical spectroscopy to characterize protein conformation and conformational changes, en *Protein Structure*, 2nd Ed. (Creighton, T. E., Ed.), pp. 261-297, Oxford University Press, Oxford.
- Schramm, H.J., Boetzel, J., Buttner, J., Fritsche, E., Gohring, W., Jaeger, E., Konig, S., Thumfart, O., Wenger, T., Nagel, N.E. y Schramm, W. (1996). The inhibition of human immunodeficiency virus proteases by interface peptides. *Antiviral Res.* **30**, 155-170.
- Schreiber, G. y Fersht, A.R. (1995). Energetics of protein-protein interactions: analysis of the barnase-barstar interface by single mutations and double mutant cycles. *J. Mol. Biol.* **248**, 478-486.
- Semisotnov, G.V., Rodionova, N.A., Razgulyaev, O.I., Uversky, V.N., Gripas, A.F. y Gimanshin, R.I. (1991). Study of the "molten globule" intermediate state in protein folding by a hydrophobic fluorescent probe. *Biopolymers* **31**, 119-128.
- Serrano, L., Horowitz, A., Avron, B., Bycroft, M. y Fersht, A.R. (1990). Estimating the contribution of engineered surface electrostatic interactions to protein stability by using double-mutant cycles. *Biochemistry* **29**, 9343-9352.
- Simmerling, C.L y Elber, R. (1995). Computer determination of peptide conformations in water: different roads to structure. *Proc. Natl. Acad. Sci. USA* **92**, 3190-3193.
- Smith, T.J., Kremer, M.J., Luo, M., Vriend, G., Arnold, E., Kamer, G., Rossmann, M.G., McKinlay, M.A., Diana, G.D. y Otto, M.J. (1986). The site of attachment in human rhinovirus 14 for antiviral agents that inhibit uncoating. *Science* **233**, 1286-1293.
- Stanfield, R.L. y Wilson, I.A. (1995). Protein-peptide interactions. *Cur. Opin. Struct. Biol.* **5**, 103-113.

- Steif, C., Weber, P., Hinz, H.J., Flossdorf, J., Cesareni, G. y Kokkinidis, M. (1993). Subunit interactions provide a significant contribution to the stability of the dimeric four- α -helical bundle protein ROP. *Biochemistry* **32**, 3867-3876.
- Stennicke, H.R., Olesen, K., Sorensen, S.B. y Breddam, K. (1997). C-terminal incorporation of fluorogenic and affinity labels using wild-type and mutagenized carboxypeptidase. *Anal. Biochem.* **248**, 141-148.
- Stryer, L. (1965). The interaction of a naphthalene dye with apomyoglobin and apohemoglobin. A fluorescent probe of non-polar binding sites. *J. Mol. Biol.* **13**, 482-495.
- Surewicz, W.K., Mantsch, H.H. y Chapman, D. (1993). Determination of protein secondary structure by Fourier transform infrared spectroscopy: a critical assessment. *Biochemistry* **32**, 389-394.
- Tang, K.S., Guralnick, B.J., Wang, W.K. Fersht, A.R. e Itzhaki, L.S. (1999). Stability and folding of the tumor suppressor protein p16. *J. Mol. Biol.* **285**, 1869-1886.
- Tang, S., Murakami, T., Agresta, B.E., Campbell, S., Freed, E.O. y Levin, J.G. (2001). Human immunodeficiency virus type 1 N-terminal capsid mutants that exhibit aberrant core morphology are blocked in initiation of reverse transcription in infected cells. *J. Virol.* **75**, 9357-9366.
- Tatischeff, I. y Klein, R. (1975). Influence of the environment on the excitation wavelength dependence of the fluorescence quantum yield of indole. *Photochem. Photobiol.* **22**, 221-229.
- Tavel, J.A., Miller, K.D. y Masus, H. (1999). Guide to major clinical trials of antiretroviral therapy in human immunodeficiency virus-infected patients: protease inhibitors, non-nucleoside reverse transcriptase inhibitors, and nucleotide reverse transcriptase inhibitors. *Clin. Infect. Dis.* **28**, 643-676.
- Taylor, H.C. y Chaiken, I.M. (1977). Active site ligand binding by an inactive ribonuclease S analogue. A quantitative affinity chromatographic study. *J. Biol. Chem.* **252**, 6991-6994.
- Turner, D.C. y Brand, L. (1968). Quantitative estimation of protein binding site polarity. Fluorescence of N-arylamino-naphthalenesulfonates. *Biochemistry* **7**, 3381-3390.
- Turner, B.G. y Summers, M.F. (1999). Structural biology of HIV. *J. Mol. Biol.* **285**, 1-32.
- Valdés, R. y Ackers, G.K. (1979). Study of protein subunit association equilibria by elution gel chromatography. *Methods Enzymol.* **61**, 125-142.

- Vogt, V.M. (1997). Retroviral virus and genomes, en *Retroviruses* (Coffin, J.M., Hughes, J.M. y Varnus, H.E., Eds.), pp. 27-69, Cold Spring Harbor Laboratory Press, Nueva York.
- von Schwedler, U.K., Stemmler, T.L., Klishko, V.Y., Li, S., Albertine, K.H., Davis, D.R. y Sundquist, W.I. (1998). Proteolytic refolding of the HIV-1 capsid protein amino-terminus facilitates viral core assembly. *EMBO J.* **17**, 1555-1568.
- von Schwedler, U.K., Stray, K.M., Garrus, J.E. y Sundquist, W.I. (2003). Functional surfaces of the human immunodeficiency virus type 1 capsid protein. *J. Virol.* **77**, 5439-5450.
- Waldron, T.T. y Murphy, K.P. (2003). Stabilization of proteins by ligand binding: application to drug screening and determination of unfolding energetics. *Biochemistry* **42**, 5058-5064.
- Wang, C.T. y Barklis, E. (1993). Assembly, processing and infectivity of human immunodeficiency virus type 1 Gag mutants. *J. Virol.* **67**, 4264-4273.
- Wei, P., Garber, M.E., Fang, S.M., Fischer, W.H. y Jones, K.A. (1998). A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA. *Cell* **92**, 451-462.
- Wells, J.A. (1991). Systematic mutational analyses of protein-protein interfaces. *Methods Enzymol.* **202**, 390-411.
- Wilk, T. y Fuller, S.D. (1999). Towards the structure of the human immunodeficiency virus: divide and conquer. *Curr. Opin. Struct. Biol.* **9**, 213-243.
- Willems, L., Kerkhofs, P., Attenelle, L., Burny, A., Portetelle, D. y Kettmann, R. (1997). The major homology region of bovine leukaemia virus p24gag is required for virus infectivity in vivo. *J. Gen. Virol.* **78**, 637-640.
- Wong, J.K., Hezareh, M., Gunthard, H.F., Havlir, D.V., Ignacio, C.C., Spina, C.A. y Richman, D.D. (1997). Recovery of replication-competent HIV despite prolonged suppression of plasma viremia. *Science* **278**, 1291-1295.
- Woody, R.W. (1995). Circular dichroism. *Methods Enzymol.* **246**, 34-71.
- Worthylake, D.K., Wang, H., Yoo, S., Sundquist, W.I. y Hill, C.P. (1999). Structures of the HIV-1 capsid protein dimerization domain at 2.6 Å resolution. *Acta Crystallog. Sect. D* **55**, 85-92.
- Wüthrich, K. (1986). *NMR of proteins and nucleic acids*. John Wiley and Sons, Nueva York.

- Xu, D., Tsai, C.J. y Nussinov, R. (1997). Hydrogen bonds and salt bridges across protein-protein interfaces. *Protein Eng.* **10**, 999-1012.
- Zhang, A.Y., Poorman, R.A., Maggiora, L.L., Heinrikson, R.L. y Kezdy, F.J. (1991). Dissociative inhibition of dimeric enzymes. Kinetic characterization of the inhibition of HIV-1 protease by its COOH-terminal tetrapeptide. *J. Biol. Chem.* **266**, 15591-15594.
- Zhang, Y. y Rollins, B.J. (1995). A dominant negative inhibitor indicates that monocyte chemoattractant protein 1 functions as a dimer. *Mol. Cell. Biol.* **15**, 4851-4855.
- Zweifel, M.E. y Barrick, D. (2002). Relationships between the temperature dependence of solvent denaturation and the denaturant dependence of protein stability curves. *Biophys. Chem.* **101-102**, 221-237.



APÉNDICE



The dimerization domain of the HIV-1 capsid protein binds a capsid protein-derived peptide: A biophysical characterization

MARÍA T. GARZÓN,^{1,5} MARÍA C. LIDÓN-MOYA,¹ FRANCISCO N. BARRERA,¹ ALICIA PRIETO,² JAVIER GÓMEZ,¹ MAURICIO G. MATEU,³ AND JOSÉ L. NEIRA^{1,4,5}

¹Instituto de Biología Molecular y Celular, Universidad Miguel Hernández, 03202 Elche (Alicante), Spain

²Centro de Investigaciones Biológicas, CSIC, 28006 Madrid, Spain

³Centro de Biología Molecular "Severo Ochoa" (CSIC-UAM), Universidad Autónoma de Madrid, Cantoblanco, 28049 Madrid, Spain

⁴Instituto de Biofísica y Física de los sistemas complejos, 50009 Zaragoza, Spain

(RECEIVED December 9, 2003; FINAL REVISION February 17, 2004; ACCEPTED February 26, 2004)

Abstract

The type 1 HIV presents a conical capsid formed by ~1500 units of the capsid protein, CA. Homodimerization of CA via its C-terminal domain, CA-C, constitutes a key step in virion assembly. CA-C dimerization is largely mediated by reciprocal interactions between residues of its second α -helix. Here, we show that an N-terminal-acetylated and C-terminal-amidated peptide, CAC1, comprising the sequence of the CA-C dimerization helix plus three flanking residues at each side, is able to form a complex with the entire CA-C domain. Thermal denaturation measurements followed by circular dichroism (CD), NMR, and size-exclusion chromatography provided evidence of the interaction between CAC1 and CA-C. The apparent dissociation constant of the heterocomplex formed by CA-C and CAC1 was determined by several biophysical techniques, namely, fluorescence (using an anthraniloyl-labeled peptide), affinity chromatography, and isothermal titration calorimetry. The three techniques yielded similar values for the apparent dissociation constant, in the order of 50 μ M. This apparent dissociation constant was only five times higher than was the dissociation constant of both CA-C and the intact capsid protein homodimers (10 μ M).

Keywords: peptide recognition; binding; calorimetry; fluorescence

During the assembly stage of HIV, thousands of copies of the viral Gag protein associate at the cell membrane and bud to form the immature noninfectious virion. Gag is cleaved

by the viral PR; this process liberates, among other proteins, the capsid protein, CA, which forms the internal protein shell of the mature virus. The mature conical capsid of the type 1 HIV (HIV-1) is composed, among other proteins, by ~1500 copies of CA. This electron-dense cone-shaped structure encapsidates the viral RNA, the nucleocapsid proteins, and other key enzymes. In vitro, CA spontaneously assembles into helical tubes and cones resembling the viral capsids (Gross et al. 1997, 1998; Ganser et al. 1999; Li et al. 2000). Several studies have shown that proper capsid assembly is critical for viral infectivity (von Schwedler et al. 1998; Tang et al. 2001; Forshey et al. 2002). Inhibition of CA assembly by mutations has a lethal effect (Gross et al. 1997; von Schwedler et al. 1998; Tang et al. 2001), and the alteration of capsid stability reduces virus replication (Forshey et al. 2002). These results support a potential antiviral

⁵These two authors contributed equally to this work.

Reprint requests to: José L. Neira, Instituto de Biología Molecular y Celular, Edificio Torregaitán, Universidad Miguel Hernández, Avda. del Ferrocarril s/n, 03202, Elche (Alicante), Spain; e-mail: jlneira@umh.es; fax: 34-966658758.

Abbreviations: CA, capsid protein of HIV (p24); CA-C, C-terminal domain of CA; CAC1, peptide comprising the second helix of CA-C; CAC1Aib, the CAC1 peptide with the anthraniloyl moiety attached to the ϵ -group of Lys182; CD, circular dichroism; IN, integrase; ITC, isothermal titration calorimetry; RT, reverse transcriptase; PR, protease; SEC, size-exclusion chromatography; TSP, 3-(trimethylsilyl) propionic acid-2,2,3,3-²H₄-sodium salt; UV, ultraviolet.

Article and publication are at <http://www.proteinscience.org/cgi/doi/10.1110/ps.03555304>.

strategy for HIV-1, based on the use of compounds able to interfere with the CA–CA recognition and then capsid assembly. Those inhibitors could be either small organic compounds or large “interfacial” inhibitors that mimic most of the interacting interface (Babé et al. 1991; Zhang et al. 1991; Garg et al. 1999). The design of either type of inhibitor requires not only a detailed structural knowledge of CA but also a thermodynamic characterization of the interactions involved in the CA–CA interfaces. The designs of inhibitors based on the oligomerization properties of several HIV proteins have been attempted in the case of integrase (IN) by using peptides (Maroun et al. 2001), and with the synthesis of small organic molecules in the case of the N-terminal domain of CA (Tang et al. 2003). In addition to the benefits provided by the design of new drugs against human immunodeficiency virus, these new molecules might shed light into the field of molecular recognition. From a theoretical and also an experimental point of view, understanding how molecular recognition occurs presents a formidable and interesting challenge (Muller-Dethlefs and Hozba 2000). The complexity of interactions between proteins and flexible target molecules is determined, in most of the cases, by the flexibility at the binding sites and the structural rearrangements occurring upon binding in both molecules (Davis and Teague 1999; Verkhiver et al. 2002), but it is not known yet to which extent in each of the molecules involved.

The CA protein of HIV-1 is formed by two independent domains separated by a flexible linker. CA is capable of different modes of association (Rosé et al. 1992; Dorfman et al. 1994; Misselwitz et al. 1995), and in fact, there is a wide body of evidence that dimerization through the C-terminal domain, CA-C, is an important driving force in virus assembly (Wang and Barklis 1993; Borsetti et al. 1998). High-resolution structures of both CA domains have been obtained (Gamble et al. 1996, 1997; Gitti et al. 1996). CA-C is a dimer, both in solution and in the crystal form (Gamble et al. 1997); furthermore, the dimerization constant of CA-C is nearly the same ($10 \pm 3 \mu\text{M}$) of that of the intact CA protein ($18 \pm 1 \mu\text{M}$; Gamble et al. 1997). Each monomer (residues 147–231 of the intact CA protein) is composed of a 3_{10} helix followed by an extended strand and four α -helices connected by short loops (Gamble et al. 1997). The dimerization interface is largely formed by the mutual docking of α -helix 2 from each monomer (residues Ser178–Val191; Fig. 1). The folding and association pathway of CA-C involves the presence of a monomeric intermediate that rearranges and dimerizes to yield the native dimer (Mateu 2002). The energetic role of each side chain at the dimerization interface has also been determined: the side chain of Trp184 (Gamble et al. 1997; del Álamo et al. 2003) and those of Ile150, Leu151, Arg154, Leu172, Glu175, Val181, Met185, and Leu189 are key for CA-C dimerization (del Álamo et al. 2003). The CA-C also has a region of 20 amino

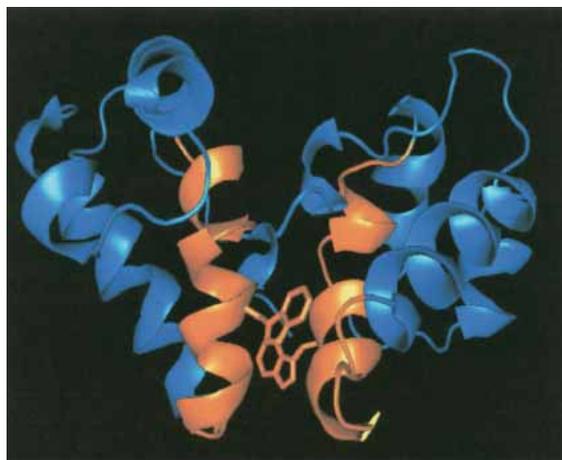


Figure 1. Structure of CA-C dimer. X-ray structure of CA-C, showing the dimeric structure of the domain. The monomers are depicted in the same color, and α -helix 2 plus the flanking residues used in the design of CAC1 in each monomer are highlighted in yellow. The side chain of Trp184 is indicated. The figure was produced with PyMOL (DeLano 2002) and from the Protein Data Bank file for CA-C (code accession number 1A43).

acids called the MHR, which is highly conserved among retroviruses. This region is hydrogen-bonded to the α -helix 2 (Gamble et al. 1997), and the MHR residues are involved in the association process, affecting the stability of the CA-C monomer (del Álamo et al. 2003).

The structural and energetic knowledge acquired on the CA-C dimerization interface can be used in the rational design of an interfacial inhibitor of CA-C dimerization, based on peptides. The successful design of such inhibitor could provide hints into the molecular basis governing CA-C–CA-C interactions, and further, it could be used as a lead compound for the development of new antiviral agents against HIV. Here, we report that a 20-mer synthetic peptide derived from the CA-C sequence recognizes CA-C, and it is able to inhibit the homodimerization of CA-C. This peptide, CAC1 (residues 175–194 of CA), comprises the amino acid sequence of the α -helix 2 in CA-C. CAC1 is mainly disordered in solution, self-associates, and interacts specifically with CA-C with an affinity only five times lower than the complete domain.

Results

Spectroscopic characterization of the CAC1 peptide and its variant, CAC1Aib

As a first step to study the binding properties between CA-C and CAC1, several biophysical techniques were used to address the conformational propensities of CAC1, a 20-mer peptide comprising the α -helix 2 of the CA-C domain of HIV-1 plus three flanking residues at each site. The se-

quence of CAC1 is Ac-E¹⁷⁵QASQEVKNWMTETLLVQNA¹⁹⁴-NH₂.

Circular dichroism measurements

The CD spectrum of CAC1 at 10 μM showed two intense minima at 220 and 195 nm, and a maximum at 205 nm (Fig. 2A). This spectrum is not characteristic either of a disordered structure or of an α-helix (Kelly and Price 2000). The intense minima and/or the maximum could be due to the presence of the sole tryptophan, as it has been shown in other proteins (Freskgård et al. 1994). The intensity of the three bands, and the shape of the CD spectrum changed as

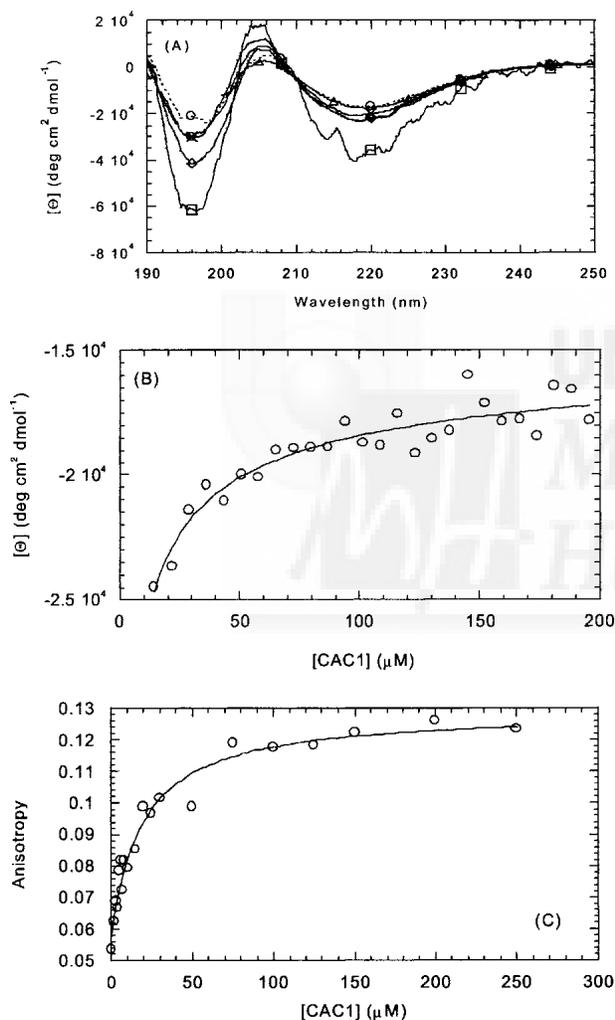


Figure 2. Determination of the self-association properties of CAC1 by circular dichroism and fluorescence anisotropy. (A) Representative spectra of CAC1 at different peptide concentrations: squares (5 μM), circles (20 μM), diamonds (30 μM), crosses (50 μM), and triangles (130 μM), all with continuous lines; and circles and dashed lines (100 μM). (B) Determination of the self-association constant by following the change in molar ellipticity at 222 nm. (C) Using the steady-state anisotropy of Trp184. The conditions were phosphate buffer (pH 7.3), 25 mM, at 298 K. The lines represent the fitting to equation 2 or the other similar equations (see text for details).

the concentration of peptide increased (Fig. 2A). All the spectra showed an isodichroic point at 210 nm (Fig. 2A), suggesting the presence of a two-state process, which can be described by a single equilibrium constant (Woody 1995; Kelly and Price 2000). Fitting of the change in mean ellipticity at 222 nm to equation 2^{'''} yielded a value for the self-dissociation constant of CAC1 of 17 ± 10 μM (Fig. 2B).

Fluorescence measurements

The fluorescence spectrum of CAC1 has a maximum at around 350 nm, corresponding to a tryptophan residue that is solvent-exposed (Pace and Scholtz 1997). It was observed that as the concentration of peptide increased, the maximum was blue-shifted, from 352 nm (at 2.3 μM) to 349 nm (at 120 μM; data not shown). This suggests that even at high peptide concentrations the tryptophan in CAC1 is not buried.

The local and global (segmental) mobility of the only indole moiety in CAC1 can be mapped by measurements of its fluorescence anisotropy (Lakowicz 1999). The anisotropy increased, as it did the ellipticity, in a hyperbolic manner from a value of 0.05 at 2 μM up to a value of 0.12 at 250 μM of CAC1 (Fig. 2C). This anisotropy shift could reflect either a change of the volume of the molecular species (global movement), or a change in the mobility of the nearby environment (local movement). The self-dissociation constant obtained from the anisotropy displacement was 16 ± 5 μM. This value was the same, within the error, than that determined by CD, in which the overall formation of secondary structure was being measured. This suggests that the anisotropy measurements were probably indicating a change in the global behaviour of the peptide.

NMR studies

The spectrum of CAC1 at 10 μM in ²H₂O (Fig. 3A) showed up-field signals clustered in those regions expected for random-coil models (Wüthrich 1986). The low-field region in ²H₂O only showed the protons of the aromatic ring of the tryptophan, which had chemical shifts close to those observed in random-coil tryptophans (data not shown). Thus, none of the amide protons remained after 5 min in ²H₂O (data not shown). All these findings further support the fact that the tryptophan was solvent-exposed, as suggested by the change in maxima wavelength of the fluorescence spectra of CAC1 (see before). Attempts to determine the structure of CAC1 by NMR failed because at the high concentrations needed to acquire two-dimensional NMR spectra, the sample precipitated (data not shown).

Design of an anthraniloyl-labeled CAC1 peptide and studies of its self-associating properties

To study the binding to CA-C by using fluorescence, a selective fluorescence probe must be used, which does not interfere with other intrinsic fluorescence probes in the pep-

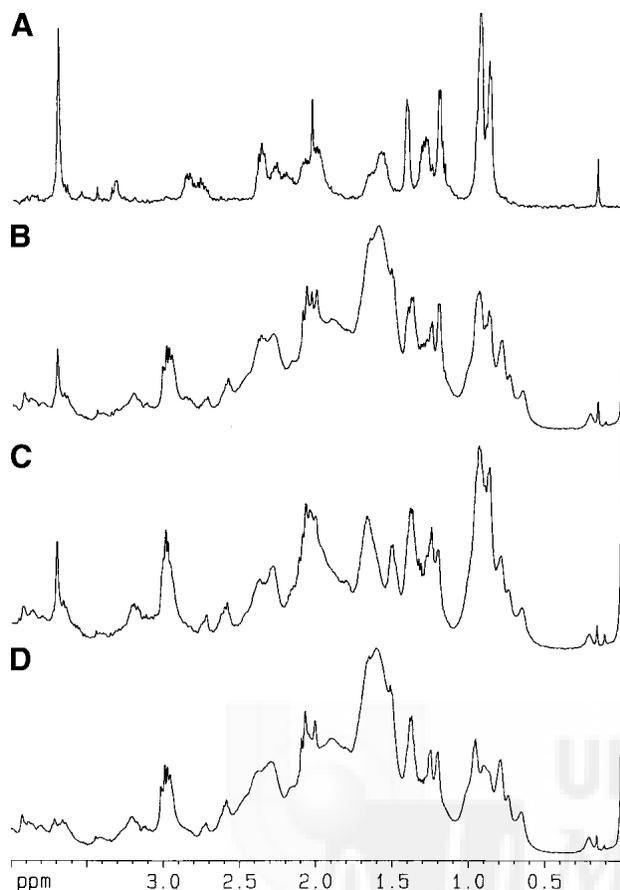


Figure 3. NMR spectra of CAC1 and CA-C. (A) Up-field region of the spectrum of CAC1 at 10 μM . (B) Up-field region of the spectrum obtained by addition of the spectra of CAC1 and CA-C. (C) Up-field region of the spectrum of the heterocomplex formed between CA-C and CAC1 at the same concentrations (10 μM of each macromolecule). (D) Up-field region of CA-C at 10 μM . The conditions were phosphate buffer (pH 7.3), 25 mM, at 298 K in $^2\text{H}_2\text{O}$ (no corrections were done for isotope effects). The sharp signal at the right side observed in all spectra, except in that of CAC1, is the NMR internal reference (TSP). The intensity of the signals for each NMR spectrum is arbitrary.

tide (such as tryptophan). We decided to label CAC1 with an anthraniloyl moiety (Stennicke et al. 1997), the maxima fluorescence emission wavelengths of which did not overlap with that of the tryptophan. The attachment of the anthraniloyl probe to the sole lysine of CAC1 during peptide synthesis yielded peptide CAC1Aib.

To check whether the introduction of the anthraniloyl moiety changed the intrinsic conformational tendencies of CAC1, we measured the self-associating properties of CAC1Aib by using two different approaches. First, we followed the change in the fluorescence of the anthraniloyl moiety upon binding to CAC1. The binding properties of CAC1Aib for CAC1 were determined by titrating a fixed amount of CAC1Aib (1.6 μM) with increasing concentrations of CAC1. Upon binding, the fluorescence intensity of

the anthraniloyl moiety at 420 nm did increase (data not shown). The self-dissociation constant determined by using equation 2' was similar, within the error, to that determined for the unlabeled peptide by CD and anisotropy measurements of the indole ring, with a value of $8 \pm 3 \mu\text{M}$. And, second, we used the change in the anisotropy of the anthraniloyl moiety as the concentration of unlabeled peptide was increased. The self-dissociation constant determined by using equation 2'' was similar to those previously shown, with a value of $9 \pm 3 \mu\text{M}$. These findings suggest that CAC1Aib and CAC1 self-associate in a similar way, and thus, both peptides are equivalent for studies of binding to CA-C.

Binding of the peptide CAC1 to the CA-C domain

We have used several biophysical techniques: first, to address whether CA-C and CAC1 interact; and, second, to quantitatively measure the value of such interaction.

CAC1-CA-C interaction

Thermal denaturation experiments. The dependence of CA-C thermal unfolding upon variation of the CAC1 concentration was firstly studied. The rationale is that, if CAC1 binds to CA-C, the observed thermal stability of the latter could be altered (Pantoliano et al. 2001). Binding of CAC1 to the native state of CA-C, in the absence of binding to the denatured state, will necessarily lead to an increase in the melting temperature; conversely, binding to the unfolded state of CA-C, in the absence of binding to the native state, will decrease the melting temperature relative to the protein in the absence of ligand (Waldron and Murphy 2003).

CA-C showed a unique thermal transition as followed by far-UV CD with a thermal midpoint at 336.1 K (Mateu 2002). CAC1, conversely, did not show any thermal transition at any concentration explored from 40 μM to 192 μM (data not shown). When thermal denaturation experiments were carried out at a CA-C concentration of 20 μM , and the CAC1 concentration was changed from 20 μM up to 60 μM , the thermal midpoint was invariably reduced (Fig. 4A). However, thermal denaturation led to precipitation of the sample. We do not know whether this precipitation was the result of the interaction of CAC1 when CA-C was partially unfolded (i.e., only partially folded CA-C species would be competent for binding), or on the other hand, CAC1 and CA-C interact and the complex showed irreversible aggregation upon heat-induced denaturation (i.e., binding would occur before the heterocomplex unfolds). In both scenarios, however, the decrease in the thermal midpoint must be due to the interaction.

SEC experiments. SEC of CA-C yielded a single peak at any protein concentration (Mateu 2002). This indicates that the equilibrium between the CA-C monomer and CA-C dimer is fast compared with the elution time (Corbett and Roche 1984). The elution volume of CA-C alone at 20 μM

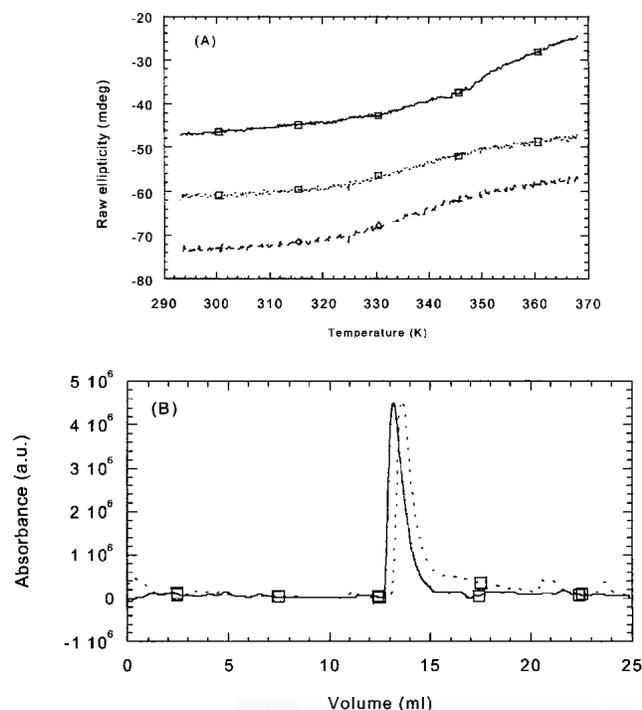


Figure 4. CA-C and CAC1 interactions followed by far-UV thermal denaturation, and gel filtration. (A) Changes in ellipticity followed at 222 nm (far-UV CD) upon thermal denaturation of CA-C isolated at 20 μM (continuous line and squares) and that of the complex (dotted line and squares) in a 20 μM /20 μM rate of protein to peptide, and (dashed lines and diamonds) 20 μM /40 μM rate of protein to peptide. The conditions were phosphate buffer (pH 7.3), 25 mM. Spectra were acquired in 0.1-cm-pathlength cells. (B) Squares and continuous lines indicate the chromatogram at 20 μM of CA-C; the squares and the dotted lines, at 20 μM of CA-C and 128 μM of CAC1. The small shoulders at the right side of the peak corresponded probably to unbound CAC1. The conditions were phosphate buffer (pH 7.3), 25 mM, and 150 mM NaCl at 298 K.

was 13.16 mL (Fig. 4B). As the concentration of CAC1 increased, the elution volume of the heterocomplex increased: 13.22 mL (at 60 μM of CAC1) and 13.52 mL (at 128 μM of CAC1). Because the inherent precision in the size of the determination of globular proteins by gel partitioning methods has been shown to be ± 0.1 mL in elution volumes (Ackers 1967), the differences found in the elution volumes in these experiments are significant. These results suggest that CAC1 interacts with CA-C and, it provides evidence that this interaction could involve the dissociation of the CA-C dimer, because the larger elution volume indicates a reduction in the size of the macromolecule.

NMR experiments. The one-dimensional NMR spectrum in $^2\text{H}_2\text{O}$ of CA-C was characterized by the presence of methyl groups up-field shifted at 0.2 and 0.5 ppm (Fig. 3D). In the amide region, a large dispersion of the amide protons between 8 and 9 ppm was observed, and the spectrum showed also a complete absence of protons between 5 and 6 ppm (data not shown). Furthermore, when dissolved in

$^2\text{H}_2\text{O}$, several amide protons remained after 8 h. All these features are expected for a folded α -helical protein, as shown by its X-ray structure (Gamble et al. 1997).

If CAC1 heterodimerizes with CA-C and this leads to dissociation of the CA-C homodimer (as suggested by the SEC experiments), the NMR signals should be expected to change (in chemical shift, line broadening, and/or intensity). Different experiments were carried out at 10 μM concentration of CAC1, 10 μM of CA-C, and 20 μM of the complex formed between CAC1 (at 10 μM) and CA-C (at 10 μM ; Fig. 3). The spectrum of the complex between CAC1 and CA-C obtained by mixing 10 μM of CAC1 and 10 μM of CA-C (Fig. 3C) was different to the sum spectrum (Fig. 3B) obtained by adding the spectra of CAC1 (Fig. 3A) and CA-C (Fig. 3D). The main differences were observed in the region between 1.5 and 2.0 ppm, where the methylene protons corresponding to the side chains of amino acids with a long alkyl chain (such as Lys or Arg) appear (Wüthrich 1986). Minor changes were also observed in the methyl region between 0.8 and 1.0 ppm (Fig. 3B,C). In the region of the H_α protons, where usually changes in secondary structure are easily detected, no changes occurring upon binding were observed, probably due to the fact that CA-C is an all α -helical protein, and there is a large number of overlapping resonances (CA-C is 86 residues long) in the region between 3.60 and 4.40 ppm. The fact that there were differences between the experimental spectrum of the CAC1–CA-C heterocomplex and the spectrum obtained by adding the spectra of the individual species (CA-C and CAC1) unambiguously indicates that there was binding between CAC1 and CA-C. Because the experiments were carried out in $^2\text{H}_2\text{O}$ (see Materials and Methods), no reliable information could be obtained from the amide region, because all the amide protons of CAC1 disappeared in the dead-time of the experiment.

We also tried to follow the changes in chemical shifts at 100 μM of CA-C concentration (data not shown), with increasing concentrations of CAC1 (ranging from 100 μM to 200 μM). However, in this range no changes in chemical shifts, signal intensity, or line broadening were observed. Thus, at these high NMR concentrations, no evidence of complex formation was observed. Similar behavior has been observed in other proteins with binding that is followed by NMR at high concentrations (Arac et al. 2003), for which such results have been interpreted as a limit for the apparent affinity constant of the corresponding complexes. This would indicate that the dissociation constant for the CAC1–CA-C heterocomplex is <100 μM .

Measurement of the apparent affinity constant of the CAC1–CA-C heterocomplex

Thermal denaturation, one-dimensional NMR, and SEC analysis provided unambiguous evidence of an interaction between CAC1 and CA-C, but none of these techniques

could be used to obtain the affinity constant. To that end, we used three additional experimental approaches.

First, the changes in the fluorescence emission intensity of the anthraniloyl moiety of CAC1Aib at 420 nm, upon CA-C binding, were analyzed (see Materials and Methods). No precipitation was observed in the explored protein concentration range (Fig. A). The value obtained for the apparent dissociation constant, according to equation 1, was $30 \pm 10 \mu\text{M}$.

As a second method to determine the dissociation constant of the CAC1–CA-C heterocomplex, we used analytical affinity chromatography, on a column containing covalently coupled CA-C. No precipitation was observed, and data could be well fit to equation 5 (Fig. 5B). The value obtained for the apparent dissociation constant of the heterocomplex was $43 \pm 2 \mu\text{M}$. The good fitting obtained with our measurements, and previous studies with analytical affinity chromatography in other proteins (Dunn and Chaiken 1974) and CA-C (Rosé et al. 1992) indicate that the CA-C protein attached to the column remains monomeric, and the heterocomplex is formed by only one molecule of CA-C

Finally, we measured the heat evolved upon CAC1–CA-C interaction by ITC (Fig. 5C). Precipitation was only observed at high CAC1 concentrations, but the apparent dissociation constant could be determined. The estimated value was $79 \pm 7 \mu\text{M}$. This value was somewhat higher than those obtained by the other techniques, but it could be due to the precipitation observed in the ITC experiments at the highest peptide concentrations. Precipitation would be hampered in the affinity experiments because of binding to the solid phase and, in the fluorescence experiments, because of the low concentrations of CA-C and CAC1Aib used. In any case, the three independent analyses yielded reasonably similar values (averaging $50 \pm 25 \mu\text{M}$) for the apparent dissociation constant of the CAC1–CA-C heterocomplex.

Discussion

A protein yet to be exploited as a potential target for HIV drugs is the capsid protein of HIV, an attractive candidate because of its essential role in virion morphogenesis (von Schwedler et al. 1998; Tang et al. 2001; Forshey et al. 2002). Agents that affect capsid assembly or its stability can be potential strong inhibitors of viral replication, as it has shown in other viruses, such as in the hepatitis B virus (Deres et al. 2003) and rhinoviruses (Smith et al. 1986). For HIV, we have undertaken the design of an inhibitor of the CA-C dimerization able to mimic, both structurally (Gamble et al. 1997) and energetically (del Álamo et al. 2003), the CA-C dimerization interface. The results described here indicate that a short peptide, CAC1, which encompasses mainly the sequence corresponding to the α -helix 2 of the CA-C domain of HIV, shows an affinity for this domain that is only about five times lower than that of

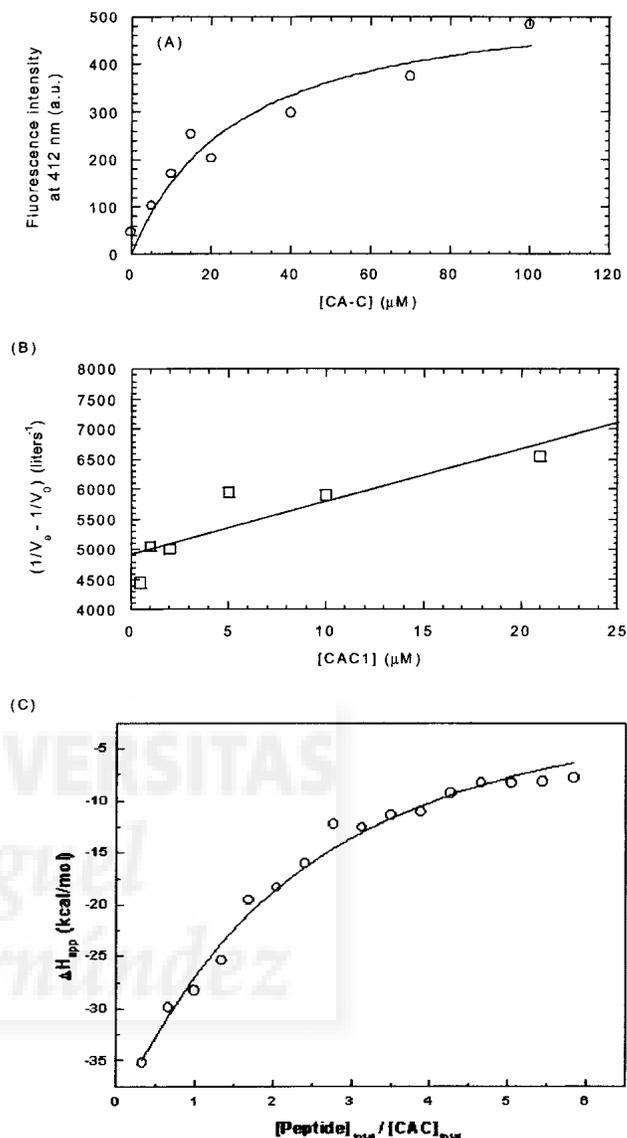


Figure 5. Binding of CAC1 to CA-C. (A) Steady-state fluorescence of the anthraniloyl-labeled CAC1 upon addition of increasing amounts of CA-C, by excitation at 320 and observation at 420 nm (a similar curve was observed by excitation at 246 nm). The value obtained for the dissociation constant was $30 \pm 10 \mu\text{M}$. (B) Affinity chromatography. The value obtained for the dissociation constant was $43 \pm 2 \mu\text{M}$. (C) ITC isotherm. The value obtained for the dissociation constant was $79 \pm 7 \mu\text{M}$. The conditions in all cases were phosphate buffer (pH 7.3), 25 mM, at 298 K. The lines represent the fitting to equation 1 for the fluorescence data (a similar curve was obtained by using equation 2), to equation 5 for the affinity chromatography results, and to the equation included with the Microcal software for a single-site model with a complex 1:1 stoichiometry (ITC data).

the intact protein, and then, it is able to mimic most of the dimerization interface of CA-C domain.

Structure of the CAC1 peptide

CAC1 is mainly disordered in solution as shown by NMR, fluorescence, and CD. However, the CD spectra showed

intense bands, which are not characteristic of fully disordered structures, but they are probably due to the sole tryptophan (Freksgård et al. 1994; Woody 1995; Kelly and Price 2000). As the concentration of peptide was increased, the ellipticity at 222 nm was reduced (Fig. 2A,B), suggesting that either some preferred helical population was reduced or the environment around the tryptophan (the sole aromatic residue) changed (Freksgård et al. 1994; Woody 1995; Kelly and Price 2000). In fact, upon increasing concentration, the indole moiety also shows a small bluelight-shifted fluorescence spectra and its anisotropy increased (Fig. 2C). The isolated peptide has a very low tendency to form helical structure, as shown by the CD spectra and the theoretical predictions obtained by using the program AGADIR (Muñoz and Serrano 1994), which yielded 2.2% of helical structure. All these data suggest that probably the environment around tryptophan changes upon self-association, although we cannot rule out that these changes are concomitantly inducing native-like helical structure. Native-like structure has been shown to be induced around bulky side chains of aromatic residues in some model peptides and proteins during their folding (Simmerling and Elber 1995; Kemmink and Creighton 1996). However, it is unclear whether the structure of CAC1 upon binding to CA-C is native-like. As it has been shown, peptides can interact with proteins in a wide variety of ways (Stanfield and Wilson 1995); in CAC1, it could be that the environment around tryptophan (and then native-like structure around it) could be kept upon CA-C binding.

The self-dissociation constant of the peptide has a mean value (averaging over all the biophysical measurements) of $12 \pm 5 \mu\text{M}$. This value is very close to that obtained for the homodimerization of the intact CA-C domain ($10 \mu\text{M}$; Gamble et al. 1997), suggesting that homodimerization of CA-C is mainly governed by α -helix-2- α -helix-2 interactions (see below). The oligomerization state of CAC1 could not be ascertained. However, the gel filtration studies of isolated CAC1 suggest that at moderately high concentrations, the peptide is a dimer. Attempts to determine the CAC1 molecular weight from the observed elution volume yielded a value of $4800 \pm 500 \text{ D}$, which is close to the expected dimer molecular weight for a monomer of 2318.56 D (data not shown).

Molecular recognition of CA-C by peptide CAC1

The CA-C domain and the CAC1 peptide interact, as it has been shown here by different biophysical techniques. However, it could be thought that the intrinsic tendency of CAC1 to self-associate is responsible for binding to CA-C through unspecific association. Negative control experiments carried out with enzyme I, a dimer protein involved in phosphorylation systems in bacteria, suggest that the interaction between CA-C and CAC1 corresponding to its α -helix 2 is

specific. These control experiments involved the determination of the Stern-Volmer quenching constant of the anthraniloyl moiety of CAC1Aib under three different solution conditions: (1) in the presence of enzyme I (EI); (2) in the presence of CA-C; and, (3) the isolated CAC1Aib peptide. Either in the presence of EI or in the absence of CA-C, the Stern-Volmer constant of the anthraniloyl moiety was lower ($1.0 \pm 0.1 \text{ M}^{-1}$) than in the presence of the CA-C domain ($1.7 \pm 0.1 \text{ M}^{-1}$). We do not know the reasons why the quenching constant of the anthraniloyl is increased upon binding, but they could be due to the fact that upon binding the peptide acquires a native-like conformation, because in the X-ray structure of CA-C (Gamble et al. 1997), the backbone of Lys182 is completely buried but its side chain is partially exposed. This suggests that binding between both macromolecules involves the side chains.

The NMR experiments also show that the most important regions involved in binding are the side chains of the amino acids of both species, as judged by the changes in the up-field region of the spectra (Fig. 3). This is further supported by the changes occurring upon binding to CA-C in the fluorescence spectra of the anthraniloyl moiety, which is attached to the side chain of Lys182. It has been shown that peptide binding may involve regular secondary structure of the binding protein (such as it happens in the MHC II class peptides), but also, as it happens in the MHC class I molecules, the side chains of the macromolecules may be the most affected regions upon peptide binding (Stanfield and Wilson 1995; Davis and Teague 1999). In CAC1 and CA-C, binding would involve the side chains of the CA-C, although changes in either peptide or protein backbones at this stage cannot be ruled out. NMR analyses of ^{15}N -labeled CA-C and its complex with CAC1 are being carried out in our laboratories to provide further insights on this subject.

The similarity between the self-association constant of CAC1 and that of CA-C and, most importantly, the relatively small difference between the dissociation constant of the CAC1-CA-C heterocomplex and that of the CA-C homodimer indicate that many important interactions for CA-C recognition are contained within α -helix 2, in agreement with previous structural (Gamble et al. 1997) and thermodynamic studies (del Álamo et al. 2003) with the CA-C homodimer. The peptide mimics most of the interactions that form the dimerization interface of CA-C. The difference between the affinity values obtained for the peptide-protein heterodimer and the protein homodimer might be due to entropic considerations. In fact, there is a large conformational entropy penalty paid upon binding of CAC1 to CA-C, mainly because of the larger flexibility of CAC1 alone compared with the CA-C monomer. Also, the difference must rely on the contribution of other critical residues for dimerization present in the intact domain, including Ile150, Leu151, Arg154, and Leu171 (del Álamo et al. 2003). The structural differences between the peptide and

the intact domain may be also in part responsible for the difference between both constants. In particular, it should be taken into account that mutation of Gln192 at the CA-C interface to Ala led to a 20-fold increase in the homodimerization affinity (del Álamo et al. 2003). This residue is included in the CAC1 sequence, but its structural environment at the C-terminal end of the peptide is probably different from that in dimeric CA-C, and its negative effect on the affinity might not be fulfilled as part of the peptide. Whatever the exact reasons, it is remarkable that, despite lacking some critical residues for homodimerization and not being folded in solution, the isolated α -helix 2 from CA-C is able to specifically bind the domain with such high affinity.

However, it is not known yet whether the heterocomplex is formed by only *one* single species of each macromolecule (i.e., *one* molecule of CAC1 and another *one* of CA-C). Several lines of evidence suggest that the heterocomplex is formed by only a single molecule of each species: (1) the elution volume in the gel filtration experiments of the heterocomplex (Fig. 4B); and (2) the fact that, within the experimental error of our measurements, when the binding is followed by the fluorescence of the anthraniloyl moiety (Fig. 5A) there is no lag phase at low peptide concentrations before the sigmoidal curve. This lag phase would indicate the formation of dimer CAC1 species, which would be the sole species capable of binding to CA-C. And, finally, the good fitting to equation 5 and previous results with analytical affinity chromatography and CA-C (Rosé et al 1992) indicate that the CA-C attached to the column is monomeric.

To conclude, we have shown that a 20-mer peptide containing the major dimerization helix of the C-terminal domain of CA from HIV-1 is able to specifically interact with and dissociate the CA-C dimer. The association constant of such peptide is modest, but it is only five times lower than the homodimerization constant of both CA-C and intact CA. Thus, CAC1 might be a model for conceiving peptidomimetics that show the same binding properties to CA-C.

Materials and methods

Materials

Imidazole, Trizma base, and NaCl were from Sigma. Standard suppliers were used for all the other chemicals. Dialysis tubing was from Spectrapore with a molecular mass of 3500 D. Water was deionized and purified by using a Millipore system.

Protein expression and purification

The wild-type CA-C protein was expressed in *Escherichia coli* BL21(DE3) and purified as described (Mateu 2002; del Álamo et al. 2003). Protein was stored and analyzed in 25 mM sodium phosphate buffer (pH 7.3). Protein stocks were run in SDS-PAGE

gels and found to be >97% pure. Purity was also confirmed by MALDI mass spectrometry analysis (data not shown). Protein concentration was calculated from the absorbance of stock solution measured at 280 nm, using the extinction coefficients of model compounds (Pace and Scholtz 1997).

Wild-type dimeric EI protein from *Streptomyces coelicolor*, used as a control dimeric protein in binding experiments (see below), was expressed in *E. coli* BL21(DE3) strain and purified by using Ni²⁺-NTA chromatography (Invitrogen). To eliminate any protein or DNA bound to the resin coeluting with the protein, an additional gel filtration chromatography step was carried out by using a Superdex 200 16/60 gel filtration column (Amersham Biosciences) running on an AKTA-FPLC machine.

Peptide design, synthesis, and purification

Peptide CAC1 (Ac-EQASQEVKNWMTETLLVQNA-CONH₂) was derived from the CA-C sequence, and comprised residues 175–194 of the wild-type CA protein. Those amino acids form α -helix 2 of CA-C (residues Ser178–Val191, i.e., most of the structural and energetic CA-C dimerization interface) plus three flanking residues on each side to avoid fraying effects at the ends of the peptide. The N and C termini were, respectively, acetylated or amidated to avoid charge effects. An anthraniloyl-fluorescent version of CAC1, CAC1Aib, was obtained by direct incorporation, during automatic synthesis, of an anthraniloyl moiety into the ϵ -amino group of the sole Lys of the peptide (Lys182). The CAC1Aib peptide showed absorption maxima at 246 and 320 nm (data not shown), as it should be expected from the anthraniloyl moiety (Stennicke et al. 1997). Purity of both peptides was assessed by mass spectrometry and HPLC chromatography and found to be >95%. Both peptides were purchased from Dr. Graham Bloomberg (University of Bristol, UK). Peptide concentrations were determined by the weight of the corresponding amount of dried peptide.

Fluorescence measurements

Fluorescence spectra were collected either on a SLM 8000 spectrofluorometer (Spectronics Instruments) interfaced with a Haake water bath, or in a Cary Eclipse spectrofluorometer (Varian) interfaced with a Peltier cell. Sample concentration was varied in the range 0.5–250 μ M for CAC1 and 0.5–200 μ M for intact CA-C. A 0.5-cm-pathlength quartz cell (Hellma) was used for measurements in the SLM spectrofluorometer, and a 1-cm-pathlength quartz cell (Hellma) was used in the Varian spectrofluorometer.

Determination of the apparent dissociation constant of the complex between CAC1 and CA-C

Increasing amounts of CA-C, in the range 0.5–120 μ M, were added to a solution of a fixed concentration of CAC1Aib (1.6 μ M) in 25 mM sodium phosphate buffer (pH 7.3), and the fluorescence was measured after 2-h incubation at 298 K. Experiments were carried out in the Varian Cary Eclipse spectrofluorometer with excitation at 246 nm (or 320 nm), and emission fluorescence was collected between 400 and 500 nm. The excitation and emission slit widths were both 5 nm, and the data pitch interval was 1 nm. Lower concentrations of the CAC1Aib (0.9 μ M) could not be monitored due to a poor signal-to-noise ratio (data not shown).

The fluorescence intensity of CAC1Aib at 420 nm changed upon binding to CA-C. The apparent dissociation constant was

calculated by fitting that fluorescence change versus the concentration of added CA-C to:

$$F_{meas} = F^0 + \Delta F_{max} \left(\frac{([CACIAib] + [CA - C] + K_D)}{4[CA - C][CACIAib]} - \sqrt{\frac{([CACIAib] + [CA - C] + K_D)^2 - 4[CA - C][CACIAib]}{4[CA - C][CACIAib]}} \right) \quad (1)$$

or its simplified form:

$$F_{meas} = F^0 + \Delta F_{max} \frac{[CA - C]}{[CA - C] + K_D} \quad (2)$$

where $[CACIAib]$ is the concentration of the anthraniloyl-labeled peptide, $[CA - C]$ is the concentration of the CA-C domain, F_{meas} is the measured fluorescence at each concentration of the CA-C domain, ΔF_{max} is the change in the fluorescence measured when all CA-C is forming the heterocomplex, F^0 is the fluorescence of the labeled peptide, and K_D is the apparent dissociation constant of the heterocomplex. Both equations yielded the same set of parameters (data not shown).

The anthraniloyl-labeled peptide stock solution was prepared by dissolving the corresponding amount in methanol (because of its low solubility in water) to yield a final stock solution with a concentration of 410 μM . Aliquots were added to the titrating solutions (either the CA-C domain or the unlabeled peptide, see below) in such small volumes that the residual final methanol concentration in the sample was <0.4%. These small concentrations of organic solvent did not change the structure or self-association properties of either the CA-C domain or the unlabeled peptide, as suggested by the CD and fluorescence spectra of both molecules (data not shown). At these small concentrations, CACIAib was soluble. Absorbance of the CACIAib (and CAC1) samples (ranging from 2 μM to 120 μM concentrations) at wavelengths >460 nm suggested the absence of high molecular weight aggregates (data not shown).

Determination of the self-dissociation constant of CAC1 and CACIAib by steady-state fluorescence

Increasing amounts of CAC1, in the range 0.5–190 μM , were added to a solution containing CACIAib (1.6 μM) in 25 mM sodium phosphate buffer (pH 7.3), and the steady-state fluorescence of the mixture was measured after 2-h incubation at 298 K. The other experimental conditions were the same as those described above for determination of the apparent dissociation constant of the heterocomplex. Smaller concentrations of the labeled peptide (0.9 μM) could not be reliably monitored because of a poor signal-to-noise ratio (data not shown).

The fluorescence intensity of CACIAib at 420 nm changed upon binding to CAC1. The apparent self-dissociation constant was calculated by fitting that fluorescence change versus the concentration of added CAC1 to:

$$F_{meas} = F^0 + \Delta F_{max} \left(\frac{([CACIAib] + [CAC1] + K_D)}{4[CAC1][CACIAib]} - \sqrt{\frac{([CACIAib] + [CAC1] + K_D)^2 - 4[CAC1][CACIAib]}{4[CAC1][CACIAib]}} \right) \quad (1')$$

(1'), which is analogous to equation 1, but here K_D is the self-dissociation constant of CAC1. Alternatively, its simplified form was used:

$$F_{meas} = F^0 + \Delta F_{max} \frac{[CAC1]}{[CAC1] + K_D} \quad (2')$$

which is analogous to equation 2.

Determination of the self-dissociation constants of CAC1 and CACIAib by steady-state anisotropy

The steady-state anisotropy, $\langle r \rangle$, is defined by (Lakowicz 1999):

$$\langle r \rangle = \frac{I_V - G I_H}{I_V + 2G I_H}, \quad (3)$$

where I_V and I_H are, respectively, the parallel and perpendicular components of the polarized fluorescence light, and G is a correction factor for the different fluorescence response to the two polarized components. The G factor was determined for each point, and it was typically 0.84 for the indole moiety. Fluorescence anisotropy was recorded in the T arrangement (simultaneous recording of the parallel and perpendicular emitted light) in the SLM 8000 spectrofluorometer. Each reported anisotropy value is the average of five independent measurements. Determination of the anisotropy of the indole moiety was followed by excitation at 280 nm and emission at 350 nm.

Determination of the anisotropy of the anthraniloyl moiety was followed by excitation at 320 nm and emission at 420 nm. The G factor was determined for each point, and it was typically 0.58 for the anthraniloyl moiety. Each reported anisotropy value is the average of five independent measurements. Increasing amounts of CAC1, in the range 0.5–190 μM , were added to a solution containing CACIAib (1.6 μM) in 25 mM sodium phosphate buffer (pH 7.3), and the steady-state anisotropy of the mixture was measured after 2-h incubation at 298 K.

Anisotropies at any particular peptide concentration were fitted for both fluorescence moieties to:

$$\langle r \rangle_{meas} = \langle r \rangle^0 + \Delta \langle r \rangle_{max} \frac{[CAC1]}{[CAC1] + K_D} \quad (2'')$$

which is analogous to equation 2'.

Quenching measurements

Quenching experiments by following the fluorescence of the anthraniloyl moiety were carried by excitation at 320 nm and collecting spectra between 390 and 450 nm at 298 K. Excitation and emission slit widths were both 5 nm. Data interval was 1 nm. Ionic strength was kept constant by addition of KCl. Also, $\text{Na}_2\text{S}_2\text{O}_3$ was added to a final concentration of 0.1 M to avoid formation of I_3^- . Data were fitted to the linear Stern-Volmer equation (Lakowicz 1999). Each reported value of the Stern-Volmer constant (see Discussion) is the averaged result of four complete independent measurements. CACIAib concentration was 2 μM in all experiments, and CA-C and EI concentrations were 4 μM .

Circular dichroism measurements

CD spectra were collected on a Jasco J810 spectropolarimeter fitted with a thermostated cell holder and interfaced with a Neslab RTE-111 water bath. Isothermal wavelength spectra were acquired at a scan speed of 50 nm/min with a response time of 2 sec and averaged over four scans at 298 K. Typically, CAC1 concentrations were 10 to 300 μM in a 0.1-cm-pathlength cell. Smaller

peptide concentrations yielded a poor signal-to-noise ratio (data not shown). All spectra were corrected by subtracting the proper baseline.

The mean residue ellipticity, $[\Theta]$, was calculated according to:

$$[\Theta] = \frac{\Theta}{(10lcN)} \quad (4)$$

where Θ is the measured ellipticity, l is the pathlength cell, c is the protein concentration, and N is the number of amino acids.

Temperature measurements

Thermal-denaturation experiments in the far-UV were performed at constant heating rate of 60 °C/h and a response time of 8 sec. Thermal scans were collected at 222 nm from 298 K to 363 K in 0.1-cm-pathlength cells with a total CA-C concentration of 20 μM, and by adding increasing amounts of CAC1 (from 20 μM to 60 μM). The possibility of drifting of the CD spectropolarimeter was tested by running two samples containing buffer, before and after the thermal experiments. No difference was observed between the scans. Every thermal denaturation experiment was repeated at least twice with fresh new samples.

Thermal denaturation experiments were also carried out with peptide CAC1. The parameter set and the CAC1 concentrations used were the same than those described above.

Determination of the self-dissociation constant of CAC1

The change in ellipticity at 222 nm was followed as the CAC1 concentration increased. It was fitted to:

$$[\Theta] = [\Theta]^o + \Delta[\Theta]_{max} \frac{[CAC1]}{[CAC1] + K_D} \quad (2''')$$

which is analogous to equation 2 (see above).

Fitting to all binding equations was carried out by a least-squares analysis using the program Kaleidagraph (Abelbeck software) on a personal computer. In all equations, the derived fitting constants, K_D , and the corresponding spectroscopic parameters are highly dependent on the values of the initial guesses used in the fitting program; then, an analysis of the χ^2 parameter was carried out to select the proper values of those parameters. In all cases, the reported apparent dissociation constants are the average of three independent measurements.

Size-exclusion chromatography

Analytical size-exclusion chromatography (SEC) experiments were carried out by using an analytical gel filtration Superdex 75 HR 16/60 column (Amersham Biosciences) running on an AKTA-FPLC system at 298 K. Flow rates of 1 mL/min were used, and aliquots of 100 μL were loaded into the column after equilibration. The column was equilibrated with four column volumes of elution buffer (25 mM sodium phosphate at pH 7.3, containing 150 mM NaCl to avoid nonspecific interactions with the column). The column was calibrated by using the gel filtration low-molecular-weight calibration kit (Amersham Biosciences). The standards used and their corresponding Stokes radii were as follows: ribonuclease A (16.4 Å), chymotrypsinogen (20.9 Å), ovalbumin (30.5 Å), and bovine serum albumin (35.5 Å; Hinkle et al. 1999). Protein and peptide elution volumes were monitored by following absorbance at 280 nm. The CA-C concentration was in all cases

was 20 μM, and increasing amounts of CAC1 were added. The mixture formed by the protein and the peptide was incubated during 2 h at 298 K. The reported elution volume is the result of four independent measurements with fresh new samples.

Nuclear magnetic resonance spectroscopy

One-dimensional ¹H NMR experiments were carried out in a Bruker DRX-500 spectrometer at 298 K with 32,000 data points and using the WATERGATE sequence (Piotto et al. 1992) to eliminate the water signal. Typically, 25,000 scans were acquired, during 15 h, and the spectral width was 6000 Hz in all cases. CA-C concentration was 10 μM (pH 7.3) in 25 mM phosphate buffer at 298 K in ²H₂O. No corrections were done for isotope effects. NMR experiments were carried out in ²H₂O, although most of the amide signals disappeared, to increase the signal-to-noise ratio at the lower concentrations used. Increasing amounts of CAC1 were added, ranging from 10 μM to 20 μM. The spectra were processed by using the BRUKER-XWINNMR software working on a personal computer workstation. An exponential window function and polynomial base line corrections were applied. The final one-dimensional ¹H NMR data contained 64,000 data points. ¹H chemical shifts were quoted relative to internal TSP. Also, experiments were also carried out with 100 μM of peptide, protein, or the mixture (100 μM of peptide and 100 μM of protein) in ²H₂O at 298 K. In these experiments, 2048 scans were acquired.

Isothermal titration calorimetry studies

Isothermal titration calorimetry (ITC) experiments were carried out using an Omega instrument (MicroCal Inc.) coupled to an external nanovoltmeter, which was used to improve the signal to noise ratio (Cooper and Johnson 1994). All measurements were performed at 298 K in 25 mM sodium phosphate buffer (pH 7.3). Prior to the experiment, CA-C was exhaustively dialyzed at 277 K against the reaction buffer. The calorimetric cell was loaded with CA-C at a concentration of 21.5 μM, and CAC1 was loaded in the syringe at a total peptide concentration of 483 μM. Microliter amounts of the ligand solution (CAC1) were added sequentially to the calorimetric cell (1.37 mL), which contained the protein solution (CA-C). The amount of power required to maintain the reaction cell at constant temperature after each injection was monitored as a function of time. In an independent experiment, the individual heats of dilution for the CAC1 peptide (mainly due to dissociation of the peptide dimer upon dilution) were determined under the same experimental conditions by making identical injections of the CAC1 solution into the titration cell, containing only buffer. The heat due to the binding reaction between CAC1 peptide and CA-C protein was obtained as the difference between the heat of reaction (CAC1/CA-C) and the corresponding heat of dilution (CAC1/buffer). The resulting isotherm (individual apparent enthalpy changes upon binding versus molar ratio, [CAC1]/[CA-C]) was fitted to a single-site model assuming that the complex CA-C-CAC1 had a 1:1 stoichiometry. The thermodynamical parameters of binding were obtained by analyzing the data with the software provided by MicroCal.

Determination of the apparent binding constant by analytical affinity chromatography

The dissociation constants of the heterocomplexes formed with CA-C and CAC1 were also determined by affinity chromatogra-

phy (Dunn and Chaiken 1974). The CA-C protein was covalently bound to an Affi-Prep10 (BioRad) gel matrix according to the manufacturer's instructions. Five milliliters of gel with coupled CA-C were packed in a FPLC column (Amersham Biosciences) and equilibrated with elution buffer (25 mM sodium phosphate at pH 7.3 and 150 mM NaCl) at 298 K. Protein or peptide samples were applied in small zones (500 μ L), eluted at 1 mL/min and continuously monitored with an on-line UV detector set at 280 nm. The column was calibrated by using free CA-C applied at a total protein (monomer) concentrations, C_t , ranging from 0.5 μ M to 20 μ M. The elution volume, V_e , for each concentration was determined as that corresponding to the maximum in the chromatogram. The self-dissociation constant (homodimerization) of CA-C, K_D , and the amount of CA-C covalently bound to the matrix, M_t , were determined as described (Rosé et al. 1992) by linear fitting of the experimental data to:

$$\frac{1}{(V_e - V_o)} = \frac{K_D}{M_t} + \frac{C_t}{M_t} \quad (5)$$

where V_o is the elution volume for a noninteracting protein (such as carbonic anhydrase). The K_D thus obtained for CA-C was $9.8 \pm 0.4 \mu$ M, which was very similar to that obtained using equilibrium sedimentation (Gamble et al. 1997), and M_t was 9.0 ± 0.2 nmole. The use of this method or a different approach based on frontal elution to calibrate affinity CAC-AffiPrep10 columns (data not shown) yielded similar M_t values. To determine the dissociation constant of the complex between CA-C and CAC1, 500 μ L aliquots of CAC1 at C_t ranging from 0.5 μ M to 20 μ M were applied to the above described column. The $\frac{1}{(V_e - V_o)}$ values obtained at different concentrations were linearly fitted, and the apparent binding constant was directly obtained by multiplying the value of the y-axis intercept (i.e., for $C_t = 0$) by the M_t value previously determined, as described (Rosé et al. 1992).

Acknowledgments

We thank Dr. F. Titgemeyer for the generous gift of the recombinant plasmid encoding the enzyme I. This work was supported by grants from the Spanish Ministerio de Sanidad y Consumo (MSC; FIS 01/0004-02), Generalitat Valenciana (CTDIB2002-6) to J.L.N., by a grant from the Spanish Ministerio de Ciencia y Tecnología (MCyT; BIO2000-1081) to J.G., by grants from the MSC (FIS 01/0004-01) and MCyT (BIO2000-0408) to M.G.M., and by an institutional grant from Fundación Ramón Areces to the Centro de Biología Molecular "Severo Ochoa." M.C.L.-M. was recipient of a predoctoral fellowship from the Spanish MSC. F.N.B. was recipient of a predoctoral fellowship from the Spanish Ministerio de Educación, Cultura y Deporte. We deeply thank May García, María del Carmen Fuster, and Javier Casanova for excellent technical assistance. We are also very grateful for the detailed suggestions provided by the referees.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

Ackers, G.K. 1967. Molecular sieve studies of interacting protein systems. I: Equations for transport of associating systems. *J. Biol. Chem.* **242**: 3026–3034.

- Arac, D., Murphy, T., and Rizo, J. 2003. Facile detection of protein–protein interactions by one-dimensional NMR spectroscopy. *Biochemistry* **42**: 2774–2780.
- Babé, L.M., Pichuantes, S., and Craik, C.S. 1991. Inhibition of HIV protease activity by heterodimer formation. *Biochemistry* **30**: 106–111.
- Borsetti, A., Öhagen, Å., and Göttlinger, H.G. 1998. The C-terminal half of the human immunodeficiency virus type 1 Gag precursor is sufficient for efficient particle assembly. *J. Virol.* **72**: 9313–9317.
- Cooper, A. and Johnson, C.M. 1994. Isothermal titration calorimetry. *Methods Mol. Biol.* **22**: 137–150.
- Corbett, R.J.T. and Roche, R.S. 1984. Use of high speed size-exclusion chromatography for the study of protein folding and stability. *Biochemistry* **23**: 1888–1894.
- Davis, A.M. and Teague, S.J. 1999. Hydrogen bonding, hydrophobic interactions, and failure of the rigid receptor hypothesis. *Agnew Chem. Int. Ed.* **38**: 736–749.
- del Álamo, M., Neira, J.L., and Mateu, M.G. 2003. Thermodynamic dissection of a low-affinity protein–protein interface involved in human immunodeficiency virus. *J. Biol. Chem.* **278**: 27923–27929.
- DeLano, W.L. 2002. *The PyMOL Molecular Graphics System*. DeLano Scientific, San Carlos, CA (<http://www.pymol.org>).
- Deres, K., Schroeder, C.H., Paessens, A., Goldmann, S., Hacker, H.J., Weber, O., Kramer, T., Niewohner, U., Pleiss, U., Stoltefuss, J., et al. 2003. Inhibition of hepatitis B virus replication by drug-induced depletion of nucleocapsids. *Science* **299**: 893–896.
- Dorfman, T., Bukovsky, A., Öhagen, Å., Hoglund, S., and Göttlinger, H.G. 1994. Functional domains of the capsid protein of human immunodeficiency virus. *J. Virol.* **68**: 8180–8187.
- Dunn, B.M. and Chaiken, I.M. 1974. Quantitative affinity chromatography: Determination of binding constants by elution with competitive inhibitors. *Proc. Natl. Acad. Sci.* **71**: 2382–2385.
- Forshey, B.M., von Schwedler, U.K., Sundquist, W.I., and Aiken, C. 2002. Formation of a human immunodeficiency virus type 1 core of optimal stability is crucial for viral replication. *J. Virol.* **70**: 8645–8652.
- Freskgård, P.-O., Mårtensson, L.-G., Jonasson, P., Jonsson, B.-H., and Carlsson, U. 1994. Assignment of the contribution of the tryptophan residues to the circular dichroism spectrum of human carbonic anhydrase II. *Biochemistry* **33**: 14281–14288.
- Gamble, T.R., Vajdos, F.F., Yoo, S., Worthylake, D.K., Houseweart, M., Sundquist, W.I., and Hill, C.P. 1996. Crystal structure of human cyclophilin A bound to the amino terminal domain of HIV-1 capsid. *Cell* **87**: 1285–1294.
- Gamble, T.R., Yoo, S., Vajdos, F.F., von Schwedler, U.K., Worthylake, D.K., Wang, H., McCutcheon, J.P., Sundquist, W.I., and Hill, C.P. 1997. Structure of the carboxyl-terminal dimerization domain of the HIV-1 capsid protein. *Science* **278**: 849–853.
- Ganser, B.K., Li, S., Kliskho, V.Y., Finch, J.T., and Sundquist, W.I. 1999. Assembly and analysis of conical models for the HIV-1 core. *Science* **283**: 80–83.
- Garg, R., Gupta, S.P., Gao, H., Babu, M.S., Debnath, A.K., and Hansch, C. 1999. Comparative quantitative structure–activity relationship studies on anti-HIV drugs. *Chem. Rev.* **99**: 3525–3601.
- Gitti, R.K., Lee, B.M., Walker, J., Summers, M.F., Yoo, S., and Sundquist, W.I. 1996. Structure of the amino-terminal core domain of the HIV-1 capsid protein. *Science* **273**: 231–235.
- Gross, I., Hohenberg, H., and Kräulisch, H.G. 1997. In vitro assembly properties of purified bacterially expressed capsid proteins of human immunodeficiency virus. *Eur J. Biochem.* **249**: 592–600.
- Gross, I., Hohenberg, H., Huckangel, C., and Kräulisch, H.G. 1998. N-terminal extension of human immunodeficiency virus capsid protein converts the in vitro assembly phenotype from tubular to spherical particles. *J. Virol.* **72**: 4798–4810.
- Hinkle, A., Goranson, A., Butters, C.A., and Tobacman, L.S. 1999. Roles for the troponin tail domain in thin filament assembly and regulation: A deletion study of cardiac troponin T. *J. Biol. Chem.* **274**: 7157–7164.
- Kelly, S.M. and Price, N.C. 2000. The use of circular dichroism in the investigation of protein structure and function. *Curr. Protein Peptide Sci.* **1**: 349–384.
- Kemmink, J. and Creighton, T.E. 1996. The physical properties of local inter-

- actions of tyrosine residues in peptides and unfolded proteins. *J. Mol. Biol.* **245**: 251–260.
- Lakowicz, J.R. 1999. *Principles of fluorescence spectroscopy*, 2nd ed. Plenum Press, New York.
- Li, S., Hill, C.P., Sundquist, W.I., and Finch, J.T. 2000. Image reconstructions of helical assemblies of the HIV-1 CA protein. *Nature* **407**: 409–413.
- Maroun, R.G., Gayet, S., Benleulmi, M.S., Porumb, H., Zaragarian, L., Merad, H., Leh, H., Mouscadet, J.F., Troalen, F., and Fermadjina, S. 2001. Peptide inhibitors of HIV-1 integrase dissociate the enzyme oligomers. *Biochemistry* **40**: 13840–13848.
- Mateu, M.G. 2002. Conformational stability of dimeric and monomeric forms of the C-terminal domain of human immunodeficiency virus. *J. Mol. Biol.* **318**: 519–531.
- Misselwitz, R., Hausdorf, G., Welfe, K., Höhne, W.E., and Welfe, H. 1995. Conformation and stability of recombinant HIV-1 capsid protein p24 (rp24). *Biochem. Biophys. Acta* **1250**: 9–18.
- Muller-Dethlefs, K. and Hozba, P. 2000. Non-covalent interactions: A challenge for experiment and theory. *Chem. Rev.* **100**: 143–167.
- Muñoz, V. and Serrano, L. 1994. Elucidating the folding problem of helical peptides using empirical parameters. *Nat. Struct. Biol.* **1**: 399–409.
- Pace, C.N. and Scholtz, J.M. 1997. Measuring the conformational stability of a protein. In *Protein structure*, 2nd ed. (ed. T.E. Creighton), pp. 253–259, Oxford University Press, Oxford, UK.
- Pantoliano, M.W., Petrella, E.C., Kwanoski, J.D., Lobanov, V.S., Myslik, J., Graf, E., Carver, T., Asel, E., Springer, A.R., Lane, P., et al. 2001. High-density miniaturized thermal shift assays as a general strategy for drug discovery. *J. Biomol. Screening* **6**: 429–440.
- Piotto, M., Saudek, V., and Sklenar, V. 1992. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR* **2**: 661–665.
- Rosé, S., Hensley, P., O'Shanessy, D.J., Culp, J., Debouck, C., and Chaiken, I. 1992. Characterization of HIV-1 p24 self-association using analytical chromatography. *Proteins* **13**: 112–119.
- Simmerling, C.L. and Elber, R. 1995. Computer determination of peptide conformations in water: Different roads to structure. *Proc. Natl. Acad. Sci.* **92**: 3190–3193.
- Smith, T.J., Krmer, M.J., Luo, M., Vriend, G., Arnold, E., Kamer, G., Rossmann, M.G., McKinlay, M.A., Dian, G.D., and Otto, M.J. 1986. The site of attachment in human rhinovirus 14 for antiviral agents that inhibit uncoating. *Science* **233**: 1286–1293.
- Stanfield, R.L. and Wilson, I.A. 1995. Protein-peptide interactions. *Curr. Opin. Struct. Biol.* **5**: 103–113.
- Stennicke, H.R., Olesen, K., Sørensen, S.B., and Breddam, K. 1997. C-terminal incorporation of fluorogenic and affinity labels using wild-type and mutagenized carboxypeptidase. *Anal. Biochem.* **248**: 141–148.
- Tang, S., Murakami, T., Agresta, B.E., Campbell, S., Freed, E.O., and Levin, J.G. 2001. Human immunodeficiency virus type 1 N-terminal capsid mutants that exhibit aberrant core morphology are blocked in initiation of reverse transcription in infected cells. *J. Virol.* **75**: 9357–9366.
- Tang, C., Loeliger, E., Kinde, I., Kyere, S., Mayo, K., Barklis, E., Sun, Y., Hunag, M., and Summers, M.F. 2003. Antiviral inhibition of the HIV-1 capsid protein. *J. Mol. Biol.* **237**: 1013–1020.
- Verkhiver, G.M., Bouzida, D., Gehlhaar, D.K., Rejto, P.A., Freer, S.T., and Rose, P.W. 2002. Complexity and simplicity of ligand-macromolecule interactions: The energy-landscape perspective. *Curr. Opin. Struct. Biol.* **12**: 197–203.
- von Schwedler, U.K., Stemmler, T.L., Kliskho, V.Y., Li, S., Albertine, K.H., Davis, D.R., and Sundquist, W.I. 1998. Proteolytic refolding of the HIV-1 capsid protein amino-terminus facilitates viral core assembly. *EMBO J.* **17**: 1555–1568.
- Waldron, T.T. and Murphy, K.P. 2003. Stabilization of proteins by ligand binding: Application to drug screening and determination of unfolding energetics. *Biochemistry* **42**: 5058–5064.
- Wang, C.T. and Barklis, E. 1993. Assembly, processing and infectivity of human immunodeficiency virus type 1 Gag mutants. *J. Virol.* **67**: 4264–4273.
- Woody, R.W. 1995. Circular dichroism. *Methods Enzymol.* **246**: 34–71.
- Wüthrich, K. 1986. *NMR of proteins and nucleic acids*. John Wiley & Sons, New York.
- Zhang, Z.-Y., Poorman R.A., Maggiora, L.L., Heinrikson, R.L., and Kézdy, F.J. 1991. Dissociative inhibition of dimeric enzymes. *J. Biol. Chem.* **266**: 15591–15594.



An extensive thermodynamic characterization of the dimerization domain of the HIV-1 capsid protein

MARÍA C. LIDÓN-MOYA,¹ FRANCISCO N. BARRERA,¹ MARTA BUENO,^{2,3}
 RAÚL PÉREZ-JIMÉNEZ,⁴ JAVIER SANCHO,^{2,3} MAURICIO G. MATEU,⁵
 AND JOSÉ L. NEIRA^{1,3}

¹Instituto de Biología Molecular y Celular, Universidad Miguel Hernández, 03202 Elche (Alicante), Spain

²Departamento de Bioquímica y Biología Molecular y Celular, Facultad de Ciencias and ³Biocomputation and Complex Systems Physics Institute, Universidad de Zaragoza, 50009 Zaragoza, Spain

⁴Departamento de Química-Física, Facultad de Ciencias, Universidad de Granada, 18071 Granada, Spain

⁵Centro de Biología Molecular "Severo Ochoa" (CSIC-UAM), Universidad Autónoma de Madrid, 28049 Cantoblanco Madrid, Spain

(RECEIVED December 29, 2004; FINAL REVISION May 9, 2005; ACCEPTED May 22, 2005)

Abstract

The type 1 human immunodeficiency virus presents a conical capsid formed by several hundred units of the capsid protein, CA. Homodimerization of CA occurs via its C-terminal domain, CA-C. This self-association process, which is thought to be pH-dependent, seems to constitute a key step in virus assembly. CA-C isolated in solution is able to dimerize. An extensive thermodynamic characterization of the dimeric and monomeric species of CA-C at different pHs has been carried out by using fluorescence, circular dichroism (CD), absorbance, nuclear magnetic resonance (NMR), Fourier transform infrared (FTIR), and size-exclusion chromatography (SEC). Thermal and chemical denaturation allowed the determination of the thermodynamic parameters describing the unfolding of both CA-C species. Three reversible thermal transitions were observed, depending on the technique employed. The first one was protein concentration-dependent; it was observed by FTIR and NMR, and consisted of a broad transition occurring between 290 and 315 K; this transition involves dimer dissociation. The second transition ($T_m \sim 325$ K) was observed by ANS-binding experiments, fluorescence anisotropy, and near-UV CD; it involves partial unfolding of the monomeric species. Finally, absorbance, far-UV CD, and NMR revealed a third transition occurring at $T_m \sim 333$ K, which involves global unfolding of the monomeric species. Thus, dimer dissociation and monomer unfolding were not coupled. At low pH, CA-C underwent a conformational transition, leading to a species displaying ANS binding, a low CD signal, a red-shifted fluorescence spectrum, and a change in compactness. These features are characteristic of molten globule-like conformations, and they resemble the properties of the second species observed in thermal unfolding.

Keywords: circular dichroism; chemical denaturation; fluorescence; NMR, protein stability; thermal denaturation

Reprint requests to: José L. Neira, Instituto de Biología Molecular y Celular, Edificio Torregaitán, Universidad Miguel Hernández, Avda. del Ferrocarril s/n, 03202 Elche (Alicante), Spain; e-mail: jlneira@umh.es; fax: +34-966658758.

Abbreviations: ASA, accessible surface area; ΔASA_{total} , the total accessible surface area exposed upon unfolding; ANS, 8-anilino-1-naphthalenesulfonate; CA, capsid protein of HIV (p24); CA-C, C-terminal domain of CA; CA-N, N-terminal domain of CA; CD, circular dichroism; DSC, differential scanning calorimetry; FTIR, Fourier transform infrared spectroscopy; GdmHCl, guanidinium chloride;

[GdmHCl]_{1/2}, the GdmHCl concentration at the midpoint of the chemical denaturation; ΔC_p , the heat capacity change of unfolding; ΔH_m , the thermal enthalpy change at the denaturation midpoint; ΔS , the entropy change; HIV, human immunodeficiency virus; LEM, linear extrapolation model; NMR, nuclear magnetic resonance; SEC, size-exclusion chromatography; TSP, 3-(trimethylsilyl) propionic acid-2,2,3,3-tetrakisodium salt; T_m , the thermal denaturation midpoint; UV, ultraviolet; V_e , elution volume in SEC experiments.

Article and publication are at <http://www.proteinscience.org/cgi/doi/10.1110/ps.041324305>.

The viral genome and several proteins of HIV are encased in a conical shell that is assembled from a single protein subunit (CA, p24) (Gelderblom 1991; Welker et al. 2000). Initially, this protein is synthesized as part of the polyprotein Gag precursor that assembles into an immature capsid at the cell membrane (Wills and Craven 1991). After proteolytic cleavage of Gag, CA forms the mature capsid that encloses the viral RNA, the nucleocapsid protein, and the reverse transcriptase (Fuller et al. 1997). In vitro, CA spontaneously assembles into helical structures and cones resembling the viral capsids (Groß et al. 1997, 1998; Ganser et al. 1999; Li et al. 2000). Several findings suggest that proper and correct capsid assembly is critical for viral infectivity (von Schwedler et al. 1998; Tang et al. 2001; Forshey et al. 2002). For instance, inhibition of CA assembly by mutations has a lethal effect (Groß et al. 1997; von Schwedler et al. 1998; Tang et al. 2001); furthermore, the alteration of capsid stability reduces virus replication (Forshey et al. 2002). CA is capable of different modes of association (Rosé et al. 1992; Dorfman et al. 1994; Misselwitz et al. 1995), and there is a wide body of evidence that dimerization through its C-terminal domain is an important driving force in virus assembly (Wang and Barklis 1993; Borsetti et al. 1998). Understanding the energetics, stability, and folding governing the CA–CA recognition interface could shed light into the poorly understood area of protein self-association and, in turn, could allow the use of such interface as a pharmacological target. Although some insight has been gained in the kinetics of protein self-association, very little is known about the structural aspects and the precise nature of interactions involved during protein self-association (Neet and Timm 1994; Honig and Yang 1995; Jaenicke and Lillie 2000). In fact, it is usually assumed that self-association proceeds via a partially exposed hydrophobic core in a molten globule-like species (Jaenicke and Lillie 2000).

The CA protein of HIV-1 is formed by two independent domains separated by a flexible linker (Gamble et al. 1996, 1997; Gitti et al. 1996; Momany et al. 1996). The N-terminal domain (residues 1–146 of the intact protein), CA-N, is composed of five coiled-coil α -helices, with two additional short α -helices following an extended proline-rich loop (Gamble et al. 1996; Gitti et al. 1996; Momany et al. 1996). The C-terminal domain (residues 147–231), CA-C, is a dimer both in solution and in the crystal form (Gamble et al. 1997); furthermore, the dimerization constant of CA-C is nearly the same ($10 \pm 3 \mu\text{M}$) as that of the intact CA ($18 \pm 1 \mu\text{M}$) (Gamble et al. 1997). Each CA-C monomer is composed of a 3_{10} helix followed by an extended strand and four α -helices connected by short loops

by the mutual docking of α -helix 2 from each monomer (residues Ser178–Val191), which buries Trp184 in dimer interface (Fig. 1). The two additional aromatic residues in each monomer, Tyr164 and Tyr169, are located in the hydrophobic core of each monomer, well away from the dimer interface. The folding and association of CA-C involves a monomeric intermediate that rearranges and dimerizes to yield the native dimer (Mateu 2002). The energetic role of the side chains at the dimerization interface has been determined by thermodynamic analysis using alanine mutants (del Álamo et al. 2003). These studies have shown that the side chain of Trp184 (Gamble et al. 1997; del Álamo et al. 2003), and those of Ile150, Leu151, Arg154, Leu172, Glu175, Val181, Met185, and Leu189 are key for CA-C dimerization (del Álamo et al. 2003).

In this work our aim has been to thoroughly characterize the thermodynamics of unfolding of a dimeric protein at different pHs and several protein concentrations, trying to get insight into the general mechanism governing the formation of the oligomer interfaces. The use of different biophysical probes, which give complementary information, is required to test how the different regions and elements of quaternary, tertiary, and secondary structure of CA-C unfold. To that end, the techniques used in this work were fluorescence, circular dichroism (CD), absorbance, nuclear magnetic resonance (NMR), Fourier transform infrared spectroscopy (FTIR), and size-exclusion chromatography (SEC). The ΔH_m and ΔC_p of unfolding of monomeric CA-C are similar to those found in other monomeric proteins of similar size. On the other hand, the enthalpy and the entropy of dissociation of dimeric CA-C are small when



Figure 1. Structure of CA-C dimer. X-ray structure of CA-C showing the dimeric structure of the domain. The monomers are depicted in different colors (gray and green). The side chains of Tyr164, Tyr169, and Trp184 are indicated as sticks in each monomer. The figure was produced with PyMOL (DeLano Scientific) and using the PDB file for CA-C (accession no. 1A43).

FI (Fig. 1). The dimerization interface is largely formed

compared with those of other small oligomeric proteins of similar size that dissociate and unfold in a coupled transition. This is due to the fact that the CA-C dimer dissociates into the monomeric species that has most of the native secondary and tertiary structure. Finally, we have found that monomeric CA-C undergoes a conformational transition during thermal unfolding before being fully unfolded. This transition leads to an intermediate, which could resemble the monomeric molten globule-like species of CA-C detected at low pH, but it is different to the species whose self-association originates the dimeric form.

Results

pH-induced unfolding experiments

The pH-dependent changes in the secondary and tertiary structure of CA-C were monitored by several spectroscopic techniques, which gave complementary information about the unfolding of the protein. Since the measured dissociation constant was about 10 μM (Gamble et al. 1997; Mateu 2002; del Álamo et al. 2003), several protein concentrations were tested: 1 mM (100% dimeric CA-C, used in the FTIR and NMR experi-

ments), 200 μM (90% dimeric CA-C), 20 μM , (~40% monomeric CA-C), and 2 μM (~75% monomeric CA-C, used in the intrinsic fluorescence, quenching, and ANS-binding experiments).

Steady-state fluorescence measurements

Fluorescence has been used to map pH-dependent transitions in the tertiary structure of the protein (Pace and Scholtz 1997) by following the changes in the maximum wavelength and the average emission intensity, $\langle\lambda\rangle$.

The emission fluorescence spectrum of CA-C at physiological pH and at a concentration of 200 μM showed a maximum at 340 nm, and thus, the spectrum was dominated by the emission of the sole tryptophan residue. As the pH was increased above neutral pH, the spectra became red-shifted toward 350 nm, due to basic unfolding of the protein. Although the pK_a of this titration could not be determined, its apparent high value is consistent with deprotonation of Tyr, Arg, and/or Lys residues (Cantor and Schimmel 1980). As the pH was decreased, the spectra were also red-shifted toward 350 nm, due to acid unfolding (Fig. 2A), yielding a titration curve with a pK_a of 4.3 ± 0.2 . This value was close to the ionization constant of the side chains of Glu, Asp, and/or the C-terminal carboxylate (Cantor and Schimmel 1980).

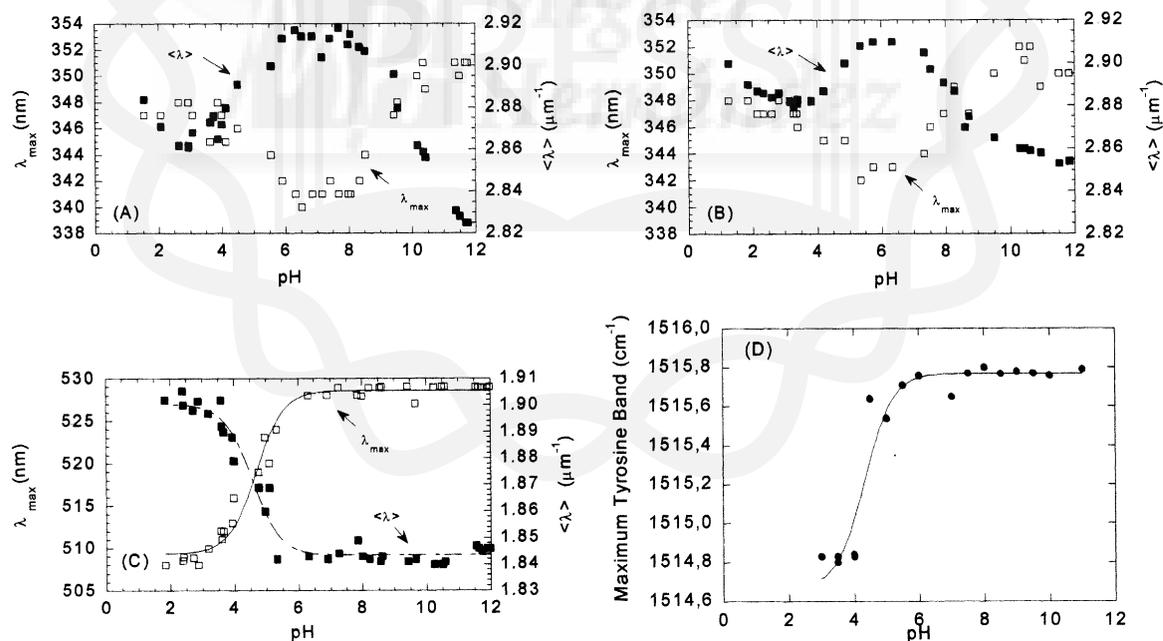


Figure 2. pH-induced unfolding of CA-C followed by fluorescence, ANS binding, and FTIR. Steady-state fluorescence: The $\langle\lambda\rangle$ (filled squares, right axis) and the maxima wavelength (open squares, left axis) are represented vs. the pH at two different CA-C concentrations. (A) 200 μM and (B) 20 μM . Experiments were acquired at 298 K. (C) ANS binding experiments, where the maxima wavelength (open squares, left axis) and the $\langle\lambda\rangle$ (filled squares, right axis) are represented vs. the pH. Protein concentration was 2 μM . The lines through the data are the fittings to Equation 5. Experiments were acquired at 298 K. (D) The maximum wavelength of the tyrosine band at different pHs as followed by FTIR. The line is the fit to Equation 5. Protein concentration was 1 mM. FTIR spectra were acquired at 278 K.

The same bell-shaped profile, with one transition midpoint at low pH and another at high pH, was observed when the $\langle\lambda\rangle$ was plotted versus the pH (Fig. 2A). The acidic transition yielded a pK_a of 4.2 ± 0.2 , which agrees with that determined using the maximum wavelength approach.

The maximum wavelength or the $\langle\lambda\rangle$ showed the same behavior as that described at high concentrations when experiments were carried out either at a 20 μM (Fig. 2B) or 2 μM (data not shown) protein concentration. At 20 μM , the maximum wavelength of the fluorescence spectrum was red-shifted (343 nm) at physiological pH (compared to the spectrum at 200 μM). This red shift was due to protein monomerization leading to increased solvent-exposure of the Trp. As the pH was either increased or decreased, the fluorescence spectra were further red-shifted toward 350 nm. The pK_a determined for the acid transition from the maxima of the spectra was 4.2 ± 0.2 , and that from $\langle\lambda\rangle$ was 4.1 ± 0.2 .

ANS-binding experiments

ANS binding is used to monitor the extent of exposure of hydrophobic regions, and to detect nonnative partially folded conformations. When ANS binds to solvent-exposed hydrophobic patches, its quantum yield is enhanced and the maxima of emission is shifted from 520 nm to 480 nm (Stryer 1965; Semisotnov et al. 1991). At low pH values, the intensity of ANS in the presence of monomeric CA-C was largely enhanced, and the maximum wavelength was 510 nm (Fig. 2C). As the pH was increased, the spectral intensity was reduced and the maximum wavelength shifted to 530 nm. Consistent effects were observed when the average emission intensity was examined: At low pHs, the $\langle\lambda\rangle$ was high, and it decreased as the pH was raised (Fig. 2C). The apparent pK_a determined from either set of data was 4.4 ± 0.2 , which agrees, within the experimental error, with those determined by the intrinsic protein fluorescence (see above).

Examination of tryptophan and tyrosine exposure by fluorescence quenching

The solvent-accessibility of tryptophan and tyrosine residues was examined by acrylamide quenching at several pHs. Three pHs were chosen based on the steady-state fluorescence results: pH 3; at neutral pH; and pH 11, in the basic region (Fig. 2A,B).

Acrylamide-quenching experiments, measured by excitation at 295 nm, yielded exponential Stern-Volmer plots in all experiments, except in that carried out at neutral pH. The K_{sv} parameter remained constant, within the error, at acidic and basic pHs, but it was smaller at neutral pH (Table 1). This suggests that at the extreme pHs, Trp184 was more solvent-exposed. Quenching control experiments in the presence of 5 M GdmHCl (pH 7), where the protein was unfolded (see below), yielded

Table 1. Quenching constants for CA-C in acrylamide at different excitation wavelengths

Conditions	280		295	
	$K_{sv}(\text{M}^{-1})$	$\nu(\text{M}^{-1})$	$K_{sv}(\text{M}^{-1})$	$\nu(\text{M}^{-1})$
pH 3	5.5 ± 0.6	1.2 ± 0.2	5.4 ± 0.2	0.67 ± 0.06
pH 7	4.4 ± 0.3	0.6 ± 0.1	3.2 ± 0.5	0.0 ± 0.1
pH 11	6.1 ± 0.2	1.2 ± 0.1	5.0 ± 0.8	0.7 ± 0.2
5 M GdmHCl	5.9 ± 0.6	1.0 ± 0.2	4.9 ± 0.6	0.4 ± 0.2

Errors are data fit errors to equation 1. The constants were obtained by the fitting to equation 1 of fluorescence intensity at 338 nm vs. the concentration of acrylamide (similar constants were obtained by fitting the intensities at 335, 336, and 337 nm; data not shown). Experiments were carried out at 298 K. The experiments at 5 M GdmHCl were carried out at pH 7.0.

values of K_{sv} similar to those observed at acidic or basic pHs. The behavior of the K_{sv} , upon excitation at 280 nm, followed the same pattern as that observed at 295 nm: higher at the extreme pHs and in 5 M GdmHCl than at neutral pH. The values of ν were the same, within error, at the extreme pHs and in 5 M GdmHCl. In general, K_{sv} and ν were larger when excited at 280 nm than at 295 nm (Table 1), probably because tyrosine and tryptophan residues were being excited at 280 nm.

Far-UV CD experiments

The far-UV CD spectrum of CA-C has the features of an α -helical protein, either at 20 μM or 200 μM , with minima at 222 nm and 208 nm (Mateu 2002), as expected from its three-dimensional structure (Gamble et al. 1997). However, interference from the aromatic residues at 222 nm cannot be ruled out (Woody 1995; Kelly and Price 2000). The shape of the spectrum did not change in the pH range explored, but the value of $[\Theta]$ at 222 nm was pH-dependent. At 200 μM CA-C, the mean ellipticity remained constant from pH 7 to 10, with a value of approximately $-13,000 \text{ deg cm}^2 \text{ dmol}^{-1}$ (Fig. 3A, inset). [F3] As the pH was increased, the absolute value of the ellipticity became progressively smaller, although a pK_a of this basic transition could not be determined because of the absence of baseline at high pHs. On the other hand, the absolute value of the ellipticity at acidic pHs showed a transition with a midpoint of 4.5 ± 0.2 . This value agrees well with those obtained by other techniques (see above).

A similar behavior was observed at 20 μM protein (Fig. 3A). The apparent pK_a of the acid transition was 4.4 ± 0.3 , which is identical, within the experimental uncertainty, to that obtained at 200 μM .

SEC experiments

Size-exclusion chromatography yielded a single peak at any pH irrespective of the concentration of CA-C used

[T1]

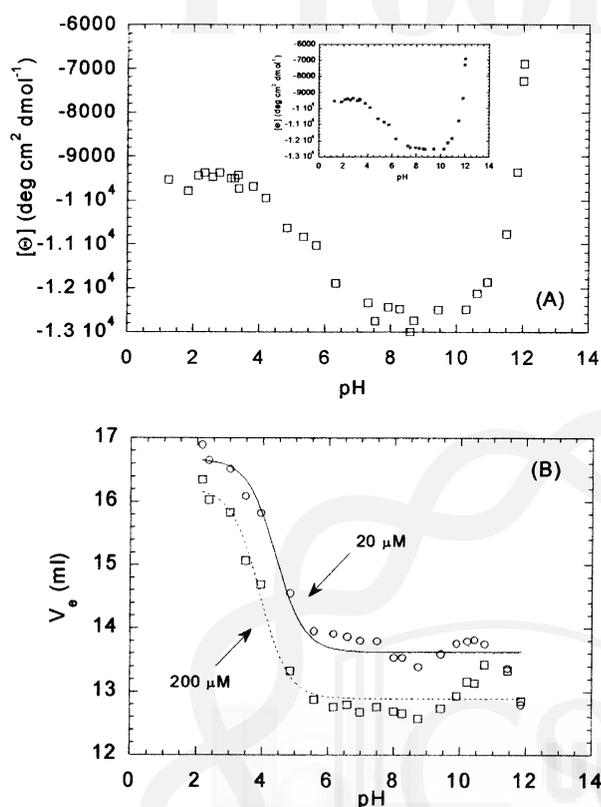


Figure 3. pH-induced unfolding of CA-C followed by CD and SEC. (A) Mean residue ellipticity at 222 nm at 20 μM of CA-C. (Inset) Mean residue ellipticity at 222 nm at 200 μM of CA-C. (B) Plots of the elution volume at 20 μM (open circles) and 200 μM (open squares) vs. the pH. The lines (continuous for 20 μM of CA-C; dashed for 200 μM) are the fits to Equation 5 (see text). All experiments were acquired at 298 K.

(20 μM or 200 μM). This single peak indicates that the exchange time between the monomeric and dimeric forms is faster than the rate of the chromatographic process; then, the differences in the elution volumes for both concentrations (Fig. 3B) must reflect the average of both species (monomer and dimer) at the concentrations used. In general, either at high pHs (pH > 11) or low pHs (pH < 2.5), the curves reflecting the behavior of the V_e tend to coalesce (Fig. 3B). There were titrations at low pHs with pK_a s of 4.3 ± 0.1 (at 20 μM CA-C) and 4.1 ± 0.1 (at 200 μM CA-C). They agree, within the error, with those measured by other techniques (see above). In both species, the V_e at low pHs is well away from the void volume of the column (8.00 mL), indicating that species observed are not oligomers. Between pH 10 and 11, a tendency to increase the V_e was observed at both concentrations, which was more evident at 200 μM (Fig. 3B).

FTIR experiments

Compared to CD, the main advantage of FTIR is its higher sensitivity to the presence or β -structure, random

coil, or some side chains, such as those of tyrosine. For instance, the maximum wave number of the tyrosine band appears, at low and neutral pHs, at 1514 cm^{-1} , and it should move toward smaller wave numbers (1505 cm^{-1}) when titration of its phenol proton occurs at basic pHs. However, the tyrosine band in CA-C did show a transition with an apparent pK_a of 4.3 ± 0.2 (Fig. 2D), similar to those previously described. This suggests that the transition may be associated with the protonation of an acidic residue close to at least one of the two tyrosine residues (Tyr164 or Tyr169).

In conclusion, the biophysical techniques used indicate that CA-C undergoes a protein concentration-independent conformational transition at low pH with a $pK_a \sim 4.3 \pm 0.2$.

Thermal-denaturation experiments

The thermal dissociation and unfolding of dimeric and monomeric CA-C has been characterized by using several biophysical techniques.

Far-UV CD experiments

Thermal denaturation of CA-C was probed at different pHs by following the change in ellipticity at 222 nm, using either a 20 μM (data not shown) or a 200 μM (Fig. 4A) protein concentration. The thermal scans at pHs 4, 5, 6, 7, 8, and 9 (Fig. 4A) revealed a sigmoidal, nonconcentration-dependent, fully reversible process (Table 2). Only at pH 10, the thermal transition was not reversible (data not shown), probably due to protein modifications occurring at the high alkalinity and temperatures. The fact that the transitions were not concentration-dependent suggests that the unfolding of a monomeric species was being observed. As the pH was increased, the T_m decreased slightly (Table 2). The plot of the ΔH_m , obtained from the van't Hoff analysis of the denaturation (Equations 6, 7), versus T_m yielded a $\Delta C_p = 1.8 \pm 0.5\text{ kcal mol}^{-1}\text{K}^{-1}$ (Fig. 4A, inset).

Near-UV CD and absorbance experiments

Changes in the ellipticity at 280 nm were followed at three selected pHs (pH 4, 7, and 8) and at a protein concentration of 200 μM (Table 2). At pH 7 and 8, a reversible sigmoidal transition was observed, but at pH 4, the steep slope of the native baseline (Fig. 4B), precluded the determination of reliable T_m and ΔH_m values. The thermal midpoints were lower than those obtained from far-UV CD measurements and similar to those observed in the fluorescence experiments (see below).

Experiments were also carried out at 20 μM CA-C to check for a concentration dependence (Fig. 4C). At this concentration, the slope of the native baseline was very

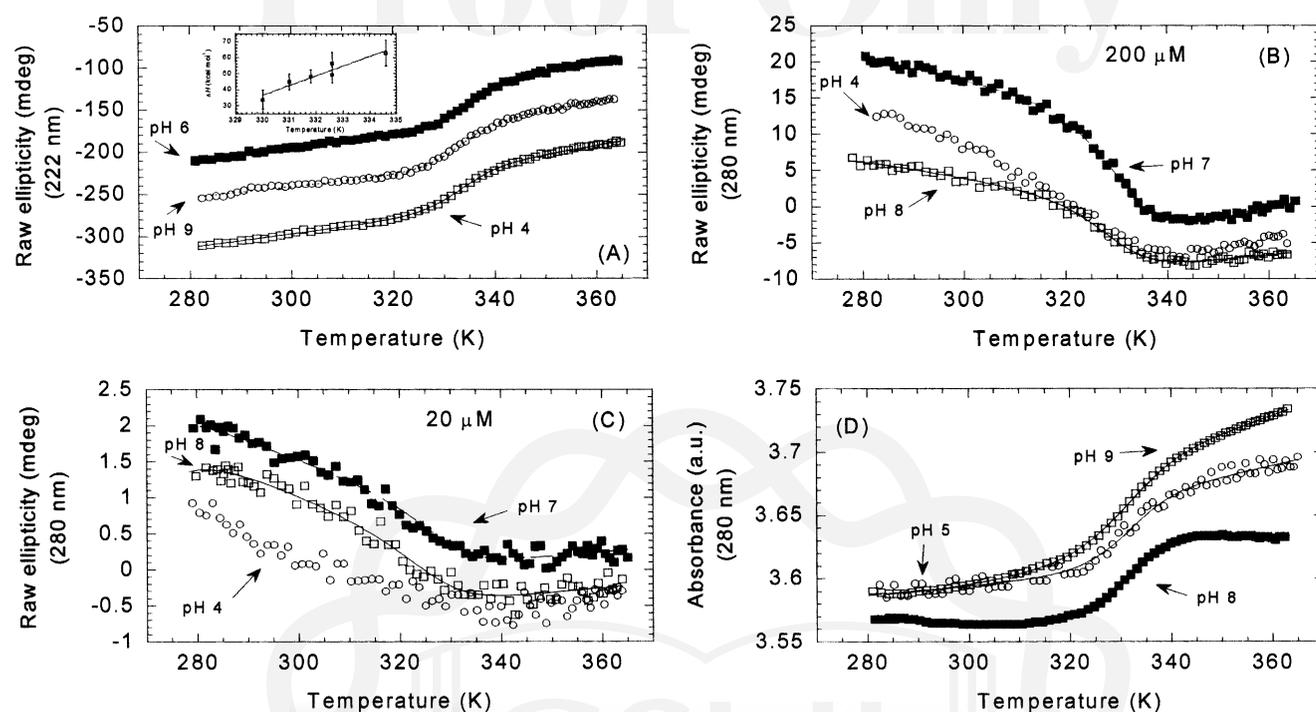


Figure 4. Thermal denaturation followed by far- and near-UV-CD, and absorbance at different pHs. (A) Selected far-UV thermal unfolding experiments at pH 4 (open squares), 6 (filled squares), and 9 (open circles) at 200 μM of CA-C. (Inset) Temperature dependence of the enthalpy change upon unfolding at 200 μM . In this temperature range, the unfolding of CA-C was characterized by a temperature-independent heat capacity change upon unfolding of $1.8 \pm 0.5 \text{ kcal mol}^{-1} \text{ K}^{-1}$. The errors in the enthalpy are fitting errors to Equation 7. (B) Selected near-UV thermal unfolding experiments at pH 4 (open circles), 7 (filled squares), and 8 (blank squares) at 200 μM of CA-C. (C) The same as (B) but at 20 μM protein concentration. (D) Selected absorbance thermal unfolding traces at pH 5 (open circles), 8 (filled squares), and 9 (open squares). The scale on the Y-axis is arbitrary and thermal unfolding traces have been shifted for the sake of clarity. Protein concentrations were 200 μM . The lines are the fittings to Equation 6, taking into account that the free energy is given by Equation 7.

steep, probably due to the substantial amount of the monomeric form present. This makes the determination of ΔH_m and T_m somewhat uncertain.

Changes in the absorbance at 280 nm were followed at pH 4, 5, 8, and 9 (Table 2; Fig. 4D) by using 200 μM protein concentration. The thermal stability of CA-C decreased at both extremes of pH (Table 2). Sigmoidal reversible curves were obtained at the four pHs; the measured T_m were close to those observed by far-UV, and, thus, higher than those observed by near-UV CD and fluorescence (see above and next paragraph). The finding that the thermal midpoint observed by absorbance agrees better with that observed by far-UV CD than with that obtained by near-UV is not new, and it has been reported in other proteins (Madonado et al. 1998; Irún et al. 2001; Campos et al. 2004). This can be rationalized by considering that the absorbance spectrum mainly reports changes in solvent exposure (Mach et al. 1995), and the exposition to solvent is dramatically affected by changes in secondary structure. On the other hand, the near-UV reports changes in the

asymmetric environment of the aromatic residues which could not be substantially altered when changes in the secondary structure occur.

Fluorescence experiments

A sigmoidal reversible transition was observed for pH 5, 6, 7, 8, and 9 at 200 μM protein concentration (Fig. 5A; Table 2). Nonreversibility was observed at pH 10, and a linear decrease in the intensity was observed at pH 4 as the temperature was raised (Fig. 5A). At 20 μM of CA-C for most of the pHs, there were not native baselines; this finding did not allow a reliable determination of the thermodynamical parameters of the transition, and most importantly, it precluded any reliable conclusion about whether the process probed by fluorescence is protein concentration-dependent or independent. Then, the calculation of T_m was carried out at a 200 μM protein concentration, where large baselines of the native and unfolded species could be observed. However, at this concentration, the steepness of the native and unfolded baselines precluded a reliable determination of ΔC_p .

Q2

Table 2. Thermal midpoints (T_m) (in K) of CA-C at several pH values determined using different biophysical techniques^a

Technique	Conditions								Protein concentration
	pH								
	3	4	5	6	7	8	9	10	
Fluorescence	ND		324.0 ± 0.2	325.7 ± 0.3	325.5 ± 0.2	324.5 ± 0.3	324.8 ± 0.5	NR	200
ANS-binding	ND	ND		318 ± 1	325 ± 1	325 ± 6	322 ± 2	ND	20
					325.2 ± 0.5				80
Anisotropy					325 ± 3 ^b				20
					326 ± 2				200
Near-UV CD	ND				326 ± 5 ^b	326 ± 5 ^b			20
	ND				327 ± 0.9	327.5 ± 0.9			200
Far-UV CD	328 ± 3	333.0 ± 0.3	332.5 ± 0.5	332.1 ± 0.3	331.7 ± 0.3	331.5 ± 0.3	NR	NR	20
	330 ± 1	331.0 ± 0.5	334.6 ± 0.3	332.6 ± 0.6	332.6 ± 0.5	331.8 ± 0.5	NR	NR	200
Absorbance	331.8 ± 0.8	330.5 ± 0.5				334 ± 0.7	328.6 ± 0.7		200
FTIR					NR				2000
	NR	NR	304.7 ± 0.3	307.2 ± 0.8	301.2 ± 0.6	NR			1000 ^d
				304.5 ± 0.6					530 ^d
				307.1 ± 0.8					1100 ^d
				307.9 ± 0.7					2000 ^d
				309.8 ± 0.8					4500 ^d
NMR	328 ± 2 ^a				334 ± 5 ^c				200

Errors are data fit errors to equation 6, using the free energy expression given by equation 7. Thermal denaturations followed by fluorescence were repeated three times at any pH and the standard deviation of the three measurements was $\pm 0.6^\circ\text{C}$. Thermal denaturations followed by ANS-binding were repeated three times at any pH, and the standard deviation was $\pm 3^\circ\text{C}$ in all cases. Each anisotropy measurement at any temperature was repeated four times, being the standard deviation ± 0.4 ; the measurements were repeated twice and the standard deviation was $\pm 4^\circ\text{C}$. Experiments in the far and near UV CD were repeated at least twice at any pH, the standard were $\pm 0.5^\circ\text{C}$ (200 μM) and $\pm 3^\circ\text{C}$ (20 μM) for the near-UV CD, and $\pm 0.4^\circ\text{C}$ for the far-UV CD experiments. Absorbance experiments were repeated twice and the standard deviation was $\pm 0.5^\circ\text{C}$. FTIR and NMR experiments were only acquired once at any pH, due to the large amounts of protein used. ND, Not determined (see text for each particular technique); NR, the thermal transition was not reversible (see text).

^a Determined from the indole signal (see text).

^b The transitions at 20 μM of CA-C have large associated fitting errors because of the poorer signal-to-noise ratio.

^c Determined from a methyl signal, with few experimental points at the unfolding baseline (see text).

^d These transitions describe the low-temperature transition observed in FTIR experiments (see text for details). The high-temperature transition was only characterized at pH 7.0.

The thermal midpoints at 200 μM of CA-C were lower than those determined by far-UV CD experiments (Table 2).

Thermal ANS binding experiments

A sigmoidal reversible transition was observed at pH 6, 7, 8, and 9 when thermal denaturation was followed by excitation at 380 nm of samples containing ANS (Table 2). The T_m were close to those obtained by fluorescence and near-UV and smaller than those observed by far-UV and absorbance, suggesting that those different techniques were reporting separated processes. In agreement with the previous spectroscopic techniques, the thermal stability decreased at the extremes of pH. The thermal midpoint was not protein concentration dependent between 20 μM and 80 μM of CA-C (Table 2). The plot of the ΔH_m versus T_m yielded a poor value of ΔC_p , due to the large fitting errors associated to ΔH_m .

Experiments at pH 4, 5, 10, and 11 did not show sigmoidal transitions (Fig. 5B). Conversely, at pH 10 and 11, where CA-C does not bind ANS (Fig. 2), in-

creasing the temperature caused a linear increase of the emission fluorescence intensity (Fig. 5B).

Thermal anisotropy measurements

The anisotropy gives a measurement of the mobility (global or local) of the tryptophan (Lakowicz 1999), as the temperature is raised. Experiments were carried out at 20 μM and 200 μM CA-C at pH 7 (Fig. 5C). The T_m were not concentration-dependent, and they agree with those obtained by near-UV CD, ANS binding, and fluorescence (Table 2).

FTIR measurements

Conversely to that observed by other techniques, thermal FTIR experiments carried out at pH 7 showed the presence of two transitions (Fig. 6A): (1) a transition with T_m of 305 K, which had not been observed by any of the previous techniques, and occurring with only small changes in the width of the amide I band, and (2) a transition accounting for most of the change in the width of the

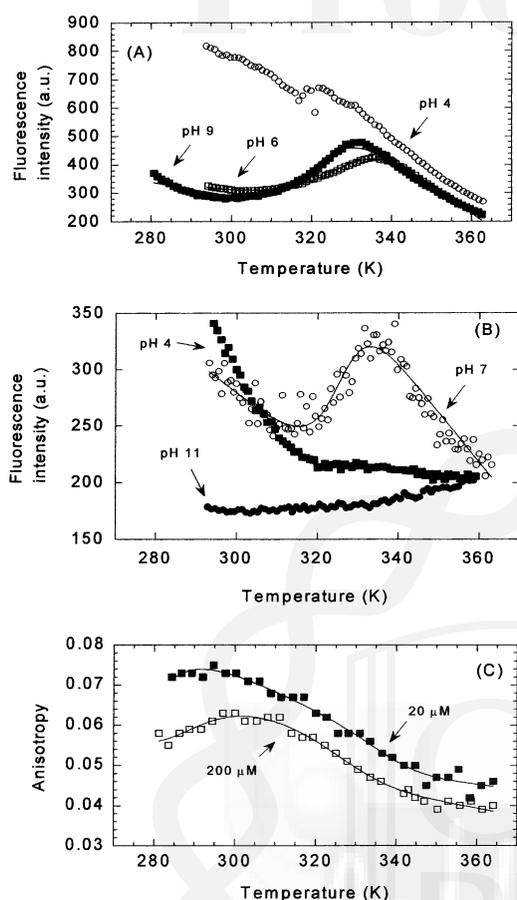


Figure 5. Thermal denaturation followed by protein fluorescence, ANS fluorescence, and anisotropy at different pHs. (A) Thermal unfolding traces followed by fluorescence at selected pHs for 200 μM of CA-C: pH 4 (open circles), pH 6 (white squares), and pH 9 (filled squares). (B) The thermal denaturation experiments followed by the emission intensity at 520 nm of 20 μM CA-C and 100 μM of ANS at several pHs: pH 4 (filled squares), pH 7 (open circles), and pH 11 (filled circles). The scale on the Y-axis for both fluorescence measurements is arbitrary. (C) The change in anisotropy at 20 μM (filled squares) and 200 μM of CA-C (blank squares) at pH 7. The lines are the fittings to Equation 6, taking into account that the free energy is given by Equation 7.

band, and with the same T_m (335 K) (Table 2) than those probed by far-UV CD or absorbance. The first thermal transition was reversible at physiological pH, as long as the temperature was not raised above 330 K. At pH 7, the transition observed at high temperatures was not reversible, probably due to the high protein concentrations used. Since it seems that the high-temperature transition was the same as that observed by the above techniques, a pH-dependent study of the low-temperature transition was carried out.

The low-temperature transition was explored at pH 4, 5, 6, 7, 8, and 9 (Fig. 6B), but it was only fully reversible at pH 6, 7, and 8. This low-temperature transition was protein concentration-dependent at pH 7 (Fig. 6C;

Table 2), and it showed an increase in the T_m as the CA-C concentration was raised. This variation can be used to determine the thermodynamic parameters governing the concentration-dependent reaction. In an oligomeric system, the dependence of T_m on C_t (where here C_t is the total molar protein concentration referred to the molar mass of the monomer) is given by $\frac{1}{T_m} = -\frac{R}{\Delta H^0} \ln(C_t) + \frac{\Delta S^0}{\Delta H^0}$, (Marky and Breslauer 1987;

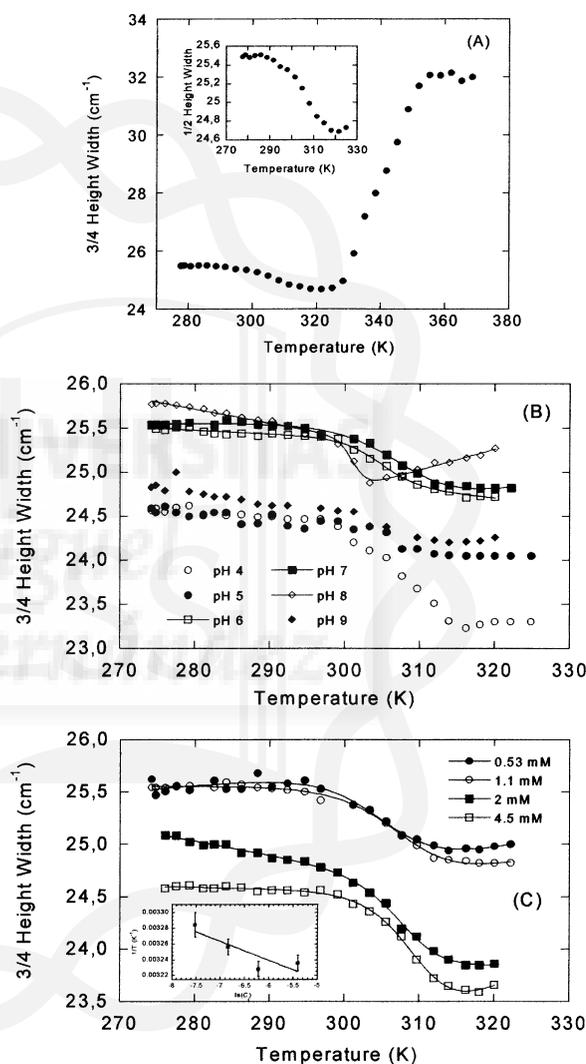


Figure 6. Thermal denaturation followed by FTIR. (A) The complete temperature range (280–370 K) for the FTIR experiments at pH 7, by measuring the 3/4 height width of the amide I band. Protein concentration was 1 mM. (Inset) Expansion showing the 3/4 of the height width of the band for the low-temperature transition. (B) The low-temperature transition observed by FTIR at different pHs (see label). Protein concentration was 1 mM. (C) The low-temperature transition at different protein concentrations (see label in the figure) at pH 7. (Inset) The protein-concentration dependence of the thermal midpoint for the low-temperature transition. The errors in the temperature are fitting errors to Equation 8. The lines are the fittings to Equation 6, taking into account that the free energy is given by Equation 8.

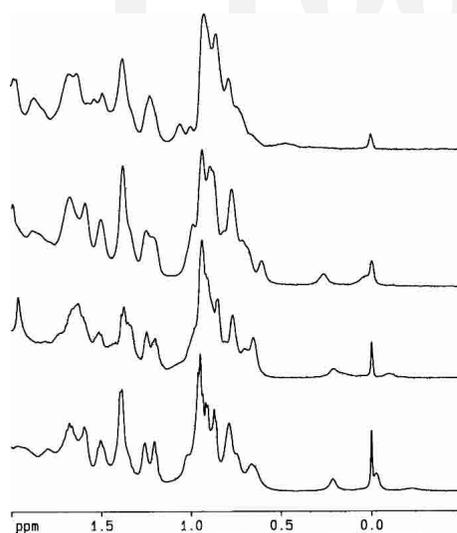


Figure 7. The up-field shifted methyl region of the NMR spectra at pH 7 and pH 3 at different temperatures. (From top to bottom) NMR spectrum at pH 3 and 320 K; pH 7 and 320 K; pH 3 and 298 K; and pH 7 and 298 K.

Steif et al. 1993), where ΔH^0 and ΔS^0 refer to the temperature-independent changes in enthalpy and entropy, respectively, under standard conditions at T_m . In CA-C, the thermodynamic parameters governing the dimer dissociation were (with $\frac{R}{\Delta H^0} = (2.7 \pm 0.4) \times 10^{-5} \text{K}^{-1}$, and $\frac{\Delta S^0}{\Delta H^0} = (30.7 \pm 0.2) \times 10^{-4} \text{K}^{-1}$) (Fig. 6C, inset) $\Delta H^0 = 74 \pm 8 \text{ kcal (mol of cooperative unit)}^{-1}$ and $\Delta S^0 = 230 \pm 30 \text{ cal K}^{-1} \text{ (mol of cooperative unit)}^{-1}$ at 1 M (standard state) protein concentration.

NMR measurements

NMR spectroscopy provides a wealth of information concerning the tertiary and secondary structure of a protein at residue level. The 1D- ^1H -NMR spectra were

acquired at pH 7 and 3 at 200 μM CA-C (and 900 μM at pH 7) (Fig. 7), where, as it has been observed by other techniques (see above), the protein is folded (pH 7), or has lost some of its tertiary structure, but it retains secondary structure (pH 3). [F7]

The 1D- ^1H -NMR spectra at 298 K and pH 3 showed a larger number of signals clustered between 8.0 and 8.5 ppm compared to that at pH 7; furthermore, there were only two methyl signals up-field shifted (Fig. 7). This result suggests that CA-C at pH 3 is less structured than at pH 7 (Wüthrich 1986).

We focused on the methyl and indol signals to follow the temperature changes, since those signals are well-separated from the rest of the resonances of the spectrum. In the spectrum at pH 7 and 330 K, all the amide protons were clustered between 8.0 and 8.5 ppm and the majority of the up-field shifted protons of the methyl region had disappeared, suggesting that most of the tertiary structure was lost. Conversely, at pH 3 and 320 K, the dispersion of the amide region and one of the two up-field shifted methyl protons observed at 298 K had disappeared (Fig. 7). Then, the NMR experiments confirmed that CA-C is less stable at low pH.

At pH 7, the protons can be classified in three groups, according to their temperature behavior: (1) Class I, protons appearing at -0.26 , -0.20 , and -0.037 ppm (Fig. 8A) were visible at low temperatures (between 278 and 288 K), but in the interval 290–310 K, the signals became broad and disappeared; (2) Class II, proton appearing at 0.2 ppm (Fig. 8B), whose chemical shift slightly increased with temperature between 280 K and 320 K, and then it did show a sigmoidal behavior. Although there were few experimental data in the unfolded region, fitting to Equations 6 and 7 was assayed and the thermal midpoint obtained was 337 ± 6 , close to those obtained by absorbance and far-UV CD; and (3) [F8]

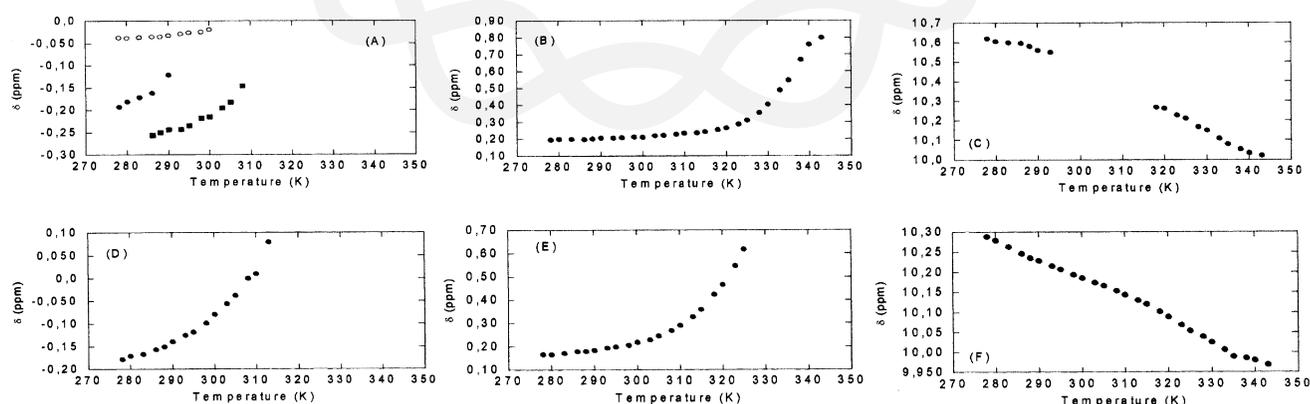


Figure 8. Temperature dependence of the chemical shifts of protons in the NMR spectra. At pH 7 (see text): (A) Class I, (B) Class II, (C) Class III. At pH 3 (see text): (D) Class I, (E) Class II (methyl proton), (F) Class II (indole proton). Conditions were 200 μM of protein in phosphate buffer (pH 7) (25 mM) or deuterated acetic acid buffer (pH 3) (25 mM) with 0.1 M NaCl.

Class III, consisting of the indole proton of Trp184 (Fig. 8C), whose chemical shift decreased with temperature. The signal became too broad between 290 K and 310 K, and it could not be observed again until temperatures higher than 310 K.

On the other hand, at pH 3, only two different classes of protons could be observed: (1) Class I, proton appearing at -0.20 ppm at low temperatures, whose chemical shift increased with temperature until 310 K, where it became very broad and disappeared (Fig. 8D); and (2) Class II, composed of the proton of the indole proton of Trp184 and a methyl proton at 0.2 ppm. Both protons could be observed in the whole range of temperatures explored. The methyl proton is probably the proton at the same chemical shift at pH 7 (Fig. 8E). Conversely to that observed at pH 7, the indole proton of the tryptophan at pH 3 was visible over all the explored temperature range (Fig. 8F).

Conformational stability and thermodynamic parameters of monomeric CA-C at pH 7

The inherently large fitting errors in the ΔH_m (obtained by some of the described techniques) yielded a ΔC_p value, for unfolding of monomeric CA-C, with a large uncertainty (1.8 ± 0.5 kcal mol $^{-1}$ K $^{-1}$). Then, it was necessary to obtain a better estimate of the value of ΔC_p by using other approaches. We have used an approach first developed by Pace and Laurents (1989), where the ΔG at a selected pH is measured at different temperatures by means of the linear extrapolation model (LEM) (Clarke and Fersht 1993; Pace and Scholtz 1997), and these data are combined with those obtained from thermal denaturations at identical pH. The GdmHCl chemical denaturation of CA-C measured by CD (where unfolding of the monomeric species is being probed) (Mateu 2002) follows the LEM. The isothermal chemical denaturation experiments by CD were

carried out at several temperatures in the range from 278 to 313 K at pH 7, and their free energies were obtained (Fig. 9). From the shape of the free-energy stability curve, the ΔH_m , ΔC_p , and T_m of the unfolding of the monomeric species could be obtained.

The m -values (Fig. 9A) were constant, within the error, in the temperature range explored. Although this tendency does not agree with the LEM model, a similar behavior has been observed in other proteins when chemical denaturations have been followed by CD (Pace and Scholtz 1997; Zweifel and Barrick 2002), and it can be rationalized by considering the presence of large slopes in the native or unfolded baselines. There was a good agreement between the data obtained from the isothermal chemical denaturation measurements and those from thermal denaturation experiments (Fig. 9B); this finding validates the use of the LEM, and most importantly, indicates that the same denatured state of CA-C is being probed by thermal and chemical denaturation measurements. A bell-shaped curve was observed when the $[\text{GdmHCl}]_{1/2}$ and the free-energy values were represented against the temperature (Fig. 9B). The temperature dependence of ΔG at pH 7 was consistent with a temperature-independent ΔC_p of 1.14 ± 0.06 kcal mol $^{-1}$ K $^{-1}$ (similar, within the error, to that obtained by far-UV measurements; see above), a T_m of 332.7 ± 0.1 K (which agrees with that determined previously by far-UV CD and NMR), and an ΔH_m of 54 ± 2 kcal mol $^{-1}$.

Discussion

The equilibrium thermal denaturation of CA-C follows a four-state mechanism

Dissociation of dimeric CA-C

The low-temperature transition probed by FTIR in thermal denaturation experiments was clearly concentration-dependent (Fig. 6), and it must reflect the thermal

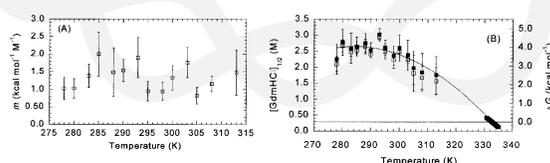


Figure 9. The GdmHCl-denaturation thermodynamical parameters at pH 7 as monitored by the change in ellipticity at 222 nm in the far-UV CD spectra. (A) Temperature dependence of the m -value from CD measurements. Errors bars are fitting errors to the LEM equation. (B) The temperature dependence of the $[\text{GdmHCl}]_{1/2}$ (left axis, open squares) and ΔG (right axis, filled squares) values. The error bars are fitting errors to the LEM equation. The errors are larger at the higher temperatures, because the native baselines were shorter. The solid line represents the nonlinear least square fit of the data to $\Delta G(T) = \Delta H(T_0) - T \cdot \Delta S(T_0) + \Delta C_p \cdot \left[T - T_0 - T \cdot \ln\left(\frac{T}{T_0}\right) \right]$, which is similar to Equation 7 except that here T_0 , the reference temperature, was taken as 298 K. The use of this equation avoids a bias in the fitting due to the larger number of experimental data around the thermal midpoint. From this equation, the ΔH_m , T_m , and ΔC_p can be easily obtained. The experimental data at the T_m (right side of figure) were obtained from CD thermal denaturation experiments. ΔG values were obtained by using the mean m -value over all the temperatures. The temperature-dependence of ΔG was consistent with a temperature-independent heat capacity change, ΔC_p , of 1.14 ± 0.06 kcal mol $^{-1}$ K $^{-1}$.

dissociation of the dimer. It is important to keep in mind that dimer dissociation was not only observed by FTIR at the very high concentrations used (~ 1 mM), but also by NMR at lower concentrations (200 and 900 μ M). Furthermore, there are two pieces of evidence, from the NMR experiments at pH 7, which suggest that a transition occurs: (1) The indole signal disappeared in the temperature range from 290 to 310 K (Fig. 7C); the tryptophan is a key residue in the dimerization interface (Gamble et al. 1997; Mateu 2002), and it must be involved in a slow conformational equilibrium, within the NMR time scale as the temperature was increased; and (2) there were methyl signals which started broadening and then disappeared at ~ 290 K (Fig. 7A,B). Since the rest of the signals of the spectrum did not broaden and disappear (data not shown), this finding must indicate that only some particular regions were involved in the transition. Recently, the changes in the breadth and the shape of the signal of an indole proton have been explained as due to the presence of intermediates species in the unfolding of the monomeric B1 domain of protein G (Ding et al. 2004).

The conformational states of monomeric CA-C

As concluded from the variation in T_m (Table 2), the stability of monomeric CA-C increased from pH 4 to a plateau around pH 6, and then decreased toward the alkaline region. Interestingly, there are two groups of techniques, which show two different midpoints. Far-UV CD, NMR and absorbance showed a T_m which was higher (~ 6 K) than that obtained by near-UV CD, ANS binding, and fluorescence anisotropy. The noncoincidence of protein unfolding curves, following different spectroscopic properties, is classical evidence of the accumulation of an equilibrium folding intermediate in the unfolding process (Luo et al. 1995, 1997).

[F10] There are at least three conformational states for the monomeric species (Fig. 10): (1) the monomeric form observed above 308 K, as a result of dimer dissociation (see before); (2) the species obtained from a conformational rearrangement, whose T_m was probed by ANS binding, fluorescence anisotropy, and near-UV CD; and (3) the final denatured state, obtained in a transition whose T_m was mapped by far-UV CD, absorbance, and NMR. In all cases, since the transitions were sharp and well separated by ~ 6 K, we fitted each transition to a two-state curve (Equations 6, 7; Table 2). The fact that different techniques show the same thermal midpoints allow to rule out that the different T_m s are the result of the fitting procedure. Moreover, to further rule out any skew in the calculations, we carried out a global fitting analysis of the far and near-UV CD data at 20 μ M and 200 μ M of protein concentration; the analysis shows that there are three states with

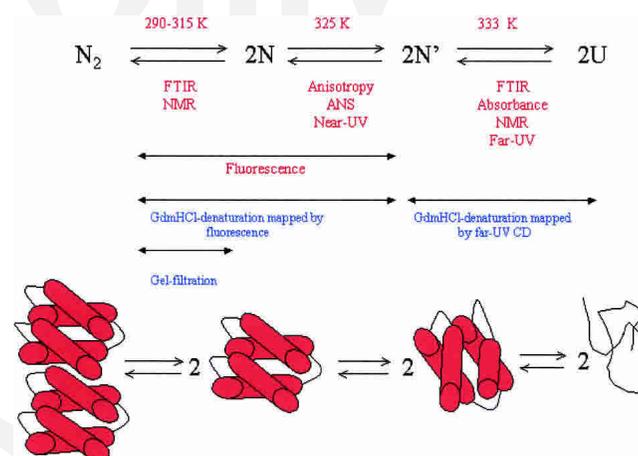


Figure 10. The equilibrium species observed in unfolding of CA-C at pH 7. U, denatured monomer; N' , partially rearranged monomer; N, folded monomer; N_2 , native dimer. The figures at the bottom of the scheme are for illustrative purposes only, and the rearranged reaction ongoing from N to N' might not involve fraying of a particular α -helix. The techniques used in this work are indicated below the arrows; the temperature where the transition was observed is at the top of the arrow. The red letters indicate the transitions mapped by thermal denaturation (this work). The blue letters indicate evidence for the detection of the different transitions by chemical denaturation experiments in previous work (Mateu 2002; del Álamo et al. 2003).

$T_{m1} = 329 \pm 3$ and $T_{m2} = 335 \pm 3$, which are similar to those obtained from the individual two-state fitting of each curve. The dominant monomeric species between both temperatures binds ANS, whereas the monomer obtained as a direct product of dimer dissociation does not. This finding has important implications for the mechanism of oligomer formation, since it has been suggested that assembly of oligomers occurs via a molten globule-like species (Ptitsyn 1995; Jaenicke and Lillie 2000). On the other hand, given the structural features of the rearranged monomer, it is tempting to suggest that this species could be similar to that observed at low pH at 298 K.

At low pH CA-C, in either the monomeric or dimeric form, undergoes a conformational transition, with a pK_a of 4.3 (Figs. 2, 3). The low pK_a suggests that the titration corresponds to one acidic group. Since the titration was also followed by the change in the amide I band corresponding to tyrosine residues in FTIR, it must be an acidic residue close to at least one of the two tyrosine residues in CA-C (Tyr164 or Tyr169), probably Asp166. It is important to stress here that this pK_a is an apparent value since it is the average of the titration midpoints of the corresponding residues in the folded and the unfolded CA-C species, appearing as the pH was lowered (Anderson et al. 1990). We could not determine from that value how many kilocalories CA-C was destabilized upon acid unfolding, but rather, we could

estimate at what pH the protein was unfolded, as it has been done in other electrostatic studies (Chen et al. 2000; Forsyth and Robertson 2000). In the acidic transition some of the CA-C tertiary structure is lost (as shown by fluorescence), but some does remain, as indicated by the up-field shifted methyl protons in the NMR spectra (Fig. 7). Furthermore, the compactness of the molecule was changed as shown by the SEC experiments (Fig. 3B). Concomitantly, hydrophobic regions of CA-C became accessible to the solvent, as indicated by ANS experiments at 2 μM of protein concentration (Fig. 2C). Then, the acidic transition must involve exposure of hydrophobic patches belonging to each monomer, and thus the species could be similar structural features to that observed in thermal unfolding experiments. The interface region of this monomeric species is not native-like, as concluded by the quenching experiments and NMR spectra at acidic pH. The quenching experiments showed that the tryptophan is as exposed at low pHs (pH 3) as it is when the protein is fully unfolded (Table 1). This solvent exposure was further confirmed by the NMR experiments at pH 3 where at the lower temperature explored (278 K), the chemical shift of the indole proton was close to that observed for random-coil models: 10.27 ppm in CA-C at pH 3 (Fig. 7F) compared to 10.22 ppm in random-coil models (Wüthrich 1986), and 10.62 ppm observed in CA-C at pH 7 (Fig. 7C).

To sum up, our thermal denaturation data suggest that unfolding of the dimeric CA-C occurs via a four-state mechanism. The dissociation step occurs at temperature midpoints about 307 K, to yield a monomeric folded intermediate, which undergoes a conformational transition at temperatures ~ 325 K, and it finally unfolds at ~ 333 K. These series of thermal intermediates agree with the findings of previous chemical denaturation experiments followed by fluorescence and far-UV CD, which monitored exclusively dimer dissociation/monomer rearrangement and monomer unfolding, respectively (Mateu 2002). It is important to stress here that the thermal transition observed by fluorescence cannot be assigned to any particular dissociation and/or conformational rearrangement, since it did not show a clear concentration-independent behavior. Probably, the thermal denaturation fluorescence is reporting several chemical events, as it has been discussed in the chemical denaturation experiments followed by fluorescence (Mateu 2002). Based on the results from FTIR, far-UV CD, anisotropy, and ANS-binding experiments, it can be suggested that thermal fluorescence is measuring dimer dissociation and monomer rearrangement occurring as the temperature changed. Finally, it must be borne in mind that the results obtained here are observed only in equilibrium, and that the sequence of

species suggested might not be observed when the kinetic folding pathway of CA-C is described.

Thermodynamic parameters governing the unfolding and dissociation of the monomeric and dimeric forms of CA-C: Comparison with theoretical methods

The ΔC_p observed upon protein unfolding is largely the result of changes in hydration of groups, which are buried in the native folded state, and become exposed to solvent upon unfolding (Robertson and Murphy 1997). Taking into account the changes in the accessible surface area, ASA, of a monomer of CA-C upon unfolding calculated from its X-ray structure (Gamble et al. 1997; Worthylake et al. 1999) ($\Delta\text{ASA}_{\text{nonpolar}} = 4298 \text{ \AA}^2$ and $\Delta\text{ASA}_{\text{polar}} = 1385 \text{ \AA}^2$), we obtain $\Delta C_p = (0.45 \pm 0.02) (\Delta\text{ASA}_{\text{nonpolar}}) + (-0.26 \pm 0.03) (\Delta\text{ASA}_{\text{polar}}) = 1.6 \pm 0.2 \text{ kcal mol}^{-1} \text{ K}^{-1}$, which is similar to that determined by far-UV CD ($1.8 \pm 0.5 \text{ kcal mol}^{-1} \text{ K}^{-1}$), and by the Pace and Laurents (1989) approach ($1.14 \pm 0.06 \text{ kcal mol}^{-1} \text{ K}^{-1}$). Also, the measured value is close to that determined in other proteins of similar size (Robertson and Murphy 1997).

The use of the above theoretical expression also allows us to estimate the ΔC_p for the CA-C dissociation reaction, assuming that the monomers retain, upon dissociation, most of the native structure they had when they were forming the dimer species. We think this is a reasonable assumption because (1) the far-UV CD spectra of the monomeric and dimeric species are similar (Mateu 2002), (2) there are protons at similar chemical shifts in the NMR spectra of monomeric and dimeric species (J.L. Neira, M. del Álamo, and M.G. Mateu, [Q3](#) unpubl.), and (3) there are only small changes in the width of the FTIR amide I band during the low-temperature transition (Fig. 6A). Based on the X-ray structure, the estimated changes in accessible surface area upon dimer formation per monomer are (Worthylake et al. 1999): $\Delta\text{ASA}_{\text{nonpolar}} = 611 \text{ \AA}^2$ and $\Delta\text{ASA}_{\text{polar}} = 245 \text{ \AA}^2$. These values yield a $\Delta C_p = 211 \pm 10 \text{ cal mol}^{-1} \text{ K}^{-1}$ per monomer. This low value explains why a dissociation transition was not observed by DSC measurements in the protein concentration range from 8 μM to 70 μM (data not shown), and why it was not observed by most of the techniques described here. The CA-C dissociation reaction is governed by a $\Delta H^0 = 74 \pm 8 \text{ kcal (mol of cooperative unit)}^{-1}$ and $\Delta S^0 = 230 \pm 30 \text{ cal K}^{-1} \text{ (mol of cooperative unit)}^{-1}$ at 1 M standard state, which are also small when compared to the corresponding values of other oligomeric proteins (Backmann et al. 1998 and references therein). However, it must be borne in mind that in many small proteins or domains, dissociation and unfolding of dimers occur concomitantly, although it does not happen in many large oligomeric proteins

(Neet and Timm 1994; Jaenicke and Lillie 2000). In CA-C, the values measured (ΔH^0 and ΔS^0) and calculated (ΔC_p) reflect exclusively the dissociation of the dimer.

The free energy of the dissociation of the dimer/conformational rearrangement of the CA-C monomer at 298 K has been obtained from chemical denaturation measurements followed by fluorescence, and it was calculated to be $12.1 \text{ kcal mol}^{-1}$ (at 1 M standard state) (Mateu 2002). Also, the free energy of the specific dissociation step at 298 K, obtained by dilution of the protein in gel filtration analysis was $6.9 \text{ kcal mol}^{-1}$ (at 1 M standard state) (del Álamo et al. 2003). If it is assumed, based on the above calculations, that the ΔC_p for the CA-C dissociation reaction is small, by using the calculated ΔH^0 and ΔS^0 , the free energy of dissociation would be $\Delta G = 5 \pm 8 \text{ kcal mol}^{-1}$ at 298 K. Although the large error in this latter value may prevent a quantitative comparison, these calculations indicate that both the dissociation and a conformational reorganization of the monomer produced may significantly contribute to the combined dissociation–rearrangement transition measured by fluorescence in chemical denaturation experiments (Mateu 2002; del Álamo et al. 2003).

Possible biological implications

Although the role of the pH in HIV maturation in vivo is controversial, the capsids of many viruses are pH-sensitive; for instance, foot-and-mouth disease virus and mengovirus are dissociated at low pH, and the capsids of rhinovirus and poliovirus are conformationally altered at low pH; furthermore, morphological changes around neutral pH have been observed in several viral systems (Johnson 1996). Moreover, the shape of the assemblies formed by CA (Ehrlich et al. 2001; Groß et al. 2000) and the association rates to form such assemblies are pH-dependent (Lanman et al. 2002). In such studies (Groß et al. 2000), a conformational change is suggested to occur between pH 6 and 7, which leads to a more extended conformation of CA at high pHs. The SEC experiments in this work indicate that there were no significant changes in the CA-C compactness at neutral and basic pHs (Fig. 3B). Only the fluorescence and CD measurements indicated a new conformational transition (Figs. 2, 3). This transition was not characterized by a solvent exposure of hydrophobic patches, but there were changes in (1) the secondary structure (Fig. 3A) and (2) the environment of tryptophan and tyrosine residues, which became solvent exposed (Table 1; Fig. 2A,B). Since HIV capsid assembly depends on both CA-C and CA-N, it is suggested, as a working hypothesis, that most of the observed differences at neutral and basic pHs in CA must be due to the CA-N domain.

On the other hand, based on our the thermal

denaturation experiments, it could be argued that at physiological temperatures CA-C should be in the monomeric form. However, it must be taken into account that in vivo, CA-C is a domain of CA, whose associative properties are highly influenced by the CA-N domain (Lanman et al. 2002). The low dimerization affinity/thermal stability of the isolated CA-C dimer might be essential to allow disassembly (del Álamo and Mateu 2005) and/or to mediate a wide range of macromolecular interactions during capsid assembly, as it has been suggested for other protein domains (Tang et al. 1999).

Conclusions

The conformational propensities of the dimeric and monomeric forms of the dimerization domain of HIV-1 have been studied. The results described stress the importance of using several biophysical techniques, which give complementary information, to describe protein unfolding. The thermodynamic parameters of the thermal dissociation and unfolding of the dimeric and monomeric species of CA-C have been determined. The enthalpy and the entropy of dissociation of CA-C are small when compared with those of other small oligomeric proteins, due to the fact that the direct product of dissociation is essentially folded. The monomeric species undergoes, before being fully unfolded, at least one conformational transition, to give a species which could resemble the molten globule-like form of CA-C detected at low pH. Then, the monomer species, whose assembly causes dimer formation, does not show the molten globule-like properties, as it has been observed in other oligomers.

Materials and methods

Materials

Trizma base and acid, phosphate salts, ANS, and NaCl were from Sigma. GdmHCl ultrapure was from ICN Biochemicals. Exact concentrations of GdmHCl were calculated from the refractive index of the solution (Pace and Scholtz 1997). Standard suppliers were used for all the other chemicals. Dialysis tubing was from Spectrapore, with a molecular weight cutoff of 3500 Da. Water was deionized and purified by a Millipore system.

Protein expression and purification

The wild-type CA-C protein was expressed in *Escherichia coli* BL21(DE3) and purified as described (Mateu 2002; del Álamo et al. 2003). Protein was stored in 25 mM sodium phosphate buffer (pH 7.3). Protein stocks were run in SDS-PAGE gels and found to be > 97% pure. Purity was also confirmed by MALDI mass spectrometry analysis (data not shown). Protein concentration was calculated by using the extinction coefficients of model compounds (Pace and Scholtz 1997).

Fluorescence measurements

Fluorescence spectra were collected in a Cary Eclipse spectrofluorometer (Varian) interfaced with a Peltier cell. A 1-cm pathlength quartz cell (Hellma) was used. Experiments were acquired at 298 K, unless indicated otherwise. Protein concentrations used were either 2 μM , 20 μM , or 200 μM , unless otherwise stated.

Steady-state fluorescence measurements

Protein samples were excited at 280 nm and 295 nm between pH 2 and 12 to characterize a possible different behavior of tryptophan and tyrosine residues. The slit widths were 5 nm for the excitation and emission wavelengths. Experiments were recorded between 300 nm and 400 nm. The signal was acquired for 1 sec and the increment of wavelength was set to 1 nm. Blank corrections were made in all spectra.

In the pH-induced unfolding experiments, the pH was measured before and after completion of experiments with an ultrathin Aldrich electrode in a Radiometer (Copenhagen) pH meter. Three-point calibration of the pH meter was performed by using standards from Radiometer. The salts and acids used in buffer preparation were pH 2.0–3.0, phosphoric acid; pH 3.0–4.0, formic acid; pH 4.0–5.5, acetic acid; pH 6.0–7.0, monosodium dihydrogen phosphate; pH 7.5–9.0, Tris acid; pH 9.5–11.0, sodium carbonate; and pH 11.5–13.0, sodium phosphate.

Steady-state ANS binding

ANS binding was detected by collecting fluorescence spectra at different pHs in the presence of 100 μM dye and at a protein concentration of 2 μM . Excitation wavelength was 370 nm, and emission was measured from 430 to 700 nm. Slit widths were 5 nm for excitation and emission wavelengths. Stock solutions of ANS were prepared in water, using a molar extinction coefficient of $6.8 \times 10^3 \text{ M}^{-1} \text{ cm}^{-1}$ at 370 nm (Mann and Matthews 1993).

Fluorescence quenching experiments

Quenching of intrinsic tryptophan and tyrosine fluorescence by acrylamide (Lakowicz 1999) was examined at different pHs. Excitation was carried out at 280 nm (for quenching of tyrosine and tryptophan residues) and 295 nm (for quenching of the tryptophan residue); emission was measured from 300 nm to 400 nm. The slit widths were set to 5 nm for both excitation and emission wavelengths. The dynamic and static quenching constants were obtained by fitting the fluorescence intensity at different wavelengths (in the range from 330 to 340 nm) to the Stern-Volmer equation (Lakowicz 1999):

$$F_0/F = (1 + K_{sv}[X])e^{v[X]}, \quad (1)$$

where K_{sv} is the Stern-Volmer constant for collisional quenching and v is the static quenching constant. The range of quencher concentrations was 0–0.7 M. Protein concentrations were 2 μM in all cases. Control experiments carried out at pH 7 with 20 and 200 μM of CA-C, after correction of inner filter effects, yielded the same quenching constants (data not shown).

Thermal denaturation measurements

Thermal unfolding curves were determined using three different techniques:

1. *Thermal denaturation following the emission fluorescence of tryptophan and tyrosine residues.* Changes were followed by excitation at 280 nm, and acquisition of the emission intensity either at 335 nm or 350 nm (both wavelengths yielded the same unfolding curves; data not shown). The excitation and emission slit widths were set to 5 nm. The scan rate was either 30 K/h or 60 K/h. Both scan rates yielded the same curve (data not shown). Experiments were repeated three times, with new samples.

2. *Thermal denaturation following the emission fluorescence of ANS.* Two different protein concentrations were assayed: 20 μM and 80 μM , with 100 μM and 160 μM of ANS, respectively. No differences were observed in the T_m s obtained at both concentrations. We used two different procedures to ascertain that no scan-rate dependence was observed. First, the temperature was raised manually in the temperature range from 290 K to 360 K; a spectrum was acquired every 3 K after 5 min of equilibration in the Peltier cell. Excitation wavelength was 380 nm, and the emission spectrum was collected in the interval range from 430 nm to 600 nm. Spectra were acquired every 1 nm. The slit widths were set at 5 nm for excitation and emission wavelengths. Second, we used the automatic thermal scan rate of the fluorimeter. Excitation wavelengths were 360 nm, 370 nm, and 380 nm, and emission wavelengths were 480 nm and 520 nm. The same unfolding curves were obtained by collecting the data at both emission wavelengths (data not shown), and similar curves were obtained at any excitation wavelength (data not shown). The excitation and emission slit widths were set to 5 nm; points were acquired every 0.2 K. The scan rate was 60 K/h. No differences were observed between the experiments acquired using the first procedure and the latter procedure; thus, the thermal-unfolding ANS experiments were not scan-rate dependent. In both procedures, experiments were repeated three times, with new samples.

3. *Thermal denaturation following the change in anisotropy.* The steady-state anisotropy, $\langle r \rangle$, is defined by (Lakowicz 1999):

$$\langle r \rangle = \frac{I_V - GI_H}{I_V + 2GI_H}, \quad (2)$$

where I_V and I_H are, respectively, the vertical and the horizontal components of the polarized fluorescence light, and G is a correction factor for the different fluorescence response to the two polarized components. The G factor was typically 0.43 for the emission wavelength of 330, and 0.56 for the emission wavelength of 350 nm. Fluorescence anisotropy was recorded in the L arrangement. Excitation wavelength was set at 280 nm and emission wavelengths were either 350 nm or 330 nm. There were no differences between both emission wavelengths, but those acquired at 350 nm and at 20 μM of CA-C had a lower signal-to-noise ratio. The excitation and emission slit widths were set to 10 nm, and the average time was 5 sec. The temperature of the Peltier cell was increased manually every 3 K and the anisotropy measurements were taken after averaging four measurements for each temperature. Each denaturation was repeated twice.

Circular dichroism and absorbance measurements

Protein concentrations were either 20 μM or 200 μM , unless otherwise stated.

Steady-state experiments

Circular dichroism spectra were collected on a Jasco J810 spectropolarimeter fitted with a thermostated cell holder and interfaced with a Neslab RTE-111 water bath. The instrument was periodically calibrated with (+) 10-camphorsulphonic acid. Isothermal wavelength spectra at different pHs were acquired at a scan speed of 50 nm/min with a response time of 2 sec and averaged over four scans.

The chemical denaturation measurements in the far-UV were performed using 20 μM of CA-C in 25 mM phosphate buffer (pH 7.3), using 0.1-cm pathlength cells (Hellma). The denaturation reaction was fully reversible, as shown by the sigmoidal curves obtained by starting from dilution of protein samples at 7 M GdmHCl (data not shown). Every chemical denaturation experiment was repeated at least three times with new samples. All spectra were corrected by subtracting the proper baseline.

Near-UV spectra at different pHs were acquired in a 0.5-cm pathlength cell (Hellma). The mean residue ellipticity, $[\Theta]$, was calculated as:

$$[\Theta] = \frac{\Theta}{(10lcN)}, \quad (3)$$

where Θ is the measured ellipticity, l is the pathlength cell (in cm), c is the protein concentration (in M), and N is the number of amino acids.

Thermal denaturation experiments

Thermal denaturation experiments were performed at constant heating rates of 60 K/h and 30 K/h and a response time of 8 sec at different pHs. Thermal scans were collected in the far-UV region at 222 nm from 298 K (or 278 K) to 363 K (or 368 K) in 0.1-cm pathlength cells. The reversibility of thermal transitions was tested by recording a new scan after the sample was cooled down to 278 K or 298 K, and comparing to the spectra obtained before heating. The possibility of drifting of the CD spectropolarimeter was tested by running two samples containing buffer, before and after the thermal experiments. No difference was observed between the scans. Every thermal denaturation experiment was repeated at least twice with new samples. In all cases, the samples were transparent and no precipitation was observed, after the heating.

Circular dichroism and absorbance thermal denaturation experiments in the near-UV were simultaneously measured at 280 nm in a Jasco J-710 spectropolarimeter using a protein concentration of 200 μM in a 0.5-cm pathlength cell (Hellma). Every thermal denaturation experiment was repeated at least twice with new samples.

Analysis of the pH, thermal, and chemical denaturation curves, and free energy determination

The average emission intensity in fluorescence spectra, $\langle\lambda\rangle$, was calculated as (Royer 1995):

$$\langle\lambda\rangle = \frac{\sum_i \frac{1}{\lambda_i} I_i}{\sum_i I_i}, \quad (4)$$

where I_i is the fluorescence intensity measured at a wavelength λ_i . This parameter was used to map the changes in the solvent exposure of aromatic residues.

The pH-denaturation experiments were analyzed assuming that both species, protonated and deprotonated, contributed equally to the physical property being observed:

$$X = \frac{(X_a + X_b 10^{n(\text{pH}-\text{p}K_a)})}{(1 + 10^{n(\text{pH}-\text{p}K_a)}), \quad (5)$$

where n is the Hill coefficient, which gives a measurement of the cooperativity of the transition; X is the physical property being observed (ellipticity, $\langle\lambda\rangle$, the maximum wavelength, V_e or the maximum wavenumber of the FTIR band); X_a is the physical property being observed at low pHs; X_b is the physical property observed at high pHs; and $\text{p}K_a$ is the apparent $\text{p}K$ of the titrating group. The apparent $\text{p}K_a$ reported was obtained from three measurements for each technique. In all experimental data in this work, the n value was ~ 1.0 .

Chemical denaturation data were obtained by following the change in the ellipticity at 222 nm in CD experiments. The denaturation data were fitted to the two-state equation:

$$X = \left(X_N + X_D e^{\left(\frac{-\Delta G}{RT}\right)} \right) / \left(1 + e^{\left(\frac{-\Delta G}{RT}\right)} \right), \quad (6)$$

where $X_N = \alpha_N + \beta_N[D]$ and $X_D = \alpha_D + \beta_D[D]$ are the corresponding fractions of the folded and the unfolded states, respectively, for which a linear relationship with denaturant is usually assumed, and X is the physical property (ellipticity) being observed, R is the gas constant, and T is the temperature in K. The chemical denaturation curves were analyzed by using the LEM: $\Delta G = m([D] 50\% - [D])$ (Pace and Laurents 1989; Pace and Scholtz 1997), where ΔG is the free energy of denaturation, $[D]$ is the denaturant concentration, and $[D]_{50\%}$ is that at the midpoint of the transition.

The change in free energy upon temperature in Equation 6 is given by the Gibbs-Helmholtz expression:

$$\Delta G(T) = \Delta H_m \left(1 - \frac{T}{T_m} \right) - \Delta C_p \left[(T_m - T) + T \ln \left(\frac{T}{T_m} \right) \right], \quad (7)$$

where ΔH_m is the van't Hoff enthalpy change. Substitution of this expression in Equation 6, leads to the calculation ΔH_m , T_m , and ΔC_p of the thermal unfolding experiments. In the FTIR experiments, where a concentration-dependent behavior was observed (see Results), the free energy is given by (Backmann et al. 1998):

$$\Delta G(T) = \Delta H_m \left(1 - \frac{T}{T_m} \right) - \Delta C_p \left[(T_m - T) + T \ln \left(\frac{T}{T_m} \right) \right] - RT \ln(2C_t), \quad (8)$$

where C_t is the molar concentration of the protein expressed in dimer equivalents.

Fitting by nonlinear least-squares analysis to the above equations was carried out by using the general curve fit option of Kaleidagraph (Abelbeck software) on a PC computer. A global fitting analysis of the far and the near-UV CD data, at 20 μM and 200 μM of protein concentration, was also carried

out as described elsewhere (Irún et al. 2001). These experiments were chosen because of their better signal-to-noise ratio. The results obtained by both fitting procedures agree within the experimental uncertainty (see Discussion).

Size-exclusion chromatography (SEC)

Analytical SEC experiments were carried out by using an analytical gel filtration Superdex 75 HR 16/60 column (Amersham Biosciences) running on an AKTA FPLC system at 298 K. Flow rates of 1 mL/min were used and aliquots of 100 μ L were loaded into the column after equilibration. The column was equilibrated with four column volumes of elution buffer (25 mM potassium phosphate [pH 7.3], containing 150 mM NaCl to avoid nonspecific interactions with the column). The column was calibrated using the gel filtration low molecular weight calibration kit (Amersham Biosciences). The standards used and their corresponding Stokes radii were ribonuclease A (16.4 Å), chymotrypsinogen (20.9 Å), ovoalbumin (30.5 Å), and bovine serum albumin (35.5 Å) (Hinkle et al. 1999). Protein elution volumes, V_e , were monitored by following absorbance at 280 nm. The CA-C concentration was in all cases either 20 μ M or 200 μ M. The reported V_e is the result of four independent measurements with fresh new samples.

Nuclear magnetic resonance spectroscopy

$1D$ - 1H NMR experiments were carried out in a Bruker DRX-500 spectrometer at 298 K with 32 K data points and using the WATERGATE sequence (Piotto et al. 1992) to eliminate the water signal. Typically, 512 scans were acquired, and the spectral width was 6000 Hz in all cases. CA-C concentration was 200 μ M at pH 7 in 25 mM phosphate buffer in 90% $H_2O/10\%$ 2H_2O or at pH 3 in deuterated acetate buffer 25 mM in 90% $H_2O/10\%$ 2H_2O . No corrections were done for isotope effects. Thermal denaturation experiments at both pHs were also carried out at 0.9 mM of protein concentration. No differences were observed either in the chemical shifts or in their temperature behavior at 0.9 mM when compared with those at 0.2 mM (data not shown).

The spectra were processed using the BRUKER-XWINNMR software working on a PC workstation. An exponential window function and polynomial base line corrections were applied. The final $1D$ - 1H NMR data contained 64,000 data points. 1H chemical shifts were quoted relative to internal TSP.

Thermal denaturation FTIR experiments

Protein was dried in a Speed Vac concentrator (Savant) and dissolved in deuterated buffer at the desired pH. The buffer was composed of 0.1 M NaCl, 0.1 mM ethylenediaminetetraacetate, 0.02% NaN_3 , 10 mM sodium acetate, 10 mM N-(1-morpholino)propane-sulfonic acid, and 10 mM 3-(cyclohexylamino)-1-propane-sulfonic acid. No corrections were done for the isotope effects in the measured pH. Samples of CA-C, at a final concentration between 0.53 mM and 4 mM, were placed amid a pair of CaF_2 windows separated by a 50 μ m-thick spacer in a Harrick demountable cell. Spectra were acquired on a Bruker FTIR-66S instrument equipped with a deuterated triglycine sulphate detector and fitted with a water bath. The cell container was continuously filled with dry air. Usually 50 scans per sample were taken, averaged, apodized with a Happ-Genzel function, and Fourier transformed to give a final resolution of 2 cm^{-1} . The

contributions of buffer spectra were subtracted, and the resulting spectra were used for analysis. The scanning rate was 50 K/h, and spectra were acquired every 2.5 K. At pH 7, the range of explored temperatures was 280–370 K, and the sigmoidal curve was fully reversible only for the low-temperature transition (see Results). The thermal experiments, carried out to detect the low-temperature transition, were acquired up to temperatures not higher than 325 K. Hence, the first thermal denaturation was fully reversible at most of the pHs (see Results). Protein concentration was 1 mM. The thermal transitions were analysed by taking the 3/4 of the height width of the band; similar results were obtained by measuring the half of the width of the height band.

Acknowledgments

This work was supported by grants from the Ministerio de Sanidad y Consumo (MSC) (FIS 01/0004-02) and Generalitat Valenciana (GV04B-402) and by an institutional grant from the Urbasa to the Centro de Biología Molecular y Celular to J.L.N.; from the Spanish Ministerio de Educación y Ciencia (MEC) (BFU2004-01411) to J.S.; and from the MSC (FIS 01/0004-01) and Ministerio de Ciencia y Tecnología (BIO2003-04445) to M.G.M.; and by an institutional grant from the Fundación Ramón Areces to the Centro de Biología Molecular “Severo Ochoa.” M.C.L.-M. was the recipient of a predoctoral fellowship from MSC. F.N.B. and M.B. were recipients of two predoctoral fellowships from the Ministerio de Educación, Cultura y Deporte. We thank Dr. Claudio Cavasotto for the calculations of the solvent-accessible surface area of folded CA-C. We deeply thank May García, María del Carmen Fuster, Javier Casanova, María T. Garzón, Helena López, and Eva Martínez for excellent technical assistance. We thank the two anonymous reviewers by their suggestions, critical thoughts, and improvements to the manuscript.

References

- Anderson, D.E., Becktel, W.J., and Dahlquist, F.W. 1990. pH-induced denaturation of proteins: A single salt bridge contributes 3–5 kcal/mol to the free energy of folding of T4 lysozyme. *Biochemistry* **29**: 2403–2408.
- Backmann, J., Schäfer, G., Wyns, L., and Bönisch, H. 1998. Thermodynamics and kinetics of unfolding of the thermostable trimeric adenylate kinase from the archeon *Sulfolobus acidocaldarius*. *J. Mol. Biol.* **284**: 817–833.
- Barth, A. and Zscherp, C. 2002. What vibrations tell us about proteins. *Q. Rev. Biophys.* **35**: 369–430.
- Borsetti, A., Öhagen, Å., and Göttinger, H.G. 1998. The C-terminal half of the human immunodeficiency virus type 1 Gag precursor is sufficient for efficient particle assembly. *J. Virol.* **72**: 9313–9317.
- Campos, L.A., Bueno, M., López-Llano, J., Jiménez, M.A., and Sancho, J. 2004. Structure of stable protein folding intermediates by equilibrium ϕ -analysis: The apoflavodoxin thermal intermediate. *J. Mol. Biol.* **344**: 239–255.
- Cantor, C.R. and Schimmel, P.R. 1980. *Biophysical chemistry*. W.H. Freeman, New York.
- Chen, H.A., Pfuhl, M., McAlister, M.S.B., and Driscoll, P.C. 2000. Determination of pK_a values of carboxyl groups in the N-terminal domain of rat CD2: Anomalous pK_a of a glutamate on the ligand binding surface. *Biochemistry* **39**: 6814–6824.
- Clarke, J. and Fersht, A.R. 1993. Engineered disulfide bonds as probes of the folding pathway of barnase: Increasing the stability of proteins against the rate of denaturation. *Biochemistry* **32**: 4322–4329.
- del Álamo, M. and Mateu, M.G. 2005. Electrostatic repulsion, compensatory mutations, and long-range non-additive effects at the dimerization interface of the HIV capsid protein. *J. Mol. Biol.* **345**: 893–906.

Q4

- del Álamo, M., Neira, J.L., and Mateu, M.G. 2003. Thermodynamic dissection of a low-affinity protein-protein interface involved in human immunodeficiency virus assembly. *J. Biol. Chem.* **278**: 27923–27929.
- Ding, K., Louis, J.M., and Gronenberg, A.M. 2004. Insights into conformation and dynamics of protein GB1 during folding and unfolding by NMR. *J. Mol. Biol.* **335**: 1299–1307.
- Dorfman, T., Bukovsky, A., Öhagen, Å., Hoglund, S., and Göttinger, H.G. 1994. Functional domains of the capsid protein of human immunodeficiency virus. *J. Virol.* **68**: 8180–8187.
- Ehrlich, L.S., Liu, T., Scarlata, S., Chu, B., and Carter, C.A. 2001. HIV-1 capsid proteins forms spherical (immature-like) and tubular (mature-like) particles *in vitro*: Structure switching by pH-induced conformational changes. *Biophys. J.* **81**: 586–594.
- Forshey, B.M., von Schwedler, U.K., Sundquist, W.I., and Aiken, C. 2002. Formation of a human immunodeficiency virus type 1 core of optimal stability is crucial for viral replication. *J. Virol.* **76**: 5667–5677.
- Forsyth, W.R. and Robertson, A.D. 2000. Insensitivity of perturbed carboxyl pK_a values in the ovomucoid third domain to charge replacement at a neighboring residue. *Biochemistry* **39**: 8067–8072.
- Fuller, T.W., Wilk, T., Gowen, B.E., Kräuslich, H.G., and Vogt, V.M. 1997. Cryo-electron microscopy reveals ordered domains in the immature HIV-1 particle. *Curr. Biol.* **7**: 729–738.
- Gamble, T.R., Vajdos, F.F., Yoo, S., Worthylake, D.K., Houseweart, M., Sundquist, W.I., and Hill, C.P. 1996. Crystal structure of human cyclophilin A bound to the amino terminal domain of HIV-1 capsid. *Cell* **87**: 1285–1294.
- Gamble, T.R., Yoo, S., Vajdos, F.F., von Schwedler, U.K., Worthylake, D.K., Wang, H., McCutcheon, J.P., Sundquist, W.I., and Hill, C.P. 1997. Structure of the carboxyl-terminal dimerization domain of the HIV-1 capsid protein. *Science* **278**: 849–853.
- Ganser, B.K., Li, S., Kliskho, V.Y., Finch, J.T., and Sundquist, W.I. 1999. Assembly and analysis of conical models for the HIV-1 core. *Science* **283**: 80–83.
- Gelderblom, H.R. 1991. Assembly and morphology of HIV: Potential effect of structure on viral function. *AIDS* **5**: 617–638.
- Gitti, R.K., Lee, B.M., Walker, J., Summers, M.F., Yoo, S., and Sundquist, W.I. 1996. Structure of the amino-terminal core domain of the HIV-1 capsid protein. *Science* **273**: 231–235.
- Groß, I., Hohenberg, H., and Kräuslich, H.G. 1997. *In vitro* assembly properties of purified bacterially expressed capsid proteins of human immunodeficiency virus. *Eur. J. Biochem.* **249**: 592–600.
- Groß, I., Hohenberg, H., Huckangel, C., and Kräuslich, H.G. 1998. N-terminal extension of human immunodeficiency virus capsid protein converts the *in vitro* assembly phenotype from tubular to spherical particles. *J. Virol.* **72**: 4798–4810.
- Groß, I., Hohenberg, H., Wilk, T., Wieggers, K., Grättinger, M., Müller, B., Fuller, S., and Kräuslich, H.G. 2000. A conformational switch controlling HIV-1 morphogenesis. *EMBO J.* **19**: 103–113.
- Hinkle, A., Goranson, A., Butters, C.A., and Tobacman, L.S. 1999. Roles for the troponin tail domain in thin filament assembly and regulation. A deletional study of cardiac troponin T. *J. Biol. Chem.* **274**: 7157–7164.
- Honig, B. and Yang, A.S. 1995. Free energy balance in protein folding. *Adv. Protein Chem.* **46**: 27–58.
- Irún, M.P., García-Mira, M.M., Sánchez-Ruiz, J.M., and Sancho, J. 2001. Native hydrogen bonds in a molten globule: The apoflavodoxin thermal intermediate. *J. Mol. Biol.* **306**: 877–888.
- Jaenicke, R. and Lillie, H. 2000. Folding and association of oligomeric and multimeric proteins. *Adv. Protein Chem.* **53**: 329–401.
- Johnson, J.E. 1996. Functional implications of protein-protein interaction in icosahedral viruses. *Proc. Natl. Acad. Sci.* **93**: 27–33.
- Kelly, S.M. and Price, N.C. 2000. The use of circular dichroism in the investigation of protein structure and function. *Curr. Protein Pept. Sci.* **1**: 349–384.
- Lakowicz, J.R. 1999. *Principles of fluorescence spectroscopy*, 2nd ed. Plenum Press, New York.
- Lanman, J., Sexton, J., Sakalian, M., and Prevelige Jr., P.E. 2002. Kinetic analysis of the role of intersubunit interactions in human immunodeficiency virus type 1 capsid protein assembly *in vitro*. *J. Virol.* **76**: 6900–6908.
- Li, S., Hill, C.P., Sundquist, W.I., and Finch, J.T. 2000. Image reconstructions of helical assemblies of the HIV-1 CA protein. *Nature* **407**: 409–413.
- Luo, J., Iwakura, M., and Matthews, C.R. 1995. Detection of an intermediate in the thermal unfolding of a cysteine-free form of dihydrofolate from *Escherichia coli*. *Biochemistry* **34**: 10669–10675.
- Luo, Y., Kay, M., and Baldwin, R.L. 1997. Co-operativity of folding of apomyoglobin pH 4 intermediate studied by glycine and proline mutations. *Nat. Struct. Biol.* **4**: 925–930.
- Mach, H., Volkin, D.B., Burke, C.J., and Middaugh, C.R. 1995. Ultraviolet absorption spectroscopy. In *Protein stability and folding (Methods in molecular biology, Vol. 40)* (ed. B.A. Shirley), pp. 91–114. Humana Press, Towota, NJ.
- Maldonado, S., Jiménez, M.A., Langdon, G.M., and Sancho, J. 1998. Cooperative stabilization of a molten-globule apoflavodoxin fragment. *Biochemistry* **37**: 10589–10596.
- Mann, C.J. and 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**: 5282–5290.
- Marky, L.A. and Breslauer, K.J. 1987. Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves. *Biopolymers* **26**: 1601–1620.
- Mateu, M.G. 2002. Conformational stability of dimeric and monomeric forms of the C-terminal domain of human immunodeficiency virus. *J. Mol. Biol.* **318**: 519–531.
- Misselwitz, R., Hausdorf, G., Welfe, K., Höhne, W.E., and Welfe, H. 1995. Conformation and stability of recombinant HIV-1 capsid protein p24 (rp24). *Biochem. Biophys. Acta* **1250**: 9–18.
- Momany, C., Kovari, L.C., Prongay, A.J., Keller, W., Gitti, R.K., Lee, B.M., Gorbalenya, A.E., Tong, L., McClure, J., Ehrlich, L.S., et al. 1996. Crystal structure of the dimeric HIV-1 capsid protein. *Nat. Struct. Biol.* **3**: 763–770.
- Neet, K.E. and Timm, D.E. 1994. Conformational stability of dimeric proteins: Quantitative studies by equilibrium denaturation. *Protein Sci.* **3**: 2167–2174.
- Pace, C.N. and Laurents, D.V. 1989. A new method for determining the heat capacity change for protein folding. *Biochemistry* **28**: 2520–2525.
- Pace, C.N. and Scholtz, J.M. 1997. Measuring the conformational stability of a protein. In *Protein structure* (ed. T.E. Creighton), 2nd ed., pp. 253–259. Oxford University Press, Oxford.
- Piotto, M., Saudek, V., and Sklenar, V. 1992. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR* **2**: 661–665.
- Ptitsyn, O.B. 1995. Molten globule and protein folding. *Adv. Protein Chem.* **47**: 83–229.
- Robertson, A.D. and Murphy, K.P. 1997. Protein structure and the energetics of protein stability. *Chem. Rev.* **97**: 1251–1267.
- Rosé, S., Hensley, P., O'Shanessy, D.J., Culp, J., Debouck, C., and Chaiken, I. 1992. Characterization of HIV-1 p24 self-association using analytical chromatography. *Proteins* **13**: 112–119.
- Royer, C.A. 1995. Fluorescence spectroscopy. *Protein stability and folding (Methods in molecular biology Vol. 40)* (ed. B.A. Shirley) pp. 65–89. Humana Press, Towota, NJ.
- Semisotnov, G.V., Rodionova, N.A., Razgulyaev, O.I., Uversky, V.N., Gripas, A.F., and Gimanshin, R.I. 1991. Study of the “molten globule” intermediate state in protein folding by a hydrophobic fluorescent probe. *Biopolymers* **31**: 119–128.
- Steif, C., Weber, P., Hinz, H.J., Flossdorf, J., Cesareni, G., and Kokkinidis, M. 1993. Subunit interactions provide a significant contribution to the stability of the dimeric four- α -helical bundle protein ROP. *Biochemistry* **32**: 3867–3876.
- Stryer, L. 1965. The interaction of a naphthalene sulfonate dye with apomyoglobin and apohemoglobin: A fluorescent probe of non-polar binding site. *J. Mol. Biol.* **13**: 482–495.
- Tang, K.S., Guralnick, B.J., Wong, W.K., Fersht, A.R., and Itzhaki, L.S. 1999. Stability and folding of the tumor suppressor protein p16. *J. Mol. Biol.* **285**: 1869–1886.
- Tang, S., Murakami, T., Agresta, B.E., Campbell, S., Freed, E.O., and Levin, J.G. 2001. Human immunodeficiency virus type 1 N-terminal capsid mutants that exhibit aberrant core morphology are blocked in initiation of reverse transcription in infected cells. *J. Virol.* **75**: 9357–9366.
- von Schwedler, U.K., Stemmler, T.L., Kliskho, V.Y., Li, S., Albertine, K.H., Davis, D.R., and Sundquist, W.I. 1998. Proteolytic refolding of the HIV-1 capsid protein amino-terminus facilitates viral core assembly. *EMBO J.* **17**: 1555–1568.
- Wang, C.T. and Barklis, E. 1993. Assembly, processing and infectivity of human immunodeficiency virus type 1 Gag mutants. *J. Virol.* **67**: 4264–4273.

- Welker, R., Hohenberg, H., Tessmer, U., Huckhagel, C., and Kräuslich, H.G. 2000. Biochemical and structural analysis of isolated matured cores of human immunodeficiency virus type I. *J. Virol.* **74**: 1168–1177.
- Wills, J.W. and Craven, C. 1991. Form, function and use of retroviral gag proteins. *AIDS* **5**: 639–654.
- Woody, R.W. 1995. Circular dichroism. *Methods Enzymol.* **246**: 34–71.
- Worthylake, D.K., Wang, H., Yoo, S., Sundquist, W.I., and Hill, C.P. 1999. Structures of the HIV-1 capsid protein dimerization domain at 2.6 Å resolution. *Acta Crystallogr. D Acta Crystallogr.* **55**: 85–92.
- Wüthrich, K. 1986. *NMR of proteins and nucleic acids*. John Wiley & Sons, New York.
- Zweifel, M.E. and Barrick, D. 2002. Relationships between the temperature dependence of solvent denaturation and the denaturant dependence of protein stability curves. *Biophys. Chem.* **101–102**: 221–237.

