



UNIVERSIDAD DE BUENOS AIRES
Facultad de Ciencias Exactas y Naturales

**“Bases moleculares de interacciones virus-hospedador:
análisis biofísico y estructural de la interacción de la
proteína E1A de Adenovirus con su blanco específico
Retinoblastoma”**

Tesis presentada para optar al título de Doctor de la Universidad de Buenos Aires en el área
Ciencias Biológicas.

Bqco. Nicolás Sebastián González Foutel

Directora de Tesis: Dra. Lucía Beatriz Chemes

Consejero de estudios: Dr. Pablo Wappner

Lugar de trabajo: Laboratorio de Estructura, Plasticidad y Función de Proteínas. IIB-INTECH-UNSAM, San Martín, Argentina. Laboratorio de Estructura, Función e Ingeniería de Proteínas. Fundación Instituto Leloir e IIBBA-CONICET, Buenos Aires, Argentina.

Buenos Aires, 2019

AGRADECIMIENTOS

En primer lugar quisiera agradecer a mi directora Lucía, que me dio la oportunidad de trabajar y sobre todo *aprender* a su lado. Por el rigor y las libertades otorgadas. Por ser la fuerza motriz de este proyecto y haber estado siempre presente.

Mis agradecimientos divididos entre los dos Institutos que me dieron lugar durante mi doctorado,

En el Instituto Leloir: a Gonzalo, quien me permitió trabajar en su laboratorio durante los primeros años de la tesis. A mis ex-compañeros del labo 209, con los que compartí desde el inicio o sólo temporalmente: Ivi, Gabyn, Gabi, Sil, Leo, Sebas, Damián, Mariano, Juli, Julia, Marikena, Esteban, Mariana, por la ayuda y porque de todos ellos he aprendido algo.

A mis compas de otros labos, en especial a La Sole por esas charlitas de pasillos. A Sole del HPLC, que me prestó mucha ayuda; a la gente de biblioteca Cardini, siempre tan predispuesta y a Diego del droguero, un genio!

A los miembros de mi Comisión de Seguimiento de Tesis: Julio Caramelo, Andrea Gamarnik y Jimena Rinaldi, por sus consejos y apoyo.

En el IIB: a Lu, que con su llegada “pobló” la 2M5. Porque con su complicidad me animó y ayudó este último tiempo.

A Lauchi y los bochincheros del IIB: Vic, Marianela, Agos, Paca, Brian, Martín, Adriano, Mile y Laura B, que me hicieron un lugarcito en las mesas redondas cuando me sumé al instituto. Y mis compas: Ceci, Paula, Ema, Mara, Flor, Maxi y Ariel por la distensión de esas charlas “enfrentadas” en estos últimos meses.

A la gente de Administración, sobre todo a Andrea, siempre servicial!

A mi *alma mater* la UNNE, por darme una educación pública y de calidad. A mis compañeros y amigos que crucé en la facu, en especial Ingrid y Lara, con quiénes compartí horas de vida y estudio, y escucharon y apoyaron todos mis planes. A la FCEN-UBA, por la educación de posgrado y la inspiración de sus muros y personas que la componen.

A JP, mi primer “mentor” y consejero en este camino, ahora amigo, quien me convenció de estudiar bioquímica y me prestó su oído muchas veces. A los miembros de “TQ” y “los porotos”, mi primer grupo de investigación, especialmente Guada, Pinty y Belén, por su empujón inicial para salir a buscar “oportunidades más grandes”.

A mis amigos, mi familia por elección, los de antes y de ahora: Elena, Favio, Lucio y Ana. A Sebas y a mi caramelera favorita: Costy, Cin, Tutty, Ailen, Jose, Gabi, por esos mensajes (incontables -de verdad-) y llamadas de alientos! A Emi y Tomás, a los “correntinos” (Santi, Joaquín, Ana, Noe y Norberto) y los “tucumanos” (Ale, Mera, el Colo), a quiénes por unas circunstancias u otras ya no están a mi lado, porque también compartieron mucho conmigo y desde su lugar me regalaron sus fuerzas para continuar con el doc.

A mis amigos y compañeros que conocí afuera, por y gracias al doctorado: Cameron y Wade y toda la gente del labo en Tampa. A Maty y la gente del labo en Budapest. A Mark, Tibor y Kris, que hicieron de mi estadía en Hungría, una experiencia inolvidable.

A mi familia, que lo es todo! A mi mamá, que no pasó día que no haya estado presente con su constante aliento, ayuda y rezos (y llamaditos semanales). A mi viejo, por enseñar con su silencio y humildad. A mi hermano Matías. A mi hermana Laura y mi sobrinito Manolo. Y a mi hermana y coequiper viajera Luciana. Porque me brindan su apoyo incondicional, y estuvieron y están a mi lado de una forma u otra en cada momento de mi vida. ***Esta tesis se la dedico a ustedes!***

A Dios y la vida, por haberme puesto toda esta gente en mi camino.



“En la vida, no existe nada que temer, solo cosas que comprender.”

Marie Curie

Bases moleculares de interacciones virus-hospedador: análisis biofísico y estructural de la interacción de la proteína viral E1A de Adenovirus con su blanco específico Retinoblastoma.

Resumen

Virus patogénicos como Adenovirus, Papilomavirus y Poliomavirus poseen proteínas capaces de interactuar con el supresor de tumores Retinoblastoma (pRb), controlador central del ciclo celular eucariótico. Dichas interacciones se establecen en sitios utilizados por pRb para unir blancos endógenos como el surco A/B donde se une el factor de transcripción E2F y el surco LxCxE donde se unen proteínas que poseen el motivo LxCxE. Estos virus hacen mímica de *motivos lineales de interacción* altamente conservados, y generalmente presentes en regiones intrínsecamente desordenadas, para interferir efectivamente con la red celular y tomar el control de la célula huésped conduciendo en algunos casos a transformación celular y oncogénesis. La oncoproteína viral E1A del Adenovirus (AdE1A) es una proteína intrínsecamente desordenada (IDP) y se une a pRb a través de dos *motivos lineales* necesarios para desplazar a E2F, el *motivo E2F* y el *motivo LxCxE*, que se encuentran separados por una región de 70 residuos o "*linker*". En este trabajo, sugerimos que ciertas características de esta proteína desempeñan un papel esencial en los mecanismos de desregulación de pRb. Dado el limitado conocimiento mecanístico y estructural de alta resolución con respecto a AdE1A, su interacción con pRb y la contribución de la región "*linker*" en esta interacción, nos propusimos caracterizarlos para revelar información que servirá para comprender mejor las interacciones virus-hospedador. En primer lugar, desarrollamos un protocolo de expresión recombinante de AdE1A y mutantes individuales de motivos, con rendimientos de 9 mg/L y pureza superior al 95%. Experimentos biofísicos evidenciaron que AdE1A es un monómero intrínsecamente desordenado con radio hidrodinámico extendido y migración anómala en gel. En segundo lugar, ensayos de interacción muestran que AdE1A se une a pRb con estequiometría 1:1 y una muy alta afinidad ($K_D = 24$ pM). Aunque los sitios individuales unen a pRb con menor afinidad ($\approx K_D = 110$ nM) que la E2F celular ($K_D = 1$ nM), el modelado del sistema como un arreglo de dos motivos unidos por un "*linker*" de longitud óptima y altamente flexible permite explicar la potenciación de su afinidad global y la efectividad de la proteína AdE1A en el desplazamiento de E2F. En tercer lugar, estudios estructurales de RMN permitieron la asignación completa de la cadena carbonada de AdE1A y su caracterización a nivel de estructura secundaria y propensiones conformacionales evidenciaron su naturaleza desordenada y flexible. AdE1A se une a pRb a través de dos sitios independientes conformados por residuos centrales de los *motivos E2F* y *LxCxE* modulados por residuos flanqueantes con función complementaria, más interacciones débiles en la región "*linker*". El análisis de la región "*linker*", sugiere que existe un fino balance entre repulsión de cargas y contactos hidrofóbicos débiles con pRb en regiones de interacción secundarias, las cuáles se encuentran altamente conservadas a nivel evolutivo y que contribuirían a su capacidad de interacción con múltiples blancos celulares. Los resultados del presente trabajo contribuyen a comprender las bases moleculares de las interacciones entre proteínas de virus y hospedador y los mecanismos mediante los cuales los virus toman control de la célula huésped, lo cual allana el camino para el desarrollo de estrategias terapéuticas contra el cáncer de etiología viral.

Palabras clave: interacciones virus-hospedador, motivos lineales, proteína intrínsecamente desordenada, AdE1A, Retinoblastoma.

Molecular basis of virus-host interactions: biophysical and structural analysis of the interaction between the Adenovirus E1A protein and its specific target Retinoblastoma protein.

Abstract

Pathogenic viruses such as Adenovirus, Papillomavirus and Polyomavirus harbor proteins that are able to interact with the Retinoblastoma tumor suppressor protein (pRb), central regulator of the eukaryotic cell cycle. These interactions are established at sites used by pRb to interact with endogenous targets, for example the A/B groove where the E2F transcription factor binds and the LxCxE groove where proteins harboring the *LxCxE motif* bind. These viruses mimic highly conserved *linear interaction motifs* generally present in intrinsically disordered regions to effectively interfere with the cellular network and take control of the host cell, leading in some cases to cellular transformation and oncogenesis. The E1A viral oncoprotein from Adenovirus (AdE1A) is an intrinsically disordered protein (IDP) and binds to pRb through two linear motifs necessary to displace E2F, the *E2F motif* and the *LxCxE motif*, which are separated by a 70-residues region or "*linker*". In this work, we suggest that certain features of this protein play an essential role in the viral-mediated pRb dysregulation mechanism. Given the limited high-resolution mechanistic and structural knowledge regarding AdE1A, its interaction with pRb and the contribution of the "*linker*" region in this interaction, we set out to characterize them to reveal information that provides a better understanding of virus-host interactions. First, we developed a recombinant expression protocol of AdE1A and single motif mutants, with yields of 9 mg/L and purity over 95%. Biophysical experiments showed that AdE1A is an intrinsically disordered monomer with extended hydrodynamic radius and anomalous gel migration, characteristic of an IDP. Second, interaction assays show that AdE1A binds to pRb with 1:1 stoichiometry and a very high affinity ($K_D = 24$ pM). Although individual sites bind to pRb with lower affinity ($\approx K_D = 110$ nM) than cellular E2F ($K_D = 1$ nM), the modeling of the system as an arrangement of two motifs tethered by a highly flexible "*linker*" of optimal length allows to explain its global affinity enhancement and the effectiveness of the AdE1A protein in the displacement of E2F. Third, high resolution structural NMR studies allowed the complete assignment of the AdE1A backbone and its characterization at secondary structure and backbone conformational propensity level, revealing its disordered and flexible nature. AdE1A binds to pRb through two independent sites formed by key residues in the *E2F* and *LxCxE motifs* modulated by flanking residues, plus weak interactions in the "*linker*" region. The analysis of the "*linker*" region suggests that evolution has kept a fine balance between charge repulsion and weak hydrophobic contacts with pRb at secondary interaction regions, which are highly conserved and would contribute to E1A's ability to function as a "hub protein", interacting with multiple cellular targets. The results of this work contribute to understanding the molecular basis of virus-host protein-protein interactions and the mechanisms by which viruses take control of the host cell, which paves the way for the development of therapeutic strategies against cancer with viral etiology.

Key words: virus-host interactions, linear motifs, Intrinsically disordered protein, AdE1A, Retinoblastoma protein.

PUBLICACIONES

Parte del trabajo desarrollado durante la presente tesis doctoral (Capítulo III) está contenido en la siguiente publicación:

Palopoli, N., González Foutel, N. S., Gibson, T. J., & Chemes, L. B. (2018). **Short linear motif core and flanking regions modulate retinoblastoma protein binding affinity and specificity.** *Protein Engineering, Design and Selection*, 31(3), 69–77. doi:10.1093/protein/gzx068

Se planea la publicación en el corto plazo un trabajo que integre toda esta tesis.

González Foutel, N.S. *et. al.* (2019). **An effective concentration "molecular ambush" mechanism underlies viral hijack by the intrinsically disordered E1A oncoprotein.** *Manuscrito en preparación.*

INDICE

	Pág
ABREVIATURAS	1
INTRODUCCION	3
A. Interacciones Virus-Hospedador	5
B. El desorden intrínseco en proteínas virales	7
C. Las proteínas con desorden intrínseco existen y son abundantes en los organismos	8
D. Proteínas Intrínsecamente Desordenadas (IDPs) y sus características	9
E. Adenovirus	25
F. La oncoproteína viral AdE1A	28
G. La relación entre AdE1A y las proteínas pRb, p107 y p130	33
H. La proteína supresora de tumores Retinoblastoma y las proteínas p107/p130	33
I. Desregulación de la función de pRb por acción de oncoproteínas virales: AdE1A	36
HIPOTESIS Y OBJETIVOS	41
CAPITULO I: Clonado, expresión y caracterización bioquímica y biofísica de AdE1A y mutantes de motivo.	47
OBJETIVO I.1. Generación de AdE1A.WT y mutantes de motivo	51
I.1.A. Biología molecular del gen Ad5E1A	51
I.1.B. Expresión y Purificación de Proteínas	58
OBJETIVO I.2. Caracterización de las propiedades estructurales-conformacionales de AdE1A	73
I.2.A. Análisis del comportamiento electroforético de AdE1A	73
I.2.B. Estudio de AdE1A por espectrometría de masa MALDI-TOF y análisis tróptico	74
I.2.C. Análisis de las propiedades hidrodinámicas y conformacionales de AdE1A	79
I.2.D. Evaluación del estados de oxidación de AdE1A.	85
I.2.E. Análisis del contenido de estructura secundaria de AdE1A.WT por Dicroísmo Circular	87
I.2.F. Desnaturalización térmica de AdE1A.WT	94
OBJETIVO I.3. Obtención de reactivos para espectroscopía de fluorescencia	95
I.3.A. Péptidos y proteínas marcados con FITC. Protocolo de marcación y rendimiento para AdE1A.WT	95

I.3.B. Purificación e identificación de AdE1A.WT-FITC	97
CONCLUSIONES I	100
CAPITULO II: Análisis del mecanismo de interacción entre AdE1A y su blanco RbAB	103
OBJETIVO II.1. Revisión de datos experimentales de interacciones a miembros de la familia de pRb mediadas por los motivos lineales E2F y LxCxE.	107
OBJETIVO II.2. Mecanismo de interacción entre AdE1A y RbAB	117
II.2.A. Estudio de las propiedades hidrodinámicas del complejo [RbAB:AdE1A] por SEC-SLS. Primer acercamiento al modelo de interacción.	117
II.2.B. Caracterización espectroscópica por Dicroísmo Circular del complejo RbAB:AdE1A. Evaluando procesos de plegamiento-interacción en la formación del complejo.	122
II.2.C. Análisis termodinámico de la interacción RbAB-AdE1A	125
Conclusiones Parciales II.2.C.	145
II.2.D. Análisis de posible efecto cooperativo alostérico “en trans” entre los motivos de interacción de AdE1A con RbAB.	145
Conclusiones Parciales II.2.D.	149
II.2.E. Evaluación de la capacidad efectiva de AdE1A y fragmentos para desplazar eficientemente E2F2 de su complejo con RbAB. Ensayos de Competición por EF	149
CONCLUSIONES II	153
CAPITULO III: Estudio estructural de alta resolución de AdE1A libre y en complejo con RbAB por Resonancia Magnética Nuclear.	155
OBJETIVO III.1. Caracterización estructural de AdE1A libre	159
III.1.A. Evidencias de desorden intrínseco en AdE1A. Asignación de resonancias de AdE1A.WT, AdE1A- Δ E y AdE1A- Δ L.	159
III.1.B. Identificación de elementos de estructura secundaria en AdE1A	166
III.1.C. Dinámica conformacional de AdE1A en estado libre. Experimentos de relajación.	176
III.1.D. Efecto de las mutaciones sobre la estructura y conformación nativa de AdE1A	183
OBJETIVO III.2. Análisis estructural del complejo [RbAB:AdE1A.WT]	186
III.2.A. Análisis de la interacción entre AdE1A y RbAB. Experimentos TROSY para proteínas libres y en complejo con RbAB.	186

III.2.B. Determinación de los regímenes de intercambio de los residuos que interaccionan con RbAB. Experimento de titulación del complejo [RbAB:AdE1A.WT]	196
CONCLUSIONES III	207
CAPITULO IV: Modelado de la interacción entre AdE1A y su blanco específico RbAB	209
IV.1. Modelado teórico híbrido de los complejos entre AdE1A.WT y mutantes con RbAB	213
IV.2. Análisis de SAXS del complejo [RbAB:AdE1A.WT]	216
IV.3. Modelado “WLC” de la interacción entre AdE1A.WT y RbAB	219
CONCLUSIONES IV	223
DISCUSION Y CONCLUSION GENERAL	225
MATERIALES Y METODOS	235
ANEXO	269

ABREVIATURAS *relevantes*

Ad5E1A	Oncoproteína viral E1A de Adenovirus serotipo 5. Nombrada para hacer diferencia de otros serotipos.
AdE1A	Oncoproteína viral E1A (del inglés <i>early</i>) de Adenovirus. Nomenclatura genérica.
AdE1A.WT	Secuencia nativa de la oncoproteína viral E1A de Adenovirus. Fragmento 36 a 146.
AdE1A-ΔE	Secuencia mutante de la oncoproteína viral E1A de Adenovirus, carente del motivo de interacción E2F.
AdE1A-ΔL	Secuencia mutante de la oncoproteína viral E1A de Adenovirus, carente del motivo de interacción LxCxE.
CD	Dicroísmo Circular, por sus siglas en inglés.
CR(1,2,3,4)	Región conservada, dominio funcional de la proteína AdE1A.
DLS	Dispersión de luz dinámica, por sus siglas en inglés.
E2F	<i>Motivo</i> de interacción viral; <i>motivo</i> de interacción celular; <i>sitio</i> complementario al motivo en la superficie de RbAB.
E2F2	Factor de transcripción celular. Nombrada particularmente en contexto de experimentos.
EF	Espectroscopía de Fluorescencia .
HAdV5	Adenovirus humano serotipo 5.
HPLC	Cromatografía líquida de alto rendimiento , por sus siglas en inglés.
HPV	Virus del Papiloma humano.
HSQC	Coherencia Cuántica Simple Heteronuclear, por sus siglas en inglés.
IDP	Proteína intrínsecamente desordenada, por sus siglas en inglés.
IDR	Región intrínsecamente desordenado, por sus siglas en inglés.
ITC	Calorimetría de Titulación isotérmica, por sus siglas en inglés.
LxCxE	<i>Motivo</i> de interacción con la secuencia LxCxE; <i>surco</i> complementario al motivo en la superficie de RbAB.
MALDI-TOF	Ionización/desorción láser asistida por matriz- Tiempo de Vuelo, por sus siglas en inglés del inglés.
MG	Estado de Glóbulo fundido, por sus siglas en inglés.
MoRF	Elemento de reconocimiento molecular, por sus siglas en inglés.
PAGE	Electroforesis en gel de Poliacrilamida , por sus siglas en inglés.
PM	Peso molecular.
PMG	Estado de pre-glóbulo fundido, por sus siglas en inglés.
PPI	Interacciones proteína-proteína.
PPII	Estructura Poliprolina tipo II.
pRb	Supresor de tumores Retinoblastoma. Nomenclatura genérica.
R_H	Radio hidrodinámico.
RbAB	Dominio central del supresor de tumores Retinoblastoma. Nombrada particularmente en contexto de experimentos
RMN	Resonancia magnética nuclear, por sus siglas en inglés.
SAXS	Dispersión de Rayos X de ángulo pequeño, por sus siglas en inglés.
SEC	Cromatografía de exclusión molecular, por sus siglas en inglés.
SLIM	Motivo corto lineal de interacción, por sus siglas en inglés.
SLS	Dispersión de luz estática, por sus siglas en inglés.
TRAM-CBP	Motivo de interacción de unión a TRAM o dominios de CBP/p300
TROSY	Espectroscopía de Relajación Transversal Optimizada, por sus siglas en inglés.
WLC	modelo de polímero flexible o cadena entrópica (del inglés <i>Worm Like Chain</i>).

INTRODUCCION

A. Interacciones Virus-Hospedador

Las enfermedades provocadas por agentes patogénicos constituyen un reto sanitario global y su costo social y económico concomitante hacen importante el interés en tratarlas debidamente. Los principales patógenos responsables de las enfermedades infecciosas a nivel global, son las bacterias y los virus. Mientras los primeros –salvo excepciones- pueden sobrevivir independientemente y ser eliminados con antibióticos, los virus constituyen un mayor reto, porque solo pueden ser contenidos en su desarrollo y multiplicación, haciendo difícil su tratamiento. En este contexto, existen múltiples esfuerzos de investigación que están orientados a entender en profundidad los mecanismos por los cuáles un virus es capaz de invadir y tomar control de la célula huésped para su propio beneficio, multiplicarse e invadir otras células. La base de la comunicación en el sistema virus-hospedador lo constituyen las interacciones físicas moleculares proteína-proteína (PPIs) tanto dentro como en la superficie celular, las cuales juegan un rol fundamental en la iniciación de la infección viral [1]. Además la supervivencia de los virus, a consecuencia de estar formados sólo por ácidos nucleicos, proteínas, algunos lípidos y otros compuestos menores y no codificar la maquinaria necesaria para su replicación, es casi completamente dependiente de estas PPIs [2]. Por lo tanto, *un conocimiento detallado de las interfaces de interacción proteicas es esencial y necesario porque permite entender mecanismos de infección viral, etiología de la enfermedad concomitante, así como el desarrollo de drogas o agentes terapéuticos efectivos contra los virus.*

Durante el curso de una infección viral, interfaces endógenas que median interacciones intraespecíficas (viral-viral o hospedador-hospedador) son blancos constantes de interfaces exógenas que median interacciones interespecíficas, es decir que las proteínas virales y celulares interactúan y compiten entre sí constantemente por determinadas interacciones [3]. Aun cuando los seres humanos pueden ser infectados por una multiplicidad de virus diferentes y existen tendencias por el estudio de determinadas familias virales (y por tanto una disponibilidad de datos parcializada) [2], un estudio basado en una estrategia de mapeo genómico de PPIs patógeno-huésped experimentalmente verificados, ha sugerido que a nivel funcional, las proteínas virales (como las bacterianas) tienden a interactuar en general con proteínas nodo o “*hub*” (aquellas que pueden interactuar con otras múltiples proteínas) y con proteínas “cuello de botella” (las cuáles son centrales en diversas vías metabólicas en la célula) [1] (Figura I.1). Esto sugeriría que los virus habrían evolucionado para interactuar con esas proteínas dado que las mismas controlan procesos y pasos celulares críticos. De esta manera pueden alterar e inhibir funciones e incluso forzar nuevas conexiones (que no están presentes

en una célula no infectada), de manera que producen cambios funcionales posteriores en muchos niveles de la red celular [4].

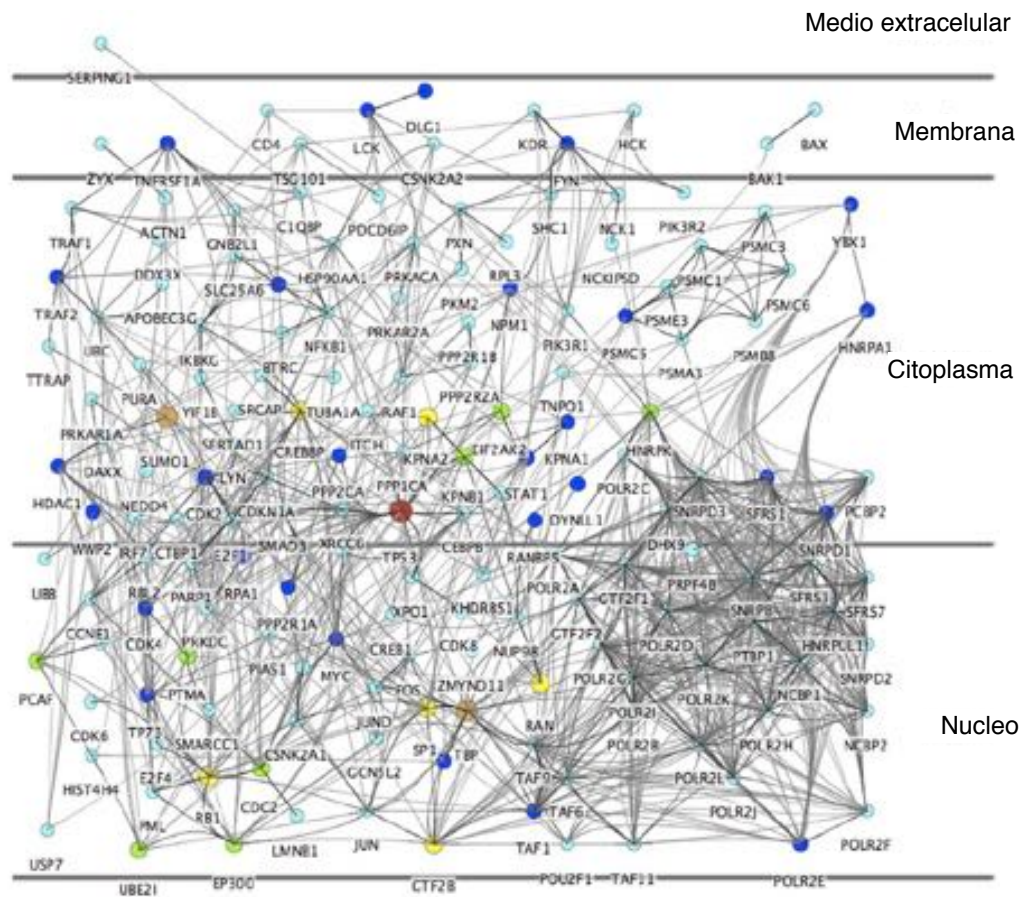


Figura 1.1. Red de interacciones entre proteínas humanas y proteínas de múltiples grupos virales. Mientras las líneas denotan la relación establecida entre las diferentes proteínas humanas en los diferentes compartimentos celulares, el tamaño y color de los círculos refieren al número de virus que interactúan con ellos (celeste: dos; azul: tres; verde: cuatro; amarillo: cinco; naranja: seis; rojo: siete). Esto demuestra como los virus interaccionan preferencialmente con proteínas humanas esenciales en la regulación de diferentes procesos celulares, subvirtiendo toda la célula con un número no necesariamente alto de PPIs. (Adaptado de [1]).

Las proteínas virales son capaces de manipular a la célula hospedadora a partir de PPIs a distintos niveles que le permiten: controlar el ciclo celular y utilizar los factores de transcripción a su favor para asegurar la transcripción de sus genes, intervenir activamente en los procesos de regulación epigenética y de la apoptosis, manipular señales de transducción, respuesta inmune, procesos degradativos y subvertir la maquinaria celular para transportar material viral a través de membranas internas o externas celulares [1, 5].

Por otro lado, estudios que integran redes de interacción con datos estructurales 3D disponibles para PPI virus-hospedador, han hallado que los virus tienden a interactuar e imitar o “hacer mímica” de interfaces proteicas hospedador-hospedador ya existentes, siendo estas

interacciones mediadas por principios funcionales, estructurales y evolutivos diferentes de aquellos que gobiernan las PPIs endógenas del hospedador. Las PPIs virus-hospedador son usualmente multivalentes porque involucran varias interfaces proteicas de interacción mediadas por contactos entre dominios proteicos globulares (globular-globular) o entre un dominio globular y una secuencia peptídica –motivo lineal de interacción- (globular-péptido). Además, estas PPIs se caracterizan por mediar interacciones transientes y estar sujetas a mayores presiones evolutivas debido a la constante “batalla” librada entre los virus para hacerse del control celular y de la célula, para evadir y reprimir estos procesos [6, 7].

B. El desorden intrínseco en proteínas virales

El tamaño de los genomas virales se caracteriza por estar notablemente colapsado. Debido a la limitación impuesta por este tamaño genómico, los virus frecuentemente presentan proteínas codificadas en marcos abiertos de lectura solapados [8] y se valen también del *splicing alternativo* para maximizar la información disponible en sus genes. En consecuencia, una simple mutación puede afectar a más de una proteína viral, sobre todo cuando las tasas de mutación viral por base es alta (en el orden de 10^{-8} a 10^{-5} en virus a ADN comparativamente con el orden 10^{-9} para bacterias y eucariotas) [9]. El reducido genoma viral y por lo tanto, el número de proteínas codificado por los virus, en comparación con organismos superiores, parecería *a priori* insuficiente para establecer la plétora de interacciones que requieren para infectar, subvertir la maquinaria celular a su favor y replicarse a nivel en el que lo hacen. Además, dominios globulares proteicos y plegamientos funcionales no toleran múltiples interfaces de interacción simultáneamente, algo que sería requerido por los virus para establecer todas sus PPIs con el hospedador [10]. Una respuesta a esto, estaría dada por *las ventajas funcionales que confiere el desorden intrínseco presente en las proteínas virales*. La falta de estructura rígida, elevada flexibilidad conformacional y otras características propias de una Proteína Intrínsecamente Desordenada (IDP, por sus siglas en inglés) le permitiría unirse a varias proteínas del hospedador a la vez, en diferentes localizaciones y contextos de la célula, permitiendo una efectiva interferencia en las redes celulares. Además, la falta de restricciones a nivel estructural en IDPs podrían hacer más tolerable la antemencionada alta tasa de mutaciones encontrada en virus [11]. *Esto lleva a preguntar: ¿Qué son las IDPs? ¿Qué características posee el desorden intrínseco que otorgaría tales ventajas?*

C. Las proteínas con desorden intrínseco existen y son abundantes en los organismos

Las proteínas constituyen uno de los componentes esenciales de la célula, críticas para el mantenimiento de sus múltiples funcionalidades y estructura, y su disfuncionalidad está asociada a trastornos y enfermedades de los más diversos en los organismos. Debido a esto, el interés en su estudio y descripción ha sido histórico. Hacia finales del Siglo XIX, ya se conocía que la catálisis enzimática era un proceso llevado a cabo por proteínas y en este contexto, Emil Fisher fue pionero al establecer el concepto de “llave-cerradura” para describir la factibilidad la misma, cuando en sus estudios sobre metilglucósidos encontró que sus isómeros α - y β - no podían ser hidrolizadas por las mismas enzimas [12]. Este concepto de “llave-cerradura” implica una necesaria complementariedad estructural para que se genere el proceso catalítico y por tanto suponer que la proteína (enzima) posee una estructura tridimensional consolidada, lo que también es apoyado por el hecho de que la desnaturalización proteica conduce a una pérdida en su actividad enzimática. Por otro lado, la introducción de la técnica de cristalografía de Rayos X, conllevó a la resolución de cristales de proteínas que presentaban estructuras bien definidas, las cuales fueron continuamente depositadas en el Banco de Datos de Proteínas (PDB, por sus siglas en inglés) [13]. Estas incontables evidencias experimentales, consolidaron por mucho tiempo la idea de que las proteínas sólo pueden poseer una estructura tridimensional bien establecida, pues esta es necesaria para que la misma cumpla su función.

Sin embargo para finales de la década de 1990, evidencias experimentales dieron cuenta de proteínas que en su estado nativo son desordenadas y que su funcionalidad reside en ese desorden, como las aportadas por Kriwaky en estudios sobre p21 [14]. Paralelamente, estudios bioinformáticos llevados a cabo sobre la base de datos SwissProt, identificaron más de quince mil proteínas que poseen segmentos de más de 40 residuos contiguos predichos sin estructuras [15]. Wright y Dyson, por su parte, revelaron que una gran proporción de secuencias de genes codifican para proteínas que están desplegadas en solución, la cual es su condición nativa y funcional [16]. Las evidencias experimentales y análisis bioinformáticos fueron y son clave en dar a entender que el desorden intrínseco en las proteínas existe y es abundante en los organismos. Estas proteínas, agrupadas hoy bajo el nombre de Proteínas Intrínsecamente Desordenadas (IDPs) [17] demuestran que la presencia de una estructura definida no es una condición *sine qua non* para la funcionalidad proteica.

C.1. Abundancia de las IDPs en organismos

Ya para principios de los 2000 estudios realizados por Ward y colaboradores, pudieron identificar que las regiones desordenadas representan un porcentaje importante de los proteomas de los organismos. Mientras que el proteoma de *archaeas* y bacterias presenta entre un 18 y un 30% de regiones intrínsecamente desordenadas mayores a 30 residuos, en eucariotas este valor asciende hasta un 35 - 50% [18](44% en el caso de los humanos [15]), sugiriendo además -junto a otros trabajos [19]- que habría una relación entre la complejidad y abundancia de regiones desordenadas en los organismos. Mas actualmente, se ha identificado que en el caso de los virus, el contenido de desorden intrínseco ascendería hasta un 30% del tamaño de su proteoma [11]. Sin embargo, dicho porcentaje variaría significativamente para distintas familias virales, como lo fue establecido en un estudio reciente por Pushker y colaboradores [20].

D. Proteínas Intrínsecamente Desordenadas (IDPs) y sus características

Las Proteínas Intrínsecamente Desordenadas (IDPs) o Regiones Intrínsecamente Desordenadas (IDRs) son aquellas que no poseen una única estructura tridimensional consolidada en solución, en condiciones fisiológicas de pH y fuerza iónica y en ausencia de un compañero de interacción [11]. Estas proteínas presentan un ensamble dinámico de conformaciones posibles en solución, pudiendo experimentar transiciones estructurales en diferentes condiciones, además de presentar tendencias estructurales locales tanto de estructura secundaria como -en ciertos casos- grados parciales de estructura terciaria [21-23]. Las IDPs e IDRs, aun en ausencia de plegamientos característicos conservan propiedades funcionales biológicas, fundamentalmente involucradas en procesos regulatorios celulares y vías señalización y reconocimiento, en tanto se encuentran negativamente asociados a procesos catalíticos [18, 24].

D.1. Características de secuencia de las IDPs

Las IDPs presentan una tendencia aminoacídica o composicional clara con respecto a las proteínas globulares. Aun cuando se caracterizan por presentar baja complejidad de secuencia (pocos tipos de aminoácidos en la secuencia primaria) [25], estas proteínas se encuentran depletadas en aminoácidos hidrofóbicos y aromáticos (ej. C, V, L, I, M, F, W, Y) – “aminoácidos favorecedores del orden”-, en tanto que están enriquecidas de aminoácidos polares y cargados (ej. S, N, Q, D, E, K, R) y disruptores de estructuras (P, G) y Alaninas – aminoácidos promotores del desorden- [21, 26]. Esto se puede ver evidenciado en el gráfico de

Uversky (Figura 1.2 A), que denota la relación *carga neta-hidrofobicidad* medias de IDPs como alta-baja, mientras que para proteínas globulares es baja-alta. Estas tendencias permiten inferir que el desorden podría ser consecuencia natural de un debilitado efecto hidrofóbico, que dirige el plegamiento de polipéptidos en estructuras terciarias definidas [23], sumado a un efecto repulsivo de cargas electrostáticas similares, que mantienen a estas proteínas en estado desplegado. Esto confiere a las IDPs, unos Radios de *Stokes* o hidrodinámicos (R_S o R_H), mas elevados que el de contrapartes globulares. Además, la alta proporción de residuos cargados les confieren a las IDPs puntos isoeléctricos (pI_{TEO}) extremos, considerándose proteínas ácidas o básicas [26].

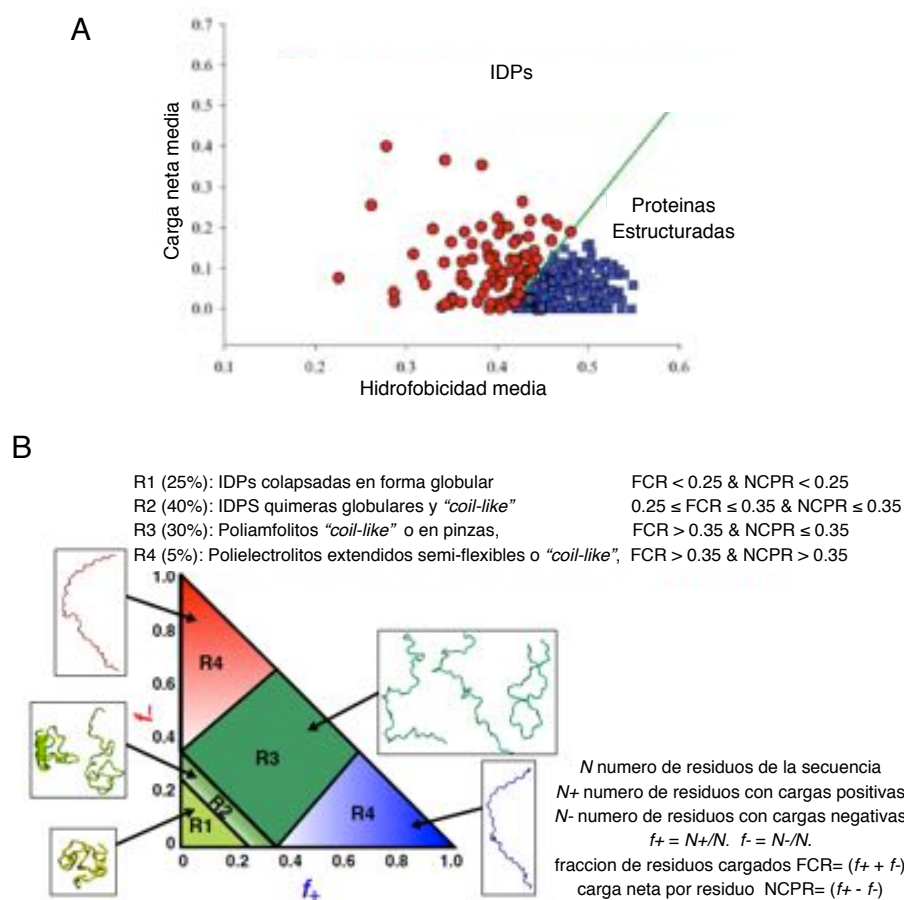


Figura 1.2. Características de secuencias de IDPs. A) Gráfico de Uversky de carga vs hidrofobicidad medias construido a partir de 275 proteínas globulares estructuradas (en cuadros azules) y 91 IDPs (en círculos rojos). La recta verde corresponde a la función $\langle R \rangle = 2.743 \langle H \rangle - 1.109$, donde $\langle R \rangle$ y $\langle H \rangle$ representan la carga e hidrofobicidad media, respectivamente (Adaptado de [27]) B) Diagrama de estados que muestra la clasificación conformacional de las distintas IDPs en función de sus secuencias. Los porcentajes presentados representan la abundancia con la que estos ensambles conformacionales de IDPs se encuentran en DISPROT [28] (Adaptado de [29]).

A su vez, las IDPs pueden identificarse con distintos sub-sets con características estructurales y funcionales diferenciales en función de sus secuencias. Estudios que combinan

evidencias experimentales y computacionales, llevadas a cabo por Pappu y colaboradores, establecieron que la información codificada en la secuencia primaria está relacionada con el ensamble de conformaciones que posee una IDP. Asumiendo una baja hidrofobicidad y bajo contenido de prolinas para las IDPs, estas pueden adoptar diferentes conformaciones en función del contenido de residuos polares y/o residuos cargados (negativos o positivos a pH fisiológico) y la distribución de las cargas en la secuencia [29-31]. Así, las IDPs se dividen en tres categorías: tractos polares, polianfolitos y polielectrolitos (fuertes o débiles). Los tractos polares están enriquecidos en residuos polares, pero depletados en residuos cargados; mientras que los polianfolitos tienen mayor contenido de residuos cargados, con un total de cargas positivas y negativas aproximadamente equivalentes y los polielectrolitos están cargados y ese exceso de cargas es positiva o negativa. En la Figura I.2 B, se esquematiza en el Diagrama de estados, cómo diferentes parámetros dependientes de las cargas de las IDPs están relacionados con su ensamble conformacional en tanto dominen procesos de repulsión y solvatación de los polipéptidos o más bien fuerzas atractivas intracatenarias. Las regiones R1, R2 y R3 comprenden el 95% de las secuencias de IDPs hasta ahora conocidas [29].

D.2. Características estructurales de las IDPs

Las IDPs no constituyen un grupo estructural homogéneo de proteínas con grado similar de desorden intrínseco, como se ha demostrado experimental y computacionalmente. Se ha propuesto entonces que las proteínas forman parte de un *continuum* conformacional [32] donde hay una gradualidad de contenido estructural (o ausencia del mismo) en un rango de completamente desordenadas y extendidas (altos R_S o R_H) a completamente ordenadas y compactas (R_S o R_H menores en comparación) (Figura I.3 A). En este modelo no hay límites conformacionales precisos y como las IDPs/IDRs son altamente dinámicas y fluctúan en un ensamble de conformaciones heterogéneas, éstas podrían poblar varios estados del *continuum*. Además, las proteínas pueden ser modulares, es decir que una misma cadena polipeptídica puede presentar una región plegada formando un dominio altamente estructurado, con funciones enzimáticas, en tanto otra región puede presentarse altamente dinámica y desordenada (IDR).

No obstante la validez del modelo del *continuum* conformacional planteado, las IDPs suelen clasificarse por una cuestión de simplicidad según el “modelo del cuarteto” [21, 27] (Figura I.3 B) en función de su contenido residual de estructura y el entorno fisicoquímico en el que se encuentren. Las categorías constituyen: “Estructurado u ordenado (G)”, “Glóbulo fundido o “*Molten globule*” (MG)”, “Pre-Glóbulo fundido o “*Pre-Molten globule*” (PMG)” y “Conformación

azarosa o “*Random coil*” (RC)”. Para este modelo, una región desordenada libre (sin estar unida a un compañero de interacción) podría poblar cualquiera de las conformaciones antemencionadas excepto por el estado “Estructurado u ordenado”.

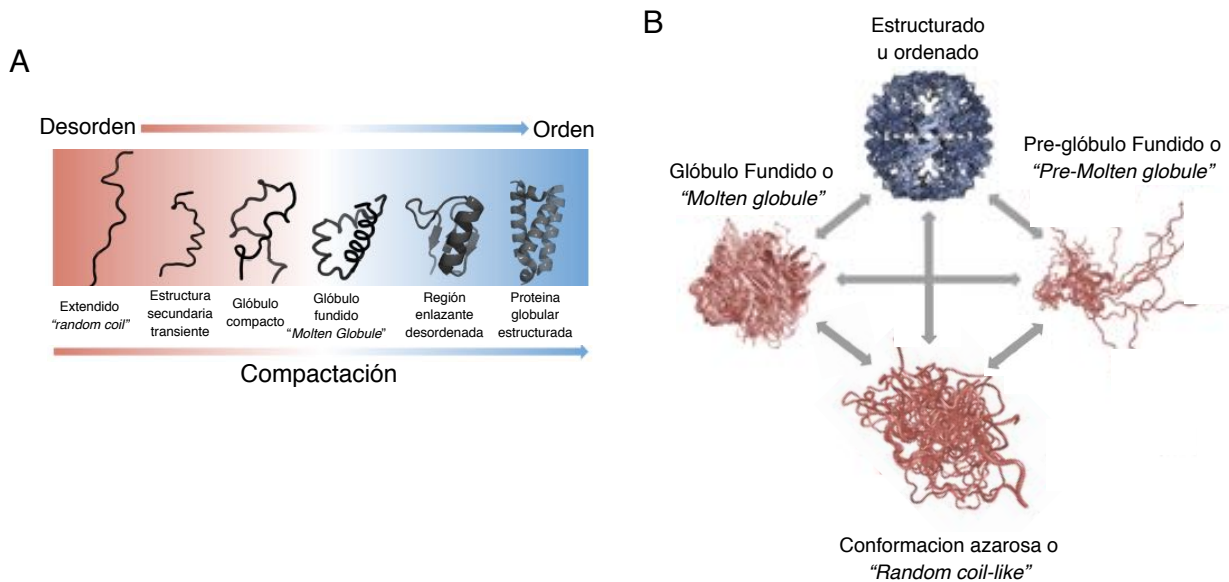


Figura 1.3. Características estructurales de IDPs. A) Representación esquemática del modelo de *continuum* conformacional de estructuras proteicas. El gradiente va desde conformaciones altamente dinámicas y expandidas (en rojo) hasta estados globulares altamente compactes y dinámicamente restringidos (en azul) B) Modelo del cuarteto de estados conformacionales proteicos, en el que sus funciones se desprenden de estas cuatro conformaciones y las transiciones entre ellas. Representaciones como ensambles conformacionales (Adaptado de [23]).

En esta clasificación:

- Las IDPs tipo “Glóbulo fundido o “*Molten globule*” (MG)” no poseen una estructura terciaria cooperativa rígida, pero presentan un grado de compactación considerable y poseen contenido de estructura secundaria estable y cierta topología funcional. Estas características dotarían a las IDPs tipo MG de regiones hidrofóbicas capaces de unir a la sonda fluorescentes hidrofóbicas como ANS (1-anilino-8-naftaleno-sulfonato). Estas presentan mayor sensibilidad a proteasas y tienen un R_H extendido alrededor de un 15% mayor al de su contraparte globular [27]. Un ejemplo de IDP MG lo constituye el dominio de interacción coactivador nuclear de la proteína de unión a CREB, asociado a la regulación del ciclo celular [23].
- Las IDPs tipo “Pre-Glóbulo fundido o “*Pre-Molten globule*” (PMG)” son menos compactas -o con R_H mas extendidos- que las IDPs MG y poseen un contenido residual de estructura secundaria transitoria. Se ha sugerido que éstas pueden interactuar con ANS de manera débil, indicando que podrían contener regiones hidrofóbicas residuales. Sus secuencias tienen bajo contenido de carga neta, lo que

las hace más compactas que las *RC-like* (abajo). Las proteínas ribosomales S12, S18, L23 y L32 y la calsecuestrina poseen naturaleza IDP PMG.

- c) Las IDPs de “Conformación azarosa o tipo *“Random coil”* (RC)” - o más apropiadamente *“Random coil-like”* (*RC-like*)- poseen una naturaleza altamente desplegada que los asemejan notoriamente (aunque no se trate de lo mismo) a lo que se definía típicamente como *“péptidos random coils”*. Para estos últimos se asume un estado ideal en el que los aminoácidos individuales no se encuentran de ninguna manera impedidos estéricamente, por lo que sus ángulos diedros (ϕ y ψ) pueden asumir con la misma probabilidad cualquier valor entre -180° y 180° , por tanto poblar cualquier zona del gráfico de Ramachandran [22]. Estas IDPs tienen propiedades hidrodinámicas altamente extendidas asociado a sus secuencias con alta carga neta y presentan muy poco a nada de estructura secundaria, aunque bajo ciertas condiciones puedan experimentar transiciones conformacionales parciales. Un ejemplo de IDP *RC-like* es la Protimosina α (ProT α) que se encuentra alta y negativamente cargado y está involucrada en remodelado cromatínico, transcripción, proliferación celular y apoptosis [33].

D.3. Características funcionales de las IDPs

Como consecuencia de sus distintas propiedades conformacionales, las proteínas globulares ordenadas y las IDPs operan de manera diferente y están involucradas en distintas funciones. Se ha demostrado que las funciones de las IDPs están comúnmente asociadas a procesos regulatorios y vías de señalización celular [18, 24], en contraposición a las funciones llevadas a cabo por las proteínas globulares, más bien involucradas en procesos catalíticos, biosíntesis y metabolismo.

Desde fines de la década de 1990, ya han sido varios los intentos por clasificar las funciones asociadas a las IDPs. Dunker y colaboradores, identificaron cerca de 30 funciones de IDPs y las agruparon en 4 categorías: cadenas entrópicas, modificaciones moleculares, reconocimiento y ensamblado [34]. Por su parte, Tompa desarrolló este esquema en seis categorías más amplias: cadenas entrópicas, efectores, sitios de modificación, recolectores, ensambladores y chaperonas [35]. Por otro lado, Gsponer y Babu han agrupado las funciones de las IDPs en tres categorías: facilitación de la regulación de procesos vía modificaciones post-traduccionales (PTM, por sus siglas en inglés), ensamblado o andamiaje y reclutamiento de otras proteínas y por último variabilidad y adaptabilidad conformacional [36]. No obstante a estas clasificaciones, bases de datos de IDPs como DISPROT y de motivos lineales de

interacción como ELM, manejan sus propios criterios de clasificación de funciones de IDPs [28, 37].

A modo simplificador, las funciones serán presentadas acorde a las categorías introducidas por Tompa:

- a) Cadenas entrópicas: esta función de las IDPs depende directamente de la naturaleza desordenada y flexibilidad de la cadena carbonada y no están involucradas en procesos de reconocimiento y plegamiento. Las regiones enlazantes o “*linkers*” flexibles, son ejemplo de cadenas entrópicas, al permitir el movimiento de dominios o sitios de interacción en los extremos de la cadena polipeptídica o las regiones espaciadoras o “*spacers*” que regulan distancias entre dominios. El dominio de proyección de la proteína 2 asociada a microtúbulo (MASP2) es un ejemplo, al repeler las moléculas que se aproximan a los microtúbulos generando el espacio necesario al citoesqueleto [23].
- a) Sitios de modificación: sobre ellos, las IDPs pueden sufrir modificaciones post-traduccionales que afecten su estabilidad, recambio-degradación, interacciones potenciales y relocalización en la célula. Las IDPs están enriquecidos en ellos y su accesibilidad está favorecida por el desorden intrínseco propio de la proteína permitiéndoles su reconocimiento y mediando interacciones transientes pero específicas con otras proteínas. Estos pueden por ejemplo, ser sitios de fosforilación o proteólisis.
- b) Efectores: las IDPs pueden modificar las funciones de su proteína compañera interactuante una vez unido a ellas. Ejemplo de ello son p21 y p27 que regulan ciclinas dependientes de quinasas (Cdks) responsables de la progresión del ciclo celular a través de efectos activadores o inhibitorios de función [38].
- c) Ensambladores: estas IDPs poseen grandes interfaces de interacción que permiten la unión de múltiples proteínas interactuantes y promueve la formación de complejos proteicos de alto orden. Los ribosomas bacterianos son un ejemplo de ello; aquí muchas de las proteínas son desordenadas y median una estabilización inicial con el ARN que conduce a su subsecuente plegamiento y agregado de otras proteínas para dar lugar al complejo ribosómico en su estado final [39].
- d) Recolectores: estas IDPs tienen capacidad de almacenar y neutralizar pequeños ligandos. Un ejemplo es la Caseína, una fosfoproteína de unión a calcio que es altamente desordenada y media la solubilización de agregados de fosfato de calcio en la leche [40].

- e) Chaperonas: corresponden proteínas que asisten a otras y al ARN a adquirir sus estructuras plegadas funcionales. Se cree que las IDPs pueden llevar a cabo esta función por sus capacidades para unir varias proteínas a la vez, estar involucrados en eventos de asociación y disociación rápidos y de plegamiento-acoplado-a-interacción (ver mas adelante). La chaperona desordenada Hsp33, por ejemplo, es regulada por estados redox y es inactiva en su forma plegada, siendo rápidamente activada por oxidantes como el ácido hipocloroso y protegiendo a otras proteínas de un desplegamiento producto de la oxidación [41].

D.4. Elementos que median interacciones en las IDPs

Como se ha visto, las IDPs cumplen múltiples funciones, la mayoría de las cuáles requieren que éstas se unan a otras proteínas (y otras moléculas) para llevarlas a cabo. Para ello disponen de diferentes regiones funcionales que se desempeñan como elementos que median interacciones: los motivos lineales de interacción (SLiMs, del inglés “*short linear interaction motifs*”), los elementos de reconocimiento molecular (MoRFs, del inglés “*molecular recognition features*”). Estos conceptos son análogos y a veces superponibles, radicando sus mayores diferencias en la longitud o tamaño que involucran y en si a dichos elementos se los define de nivel de secuencia (SLiMs) o a nivel estructural (MoRFs) [21].

D.4.1. Motivos lineales de interacción

Los SLiMs, o también conocidos sencillamente como motivos lineales (LM), son segmentos de la secuencia primaria, de una longitud generalmente menor a 15 residuos embebidos en una región desordenada y que modulan las funciones proteicas [42]. Estos motivos pueden presentar un sesgo composicional hacia aminoácidos cargados, hidrofóbicos y prolinas y estar depletados en glicinas y alaninas [43]. La distribución de los residuos que componen a los LM- y por los que este media sus interacciones- no es azarosa, sino que siguen lo que se conoce como “expresión regular” (Figura I.4 A), un patrón que determina que aminoácidos están permitidos o prohibidos en cada posición del motivo. Esto define posiciones “fijas” que son centrales en el motivo y poseen alto grado de conservación evolutiva y proporcionan el mayor aporte a la especificidad y afinidad de las interacciones que median con otras proteínas y posiciones “comodines” que pueden ser variables en secuencia. Residuos en su vecindad, los “adyacentes” y “flanquantes”, están menos conservados y se ha visto que también pueden modular en grado variable las interacciones de los SLiMs [44, 45].

Estos motivos pueden actuar como sitios de modificación: para eventos post-traduccionales como clivaje proteolítico, fosforilación e isomerización *cis-trans* y también como ligandos: promotores de la formación de complejos, anclaje o “*docking*” y sitios blanco para permitir relocalización sub-celular de proteínas.

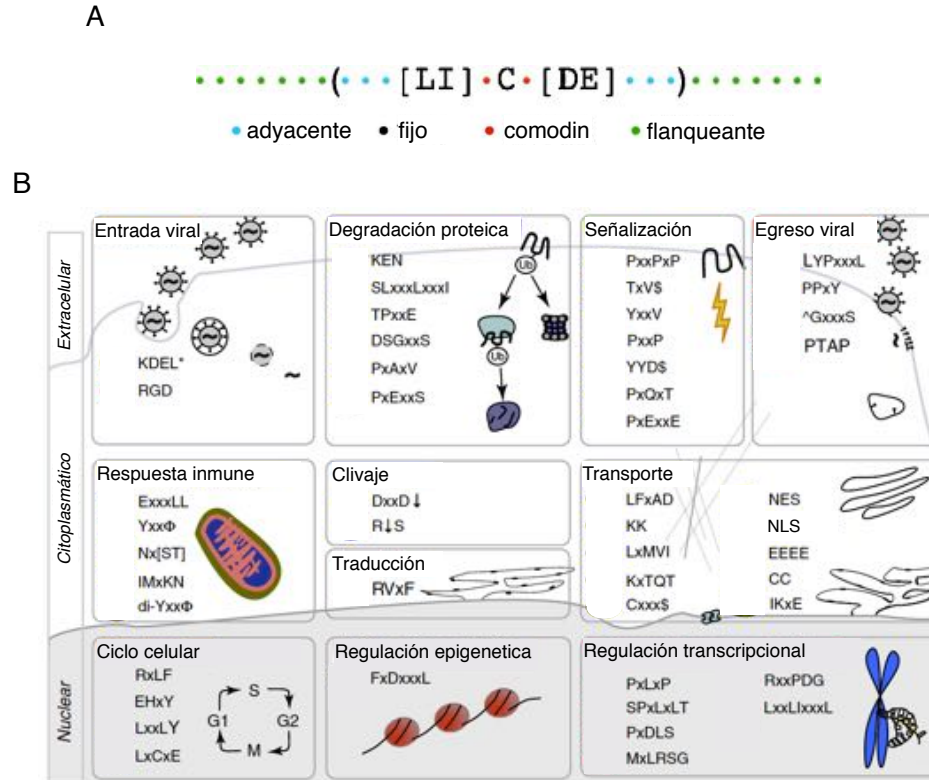


Figura 1.4. SLiMs en IDPs. A) Representación de la expresión regular de los SLiMs, con sus distintas posiciones señaladas. En el ejemplo, la expresión regular del motivo LxCxE, presente en blancos celulares y virales de pRb y discutida más adelante en esta tesis B) Representación selectiva de SLiMs que están presentes en proteínas IDPs virales y que muestran el rango diversificado de funciones en lo que intervienen a nivel celular al hacer mímica de los mismos (Adaptado de [5]).

Debido a la naturaleza corta (pocos residuos) de los SLiMs, múltiples de estos motivos pueden estar presentes en una misma región proteica desordenada –o incluso en “*loops*” entre dominios globulares- (se dice que las IDR están “enriquecidas” en ellos) y algunos pueden presentarse solapados. En consecuencia, estos motivos son módulos estructuralmente flexibles y se pueden adaptar a las diferentes proteínas con las que interactúan. Por otro lado, las interacciones en las que están involucradas implican superficies de contacto más pequeñas y afinidades un tanto menores (de moderadas a bajas: K_D 0.5–10 μ M [46]) en comparación con interacciones entre dominios globulares, por lo que participan de interacciones de carácter transientes.

Otra consecuencia no menor de su corta longitud, es que estos motivos tienen una alta propensión a evolucionar de manera convergente y como consecuencia de unas pocas

mutaciones puntuales, poder emerger en proteínas no relacionadas entre sí [5, 42]. Como corolario de ello, virus (y bacterias) han evolucionado haciendo “mímica” de los estos SLiMs ya presentes en las proteínas de los hospedadores, permitiéndoles competir efectivamente por interacciones en la célula, manipulando redes de PPIs fundamentales en el hospedador. En la Figura I.4 B, se muestran los múltiples ejemplos de SLiMs de los cuáles los virus hacen mímica y utilizan durante los distintos estadios de una infección viral, desde su ingreso a egreso y en los diferentes compartimentos celulares, demostrando el alto nivel adaptativo de los mismos y su capacidad para poder manipular vías y mecanismos celulares de los más diversos a su favor [5].

D.4.2. Elementos de reconocimiento molecular

Los MoRFs, corresponden a segmentos de 10 a 70 residuos de longitud que son faltos de estructura, dentro de regiones intrínsecamente desordenadas mas largas e intervienen y promueven activamente interacciones proteína-proteína. Estos elementos se caracterizan por experimentar transiciones de desorden-a-orden simultáneamente al producirse la interacción con otra proteína (un proceso que también se denomina plegamiento-acoplado-a-interacción). Estos MoRFs podrían tener un contenido residual de estructura previo a las interacciones y adoptar finalmente una forma consolidada de α -hélice, lámina- β o una estructura irregular (pero rígida) tras la formación del complejo, dando lugar a los α -MoRFs, β -MoRFs y ι -MoRFs y también a formas “complejas” producto de la mezcla de las anteriores [47].

Muchos ejemplos de estas estructuras se encuentran depositadas en el PDB. Una proteína que contiene múltiples MoRFs y que es fundamentalmente desordenada en ausencia de sus interactores es la proteína p53 (Figura I.4). Los primeros MoRFs que presenta p53 se encuentran en su dominio de transactivación (N-TAD), entre los residuos 20 a 50 (Figura I.5 Área punteada I.). Las estructuras cristalinas de los complejos revelan que p53 presenta α -MoRFs (α -hélice) al interactuar con RPA70 y Mdm2 (Figura I.5 Área punteada I. a y b, respectivamente) aun cuando en ausencia de sus compañeros de interacción es una región que se encuentra desordenada. No obstante a ello, existe evidencia experimental de que esta región presenta una mínima tendencia estructural a α -hélice en su estado libre [48]. Estos elementos pre-formados servirían de contacto inicial para eventos de interacción con propiedades termodinámicas y cinéticas asociadas distintas de aquellos en las que la región interactuante está completamente desordenada.

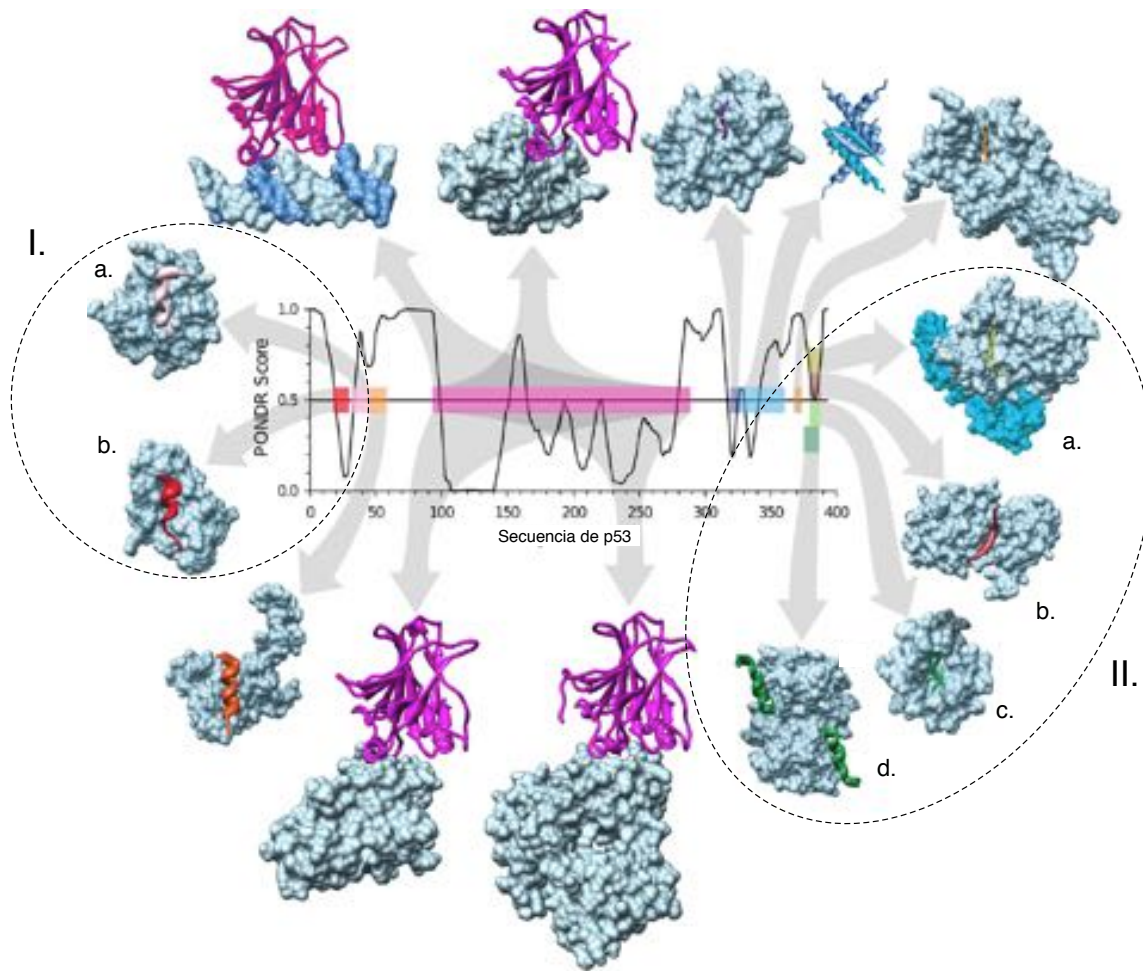


Figura I.5. MoRFs en p53 y su plasticidad conformacional como IDP. Representación de las múltiples interacciones de las que participa p53 y hay estructuras resueltas disponibles en el PDB. En el centro se encuentra una predicción de desorden de p53 en función de su secuencia, donde los valores por debajo de 0.5 indican regiones ordenadas y por encima, desordenadas. El segmento magenta corresponde al Dominio Central ordenado de p53. Los otros colores corresponden a las secuencias involucradas en distintos complejos. (Adaptado de [23]).

En su dominio C-terminal (CTD) p53 también contiene MoRFs que pueden interactuar con múltiples proteínas. Se puede observar como una única pequeña secuencia predicha totalmente desordenada puede adoptar distintas estructuras resultantes al formar complejos con diferentes proteínas (Figura I.5 Área punteada II.), evidenciando: ausencia de elementos pre-formados, promiscuidad y plasticidad conformacional de las regiones desordenadas involucradas en interacciones. La secuencia entre los residuos 370 y 390 de p53 es capaz de estructurarse en α -hélice en complejo con S100 $\beta\beta$ (Figura I.5 Área punteada II. d), mientras que forma una lámina- β con Sirtuina (Figura I.5 Área punteada II. b), y ι -MoRF en complejos con Cdk2-Ciclina A (Figura I.5 Área punteada II. a) y con el dominio de bromo de CBP (Figura I.5 Área punteada II. c) [23].

D.5. Características de las interacciones mediadas por IDPs

En la actualidad se considera un paradigma que las interacciones entre proteínas (o proteínas y otras moléculas) se produzcan por complementariedad de interfaces que definen afinidad y especificidad de unión. En consecuencia, en el caso de las IDPs se acepta que estas requieren experimentar procesos de plegamiento-acoplado-a-interacción cuando forman complejos [23]. Por un lado, a nivel termodinámico, esto implica que las interacciones mediadas por regiones desordenadas son específicas pero generalmente transientes, ya que al producirse la unión, una región desordenada y flexible pierde libertad conformacional (entropía), lo cual reduce la energía libre global de la interacción, conduciendo a interacciones más débiles en comparación con regiones plegadas que interactúan con la misma entalpía y/o especificidad [23], aunque esto no implica que interacciones simultáneas de múltiples regiones desordenadas conduzcan a interacciones globales mucho más afines (ver más abajo). Por otro lado, a nivel mecanístico, surge el cuestionamiento de si el plegamiento de la región desordenada se produce antes, durante o después de la interacción. Esto último ha llevado a proponer que esto ocurriría utilizando dos mecanismos: el ajuste inducido y la selección conformacional [16, 47] (Figura I.6) los cuales tendrían o uno u otro un rol dominante, dependiendo de las concentraciones individuales de proteínas y de la cinética de interacción.

En la selección conformacional (Figura I.6), una población parcialmente estructurada del ensamble conformacional de la IDP en estado libre es preferencial para mediar la interacción con la interfaz de su proteína compañera. Una evidencia experimental se encuentra en lo descrito para los residuos del N-TAD de p53 que interactúan con RPA70 y Mdm2 (Figura I.5 Área punteada I.). Es importante destacar que la pre-existencia de estructura no implica que el mecanismo de selección conformacional esté mediando la interacción, como fue probado en el sistema KID-KIX [49]. En el mecanismo de ajuste inducido (Figura I.6), la región desordenada experimenta una transición de desorden-a-orden en forma simultánea con el proceso de interacción. Un ejemplo de ello estaría dado en el otro extremo de p53, su dominio CTD (Figura I.5 Área punteada II.), en el que una sin evidencias de pre-estructura se pliega en distintas conformaciones finales. No obstante, estos mecanismos no son excluyentes y las IDPs utilizarían una combinación de ellos (modelo sinérgico) en los procesos de interacción [50].

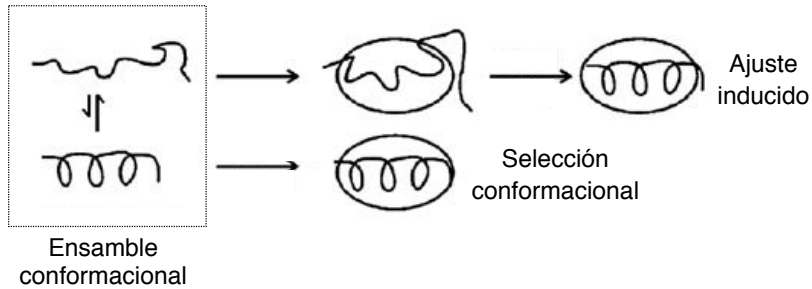


Figura 1.6. Mecanismos de interacción posibles para IDP. Representación esquemática mostrando los dos posibles mecanismos por los cuáles las IDPs pueden interactuar con otras proteínas a selección conformacional o el ajuste inducido (Adaptado de [51]).

Con el reconocimiento de la importancia biológica de las IDPs, los estudios que intentan comprender los mecanismos por los cuáles estas proteínas interactúan son cada vez más numerosos. Las evidencias experimentales dan cuenta que los mecanismos de interacción que involucran IDPs son muy diversos.

Por un lado a pesar de que una determinada región desordenada pueda experimentar proceso de plegamiento-acoplado-a-interacción descrito anteriormente, ahora es sabido que frecuentemente la estructuración de estas proteínas puede no ser completa y por lo tanto, permanecer parcialmente desordenada aun cuando está formando el complejo. Esto llevó a desarrollar el término complejos “fuzzy” para describir tales casos [52]. En los complejos “fuzzy”, las IDPs pueden presentar promiscuidad estructural estática, es decir que pueden estructurarse en más de una forma estable una vez unido en complejo o implicar que gran parte de esta conserve su flexibilidad y desorden. Esto conduciría a ventajas funcionales como la interacción con múltiples proteínas en forma simultánea, la posibilidad de “pescar” otras proteínas con las que requieran interactuar o intervenir (promoviendo o inhibiendo) la interacción de segmentos que experimentan plegamientos [21]. Por ejemplo Artemis, una nucleasa que juega un rol esencial en la inmunidad adaptativa, reparación de ADN e integridad genómica en células de mamíferos, es capaz de unirse al dominio de unión de ADN (DBD) de Ligasa IV, a través de su dominio C-terminal el cual es altamente desordenado. Tras producirse la unión de Artemis, los residuos 485 a 495 se estructuran en α -hélice y el resto permanece desordenado (Figura 1.7 A II.), formando un complejo “fuzzy”, de modo que al retener desorden puede interactuar por la misma región con otras proteínas, como la quinasa ADN-PKcs [53].

Como se ha mencionado, las IDPs pueden interactuar a través de SLiMs y mediar interacciones específicas pero poco afines. No obstante, se ha visto que múltiples SLiMs en una IDP pueden mediar individualmente interacciones débiles pero colectivamente participar de interacciones de muy alta afinidad. Las nucleoporinas intrínsecamente desordenadas (Nups)

que forman parte del complejo del poro nuclear (NCP) por ejemplo, poseen dominios con múltiples motivos FG que les permiten interactuar con transportadores nucleares como Importina β de manera polivalente. Se ha visto que cada uno de estos motivos median interacciones milimolares, pero en conjunto producen interacción efectiva del orden nanomolar (en términos de constantes de disociación). Llamativamente, estas interacciones son muy rápidas y se establecen a nivel local, de manera que se mantiene una alta plasticidad proteica aun formado el complejo (Figura 1.7 A III.), sin haber estructuración alguna en las Nups-IDP, asegurando así un transporte rápido y específico a través del NCP [54]. La observación de este tipo de interacción se distingue de otros en los que está mediada por un solo SLiM, pero también porque no requiere transición conformacional (plegamiento-acoplado-a-interacción). Respecto de esto nuevas evidencias, dan cuenta de un caso más extremo, en el que una IDP puede interactuar con otra IDP para formar complejos en los que el desorden persiste completamente en ambas proteínas tras producirse la interacción, la cual es de muy alta afinidad y está mediada por interacciones electrostáticas. Borgia y colaboradores han observado que la Histona H1 (IDP cargada positivamente) y su chaperona nuclear Protimosina α (ProT α) (IDP cargada negativamente) involucradas en la remodelación cromatínica, dan lugar a un complejo de afinidad en el orden picomolar y en el que persisten sus caracteres altamente desordenados, flexibles y dinámicos [33] (Figura 1.7 A IV.). Este panorama confirma a las IDPs como elementos con características particulares que los hacen capaces de formar complejos con grado muy variable de desorden asociado y una plétora de mecanismos moleculares subyacentes mediando dichas interacciones.

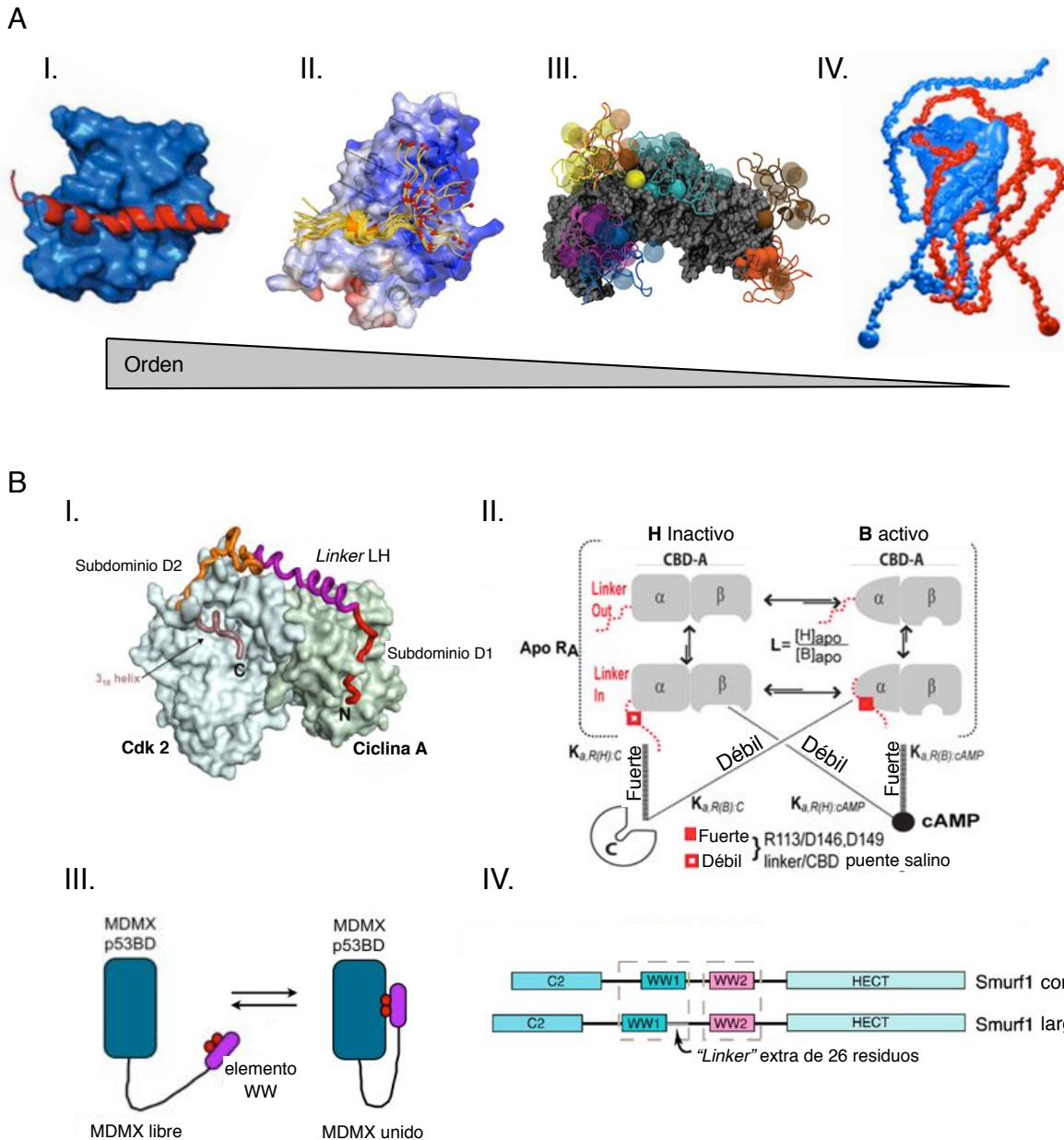


Figura I.7. Distintos tipos de complejos formados por IDPs y el rol de "linkers" desordenados en diversas interacciones. **A)** Representación de la plétora de conformacional de complejos en los que las IDPs se encuentran involucrados. **I.** Interacción con una proteína ordenada para formar un complejo cuasi-globular. Dominio C-terminal del factor de transcripción hematopoyético c-Myb y su interacción directa con el dominio KIX del co-activador transcripcional CBP. El dominio C-t de c-Myb es desordenado y estabiliza α -hélice al unirse al dominio ordenado KIX de CBP (Adaptado de [49]). **II.** Un complejo "fuzzy" en el que se retiene desorden tras formarse el complejo. Una pequeña región del dominio C-terminal de Artemis se pliega en hélice (líneas amarillas) al unirse al dominio ordenado de unión a ADN de Ligasa IV (superficie azulada), mientras el resto del dominio permanece en desorden (Adaptado de [53]). **III.** Un complejo formado por una proteína globular y una proteína desordenada que mantiene completamente su plasticidad tras formarse el complejo. Los Nups (en colores) se unen a través de sus múltiples motivos FG (pueñas bolitas) de manera rápida y polivalente a Importina- β (superficie gris) mediando un complejo de alta afinidad y en el que Nups retiene su desorden y flexibilidad (Adaptado de [54]). **IV.** Un complejo formado por dos IDPs que retiene alto grado de desorden tras acomplejarse. Histona H1 (azul) y su chaperona nuclear ProTa (rojo) unidas por interacciones electrostáticas en un complejo de muy alta afinidad (Adaptado de [33]). **B)** Ejemplificaciones de distintas regiones "linkers" con funcionalidades diversas. **I.** Dominios KID de p21-p27 unido al complejo Ciclina-Cdk por sus dominios D1 (cilindro rojo) y D2 (cilindro naranja) y estabilizados por un "linker" formando una hélice extendida (cilindro violeta) (Adaptado de [55]). **II.** Estabilización de las formas activa (B) e inactiva (H) por acción de la formación de un puente salino mediado por el "linker"-línea roja punteada- en ApoRA ("linker"+ apo CBD $^{\alpha-\beta}$) (Adaptado de [56]). **III.** MdmX y su "linker" flexible interno que regula la interacción entre el motivo o elemento WW y el dominio p53BD por un mecanismo de concentración efectiva (Adaptado de [57]). **IV.** Representación esquemática de las proteínas E3 ubiquitin ligasas Smurf1 y sus dominios WW que median las interacciones y el "linker" flexible extra que disminuye la afinidad de esta proteína por sus blancos (Adaptado de [58]).

Por otro lado, se ha descrito que las interacciones entre proteínas están mediadas por elementos funcionales como SLiMs y MoRFs, o incluso grandes superficies de dominios. Estos elementos funcionales pueden estar conectados entre sí por regiones enlazantes cuyo desorden y flexibilidad asociados permiten, por ejemplo su movimiento. No obstante, mucha evidencia ha sido expuesta acerca de que estas regiones enlazantes o más sencillamente “*linkers*”, cumplirían funciones más complejas que la de ser meros conectores y contendrían mucha información codificada en ellas, esencial para su funcionalidad. Así, estas regiones “*linker*” podrían participar activamente en la interacción mediada por regiones de reconocimiento por ejemplo estabilizando los complejos formados. Esto sucede en el caso de la interacción entre los reguladores del ciclo celular p21 y p27 y los complejos Ciclina-Cdk. p21 y p27 son IDPs con funciones inhibitorias de quinasas que las llevan a cabo a través de sus dominios KID, formados por un dominio D1 que contiene el SLiM RXL de unión a ciclinas, un dominio D2 de unión a Cdk y al una región “*linker*” LH. En el estado libre estas proteínas muestran un alto grado de desorden y la región “*linker*” LH parece poblar parcialmente estructura hélice, la cual establece contactos con el complejo Ciclina/Cdk, luego de la unión del SLiM RXL, permitiendo la posterior unión de D2 a Cdk (Figura 1.7 B I.). Este “*linker*” ofrece en este caso una gran superficie de estabilización a los complejos cuando D1 y D2 reconocen sus sitios de interacción complementarios [55].

En otros casos, los “*linkers*” desordenados podrían mediar interacciones muy débiles e incluso con ello participar activamente de procesos de regulación alostérica, algo que en el concepto clásico está asociado mecánicamente a proteínas globulares que experimentan cambios conformacionales que modifican afinidad en otros sitios alejados de interacción. Akimoto y colaboradores, han descrito un caso en el que una región interna de la subunidad R de PKA, un “*linker*” flexible y dinámico, controla el equilibrio entre las formas activa (H) e inactiva (B) del dominio $CBD^{\alpha-\beta}$ de dicha subunidad. Ellos han visto que la formación de un puente salino mediado por un único residuo del “*linker*” con residuos del dominio CBD^{α} , estabiliza preferencialmente la forma activa. Cuando este puente se rompe, el equilibrio conformacional se desplaza hacia la forma inactiva H (Figura 1.7 B II.) [56]. En este caso el “*linker*” provee un nivel adicional de regulación funcional en proteínas modulares. En otros casos, como el del “*linker*” interno de MdmX, el mismo también juega un rol regulatorio en las funciones de esta proteína a través de otro mecanismo. MdmX que es un regulador negativo de la actividad de p53, posee un motivo WW interno de autoinhibición que compite con p53. El motivo WW y el sitio de la misma MdmX por el que compite con p53, distan por una región “*linker*” que influiría en la afinidad de interacción. En sus experimentos, Borchers y

colaboradores demostraron que cuando se produce la interacción interna en ausencia del “*linker*” que los conecte (eliminado por mutación), se unían con una afinidad 400 veces menor que en su presencia. De este modo sugirieron el “*linker*” flexible es responsable de la alta afinidad por un mecanismo de concentración efectiva, que aumentaba la disponibilidad de WW sobre su sitio interno de interacción (Figura I.7 B III.) y que dicho mecanismo autoinhibitorio sería importante en el proceso regulatorio de p53 [57].

Otros grupos de investigación también han puesto su atención en que los “*linkers*” no son actores pasivos en las interacciones y cómo modificando propiedades como su longitud o secuencia se pueden generar efectos diversos en dichas uniones. En su trabajo, Chong y colaboradores estudiaron a Smurf, una ubiquitin E3 ligasa que participa en la degradación de TFG- β y otros blancos. La isoforma Smurf 1 se une al motivo PY de Smad7 –a fin de marcar esta última para degradación- a través de sus 2 dominios WW, primero a partir del WW2 que es esencial y luego por WW1 que estabiliza contactos. Sin embargo encontraron que una forma mas larga de esta isoforma, con el agregado de un segmento desordenado de 26 residuos al “*linker*” entre ambos dominios WW unía solo parcialmente a Smad7, por una afinidad reducida al mismo (Figura I.7 B IV.). Al notar también el grado de conservación evolutiva en las longitudes de estas regiones “*linker*” entre dominios WW, sugirieron que un largo conservado de esta región es algo necesario para la correcta interacción entre este grupo de proteínas [58]. Estas últimas evidencias, junto a muchas otras no expuestas, demuestran que las regiones enlazantes o “*linkers*” desordenados y flexibles que en principio se caracterizaban como meros conectores o espaciadores, pueden cumplir numerosas funciones esenciales en las interacciones.

Se ha visto entonces que las interacciones mediadas por IDPs implican una plétora de mecanismos diversos, desde aquellos en los que se median interacciones a nivel local y monovalentes a interacciones polivalentes de alta afinidad y largo alcance; interacciones que implican procesos de plegamiento-estructuración o aquellos en los que esto no es requerido para mediar especificidad y afinidad y el alto grado de desorden es mantenido tras formarse el complejo, o que implican contactos electrostáticos o hidrofóbicos. La influencia de regiones “*linkers*” que no son pasivos y median distintas funciones, enriquecen (y complejizan) este campo de estudio. Todas las evidencias experimentales que han surgido en los últimos años, por lo tanto demuestran que lejos de aclararse, aun queda mucho por entender acerca de los mecanismos de interacciones mediados por IDPs.

La descripción de las propiedades y ventajas adaptativas y funcionales del desorden intrínseco en las proteínas ha sido presentado. Esto permitiría entender cómo los virus a pesar de su pequeño genoma, son capaces de interferir efectivamente en las redes celulares del hospedador. El modelo de estudio presentado en esta tesis corresponde a la proteína E1A de Adenovirus (AdE1A) y la interacción establecida con la proteína supresora de tumores Retinoblastoma (pRb), reguladora del ciclo celular en eucariotas. *Para contextualizar la presente investigación cabe entonces preguntarse: ¿Qué son los Adenovirus? ¿Qué características posee AdE1A y cómo se relaciona con pRb?*

E. Adenovirus

E.1. Familia *Adenoviridae*

La familia *Adenoviridae* constituye un grupo de virus ubicuos no envueltos, de 70-90 nm de diámetro estructuralmente formados por una cápside externa y un núcleo interno. La cápside es icosaédrica y está formada por 252 capsómeros (240 hexonas y 12 pentonas). Las pentonas tienen asociadas las fibras elongadas que protruyen del virión y le dan su característica morfología (Figura I.8 A). A nivel interno, se encuentra su genoma, una única molécula lineal de ADN doble cadena que se encuentra colapsada con múltiples proteínas similares a histonas. El material genómico tiene un tamaño variable de entre 25 y 70 mil pb, contiene una repetición terminal invertida (ITR por sus siglas en inglés) de entre 36 y 200 pb y en el extremo 5' codifican para una proteína terminal (TP) que funciona como cebador en la replicación viral y asegura que el genoma viral esté adecuadamente replicado [59]. En total codifican de entre 20 a 80 polipéptidos -la mayoría de ellos por un mecanismo de *splicing*- de los cuáles la tercera parte corresponde a proteínas constitutivas del virión [3]. No hay reportes de que contengan lípidos y sólo algunas proteínas se presentan glicosiladas [60, 61]. Esta familia está subdividida en 5 géneros diferenciados por serología, distancia filogenética y organización genómica: *Mastadenovirus*, *Aviadenovirus*, *Atadenovirus*, *Siadenovirus* e *Ichtadenovirus* en los que se encuentran agrupados más de 250 serotipos de adenovirus caracterizados, los cuales infectan a un amplio rango de vertebrados [60].

E.2. Género *Mastadenovirus*

Los *Mastadenovirus* (del griego *mastos*: mama) corresponden a un grupo que sólo infectan a mamíferos. Su ADN doble cadena comprende entre 30 y 38 mil pares de bases (Figura I.8 B). Este grupo codifica para un set de proteínas exclusivas que incluyen: proteína V,

involucrada en el transporte del ADN viral al núcleo de la célula infectada, proteína IX, que actúa como activador transcripcional y re-organizador nuclear y a la mayoría de las llamadas E [60] (de “early”, temprano) transcriptas en la fase temprana de la infección viral (Figura 1.8 B) porque están involucradas en la replicación del genoma y su transcripción [4]. A este género pertenecen los mas de 60 serotipos de Adenovirus identificados que infectan seres humanos [62], los cuáles han sido subdivididos en siete especies: de HAdVA a HAdVG, de acuerdo a sus propiedades biológicas, falta de neutralización cruzada, relación genómica y patología [63].

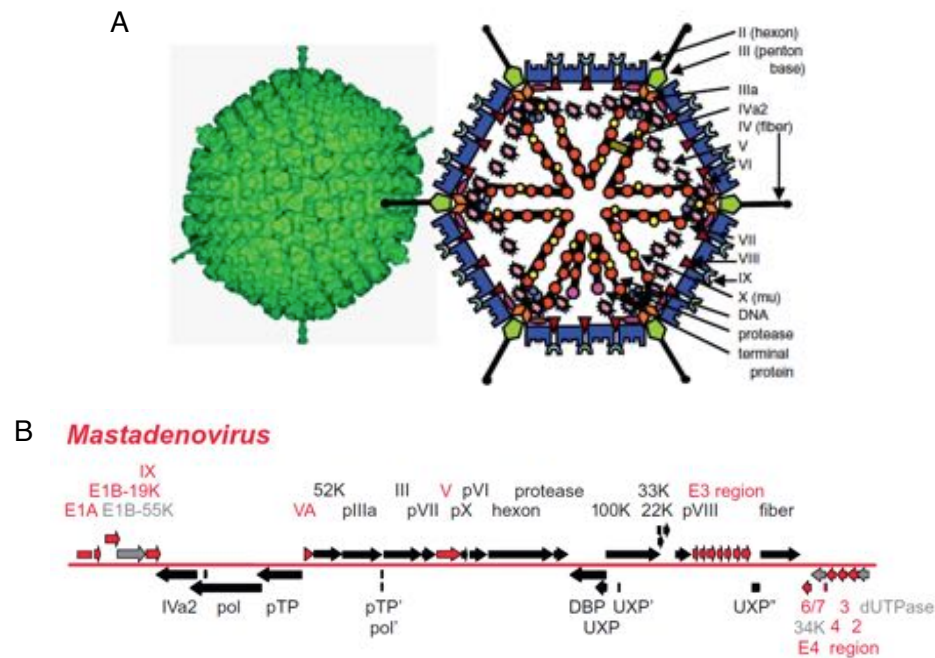


Figura 1.8. Adenoviridae-Mastadenovirus. A) Reconstrucción de una vista por Crio-EM y representación esquemática de un miembro de la familia *Adenoviridae*, género *Mastadenovirus* (del virus HAdV2) mostrando la morfología típica de estos virus. B) Esquema de la organización genómica de un *Mastadenovirus*. Las flechas negras indican genes conservados entre los 5 géneros y las rojas las propias de este género (Adaptado de [60]).

E.3. Ciclo de vida

El proceso de infección de estos virus se inicia cuando utilizan sus fibras elongadas para interactuar con múltiples receptores celulares (como CAR, MCP y CD46) en la célula huésped, lo que resulta en su fijación. La internalización subsiguiente se logra por una endocitosis mediada por integrinas a las que las pentonas se unen [64]. La disolución del endosoma o vacuola fagocítica se produce luego por la actividad tóxica de la pentonas y el virión se libera dentro del citoplasma. Allí, las proteínas de la cápside son digeridas por los lisosomas de la célula huésped y proteasas virales, generando una partícula viral esférica parcialmente descubierta (virión maduro). Éste migra hacia el núcleo a través de los microtúbulos y se trasloca a través de los poros nucleares para liberar la carga genómica [59,

65, 66]. En la fase temprana de la replicación, se transcriben los genes de las regiones E, que permiten la activación de otros genes virales, evitando muerte prematura de la célula infectada y alternado la expresión de las proteínas del hospedador para producir la replicación del ADN viral. La fase tardía del ciclo de estos virus, está centrada en la producción de cantidades suficientes de proteínas estructurales (genes tardíos) para empacar el material genético replicado. Finalmente se produce el ensamble de las partículas y la lisis celular inducida para esparcir los nuevos viriones [66] (Figura 1.9). Este proceso lítico ocurre por ejemplo, en infecciones autolimitadas en células humanas, en los que el virus no presenta actividad transformante lo que implicaría inmortalización celular.

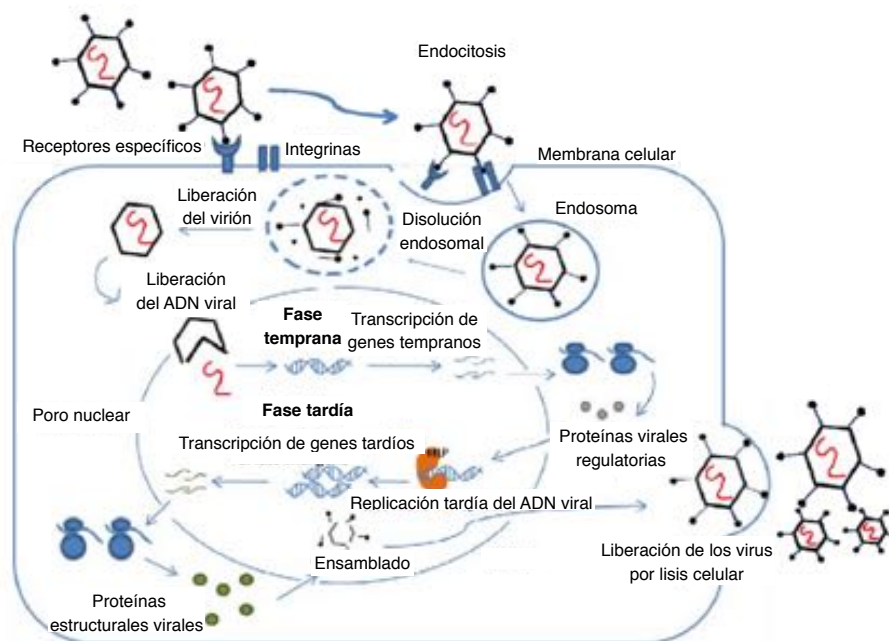


Figura 1.9. Ciclo de vida de Adenovirus. (Adaptado de [66]).

E.4. Patologías asociadas en humanos

Estos virus infectan principalmente células de tejido mucoepitelial, por lo que están involucrados en infecciones del tracto respiratorio superior en niños y adultos, desde del resfrío común a neumonía (fundamentalmente especies HAdVB1, HAdVC y HAdVE). Al persistir en el sistema linfático e infiltrarse en los distintos órganos, incluido la barrera hematoencefálica, también son responsables de enfermedades gastrointestinales como gastroenteritis (HAdVA, HAdVF), oftalmológicas como queratoconjuntivitis (HAdVD y HAdVE), genitourinarias como cistitis hemorrágica aguda (HAdVB2) y neurológicas [61, 67]. Estas infecciones generalmente son autolimitadas en pacientes inmunocompetentes tratados paliativamente y pueden ser fatales en casos de pacientes inmunocomprometidos y menos frecuentemente en individuos

sanos. Si bien en seres humanos no se probado su oncogenicidad, se ha visto que ciertos adenovirus humanos presentan actividad transformante en cultivos *in vitro* de células renales murinas –HadV(especie C)5 y HadV(especie A)12- y que son capaces de generar tumores en hámsters y ratas recién nacidos [68, 69] y ratones inmunocompetentes –HAdA12- [70]. Debido a estas propiedades, se consideran a los adenovirus “virus oncogénicos pequeños a ADN doble cadena” junto a los Poliomavirus (como Virus del Simio 40 -SV40-) y Papilomavirus (como el Virus del Papiloma Humano –HPV-).

F. La oncoproteína viral AdE1A

Como se ha mencionado, las proteínas virales que se transcriben en la fase temprana de la replicación viral de los Adenovirus (proteínas E), entre ellas la proteína E1A (de aquí en adelante AdE1A), son esenciales para la transcripción y replicación del genoma viral. Trabajando sobre el genoma de HAdV5 hacia principios de la década de 1980, se identificó originalmente a AdE1A, como la proteína responsable de que estos virus fuesen capaces transformar células *in vitro* y causar tumores en animales de experimentación [68, 71], la cual presenta una máxima capacidad transformante cuando coopera con E1B y *ras* activado [70]. Estas propiedades la definen como una *oncoproteína*.

El gen HAdV5-AdE1A codifica para 5 isoformas de esta proteína de 289 (13S), 243 (12S), 217, 171 y 55 residuos. Las dos primeras se transcriben inmediatamente producida la infección, en tanto que las otras tres isoformas menores se producen por *splicing* en la fase tardía de la infección [72]. La isoforma mayor contiene 5 dominios funcionales: N-terminal, CR1, CR2, CR3 y CR4 (Figura I.10, Figura I.11 E), dónde CR refiere a regiones o módulos conservados por homología, algo que se había visto entre adenovirus humanos y de simios al inicio de las descripciones de esta proteína [73]. La isoforma 12S, difiere de la primera en 46 residuos que corresponden al CR3 [74]. Parches de secuencias en CR1 y CR2 guardan homología con proteínas como E7 de HPV y LT de SV40. Adicionalmente a estos 5 dominios conservados, AdE1A presenta tres bloques interdominios (IDR1-2, IDR2-3, IDR3-4) que no están igualmente presentes en todas las secuencias de la proteína [75] (Figura I.10). Estas IDRs que en algunas secuencias de AdE1A pueden presentar el mismo nivel de conservación que los dominios CR [75] serían también funcionalmente importantes, como se ha sido probado para IDR2-3 que posee función determinante de oncogenicidad en HAdV12 E1A [70] y la necesaria flexibilidad aportada por la región IDR1-2 que coopera en la formación del complejo ternario AdE1A-CBP/p300-pRb [76].

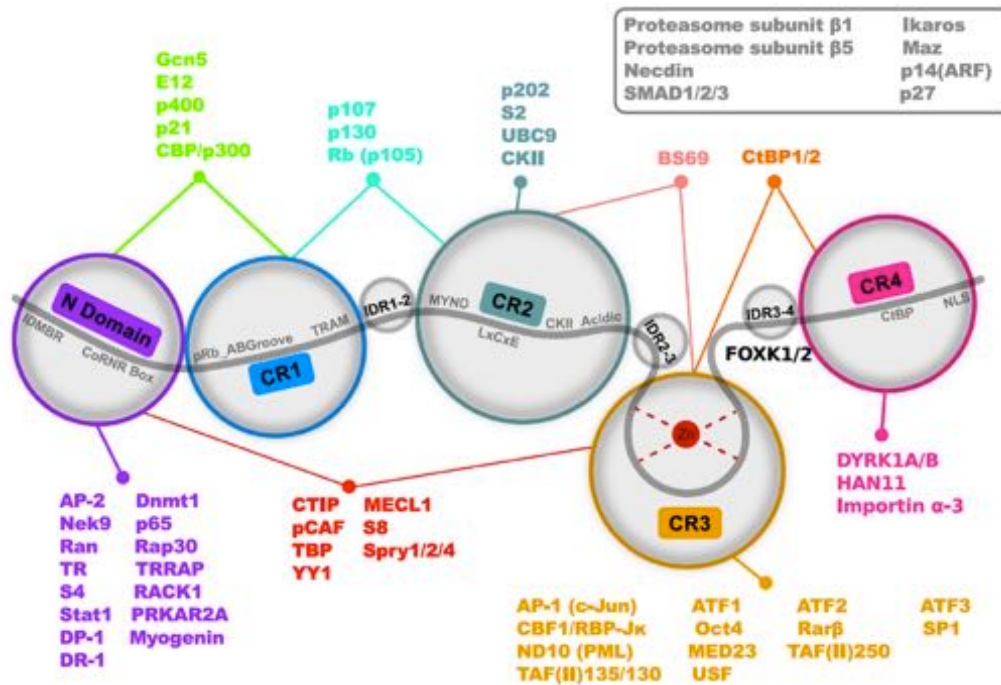


Figura I.10. Representación esquemática de los dominios e interacciones primarias de AdE1A. Los dominios N, CR1, CR2, CR3 y CR4 y regiones interdominios IDR1-2, IDR2-3 e IDR3-4, están señalados en círculos de distintos colores y cuyo tamaño no está relacionado al tamaño de los dominios. La localización aproximada de los SLiMs que posee AdE1A están sobre la línea gris que representa la cadena polipeptídica. Los blancos celulares de AdE1A están agrupados en función de los dominios de la proteína viral con los que interactúan. En el recuadro gris, se encuentran blancos cuyos contactos con AdE1A no han sido mapeados (Adaptado de [75]).

La actividad biológica de AdE1A involucra PPIs con un vasto número de moléculas diferentes del hospedador, lo que lo hace una proteína nodo o “hub” para la infección [77]. Con sus interacciones primarias con más de 50 blancos celulares, AdE1A haría contactos potenciales con mas de 2000 otras proteínas en el hospedador [4]. En contraposición a otros “hubs” AdE1A no posee actividad enzimática o de interacción con ADN y en su lugar actúa de manera modular, en términos de SLiMs [4, 5] (Figura I.10). AdE1A está densamente empacada con SLiMs con los que media sus interacciones con otras proteínas. El dominio N-terminal contiene a: región intrínsecamente desordenada de múltiples interacciones (IDMBR) y al motivo CoRNR Box. En el CR1 se encuentra los motivos: ABGroove -aquí llamado *motivo E2F*- (el cual también está presente en E7 de HPV) y el dominio de interacción a TRAM-CBP. El CR2 contiene el motivo de interacción al dominio MYND y al *motivo LxCxE*, el sitio de fosforilación de Caseína quinasa II (CKII) y el parche ácido, que actuarían como un módulo funcional y están compartidos con E7 de HPV y el antígeno T grande (LT) de SV40. En el CR3, se encuentra el dominio de unión a Zn y sitios de interacción a factores de transcripción y en CR4, el motivo de unión a CtBP y la señal de localización nuclear (NLS) [75].

Funcionalmente, estos motivos le permiten a AdE1A interactuar con blancos celulares que conducen -entre otras actividades- a: inactivar a proteínas reguladoras del ciclo celular (pRb, p130 y p107) permitiéndole tomar control del mismo; secuestrar el transporte mediado por microtúbulos (α -importinas) para relocalizarse en el núcleo donde su presencia es necesaria para otras funciones; alterar la funcionalidad de co-activadores transcripcionales CBP/p300 y otros factores reguladores de la estructura cromatínica (PCAF, BS69), bloqueando diferenciación celular; secuestrar a proteínas adaptadoras o “*scaffolds*” (DCAF7) que anclan enzimas (quinasas DYRK1A/B) relocalizando y manipulando su actividad catalítica, afectando la subsecuente actividad de factores de transcripción (TFs) y por tanto subvirtiendo de la maquinaria transcripcional a su favor y alterando la expresión de genes del hospedador; interferir funciones enzimáticas (RuvBL1, Bre1) conducentes a la evasión de la respuesta inmune mediada por IFN; interactuar con proteasomas (Proteasoma 26S) controlando la abundancia de proteínas del hospedador durante la infección [1, 4]. No obstante al perfil transformante de estas actividades mencionadas, también se ha visto que las mismas interacciones mediadas por SLiMs le permiten a AdE1A actuar en células ya transformadas como anti-oncogén, al suprimir metástasis, angiogénesis y tumorigenicidad *in vivo*, activar apoptosis e inducir diferenciación celular [77], lo que demuestra la complejidad de las interacciones medidas por esta proteína.

A nivel estructural AdE1A no ha sido extensamente caracterizada hasta el momento. No existe estructura cristalográfica de Rayos X disponible para esta proteína, a pesar de intentos de obtenerla [77]. La observación de que AdE1A aun retenía su habilidad para inducir la transcripción y traslocarse al núcleo en células *in vitro* después de haber sido tratada con alta temperatura a ebullición, sugirieron que se trataba de una proteína “desestructurada” [78]. Predicciones hechas en función de su carga neta e hidrofobicidad medias, apuntaron a que esta proteína era “nativamente desplegada” [26] al igual que lo hicieron otros análisis bioinformáticos en función a su secuencia [79, 80].

Las primeras aproximaciones estructurales de esta proteína se han alcanzado por experimentos de Resonancia Magnética Nuclear a partir de estudios de fragmentos de la misma. Los mismos han demostrado que en términos generales AdE1A presenta alto grado de desorden (Figura I.11 A). Por un lado, el dominio N-terminal de HAdV5 AdE1A era predicho como una mezcla entre desestructurado y estructurado por la presencia de una α -hélice [77]. Análisis de RMN, permitieron observar a Ferreon y colaboradores que el segmento entre residuos 13 y 21 presentaba una significativa propensión a α -hélice [76] (Figura I.11 B). Más

tarde, en un estudio independiente Hosek y colaboradores hicieron la misma observación, en donde el N-terminal puebla conformación α -hélice entre los residuos 11 y 22 (Figura I.11 A). Los dominios CR2 y CR4, predichos desordenados, fueron confirmados como tales por estos mismos estudios estructurales en los que usaron secuencias de AdE1A tanto de HAdV5 como de HAdV12 [76, 81] (Figura 1.11 A, B). El dominio CR3, contiene un doble motivo CxxC de unión a Zn -como también está presente en E7 de HPV formando parte de un dominio globular- y el mismo se ha predicho contenía estructura definida [77]. Hosek y colaboradores en su estudio reciente han identificado la presencia de dos α -hélices (entre residuos 153–165 y 170–180) en esta región de AdE1A [81] (Figura I.11 A). En el caso del dominio CR1, también predicho desordenado, ha habido mayores aportes estructurales. En el único reporte de estructura obtenida por Rayos X, Liu y Marmorstein lograron cristalizar el complejo formado por el dominio central de pRb (RbAB, ver mas adelante) junto con el *motivo E2F* viral (entre residuos 41 a 49) (PDBID:2R7G), y observaron que este motivo se encuentra plegado en conformación α -hélice en el estado unido (Figura I.11 C, cinta violeta). En el trabajo de Ferreon y col. antes mencionado, han encontrado por RMN que el CR1 entre los residuos 59 a 78 (motivo TRAM-CBP) se asocia con el dominio TAZ2 de los co-activadores transcripcionales CBP/p300 estructurando α -hélice (principalmente entre residuos 59 a 66) y estableciendo múltiples contactos hidrofóbicos (Figura I.11 D, cinta marrón). Mas adelante, Habertz y colaboradores encontraron que la misma región se pliega en α -hélice también al interactuar con el dominio NCBD de CBP/p300 [82].

En algunos estudios complementarios, también se ha probado que determinadas regiones de AdE1A experimentan transiciones desorden-a-orden en distintas condiciones, como los sitios CtBP y TBP en el CR4, los cuáles adquieren conformaciones giro- β y α -hélice respectivamente en presencia de trifluoretanol [83].

En un primer acercamiento acerca del fragmento de AdE1A utilizado en el presente trabajo, el cual contiene CR1 y al CR2 y la región enlazante o "*linker*" (Figura I.11 E) se puede observar que el mismo se predice desordenado (Figura I.11 G) y presenta parches hidrofóbicos involucrados en interacciones. Además, se ve que se trata de una proteína ácida, con alto contenido de residuos cargados y prolinas (Figura I.11 H).

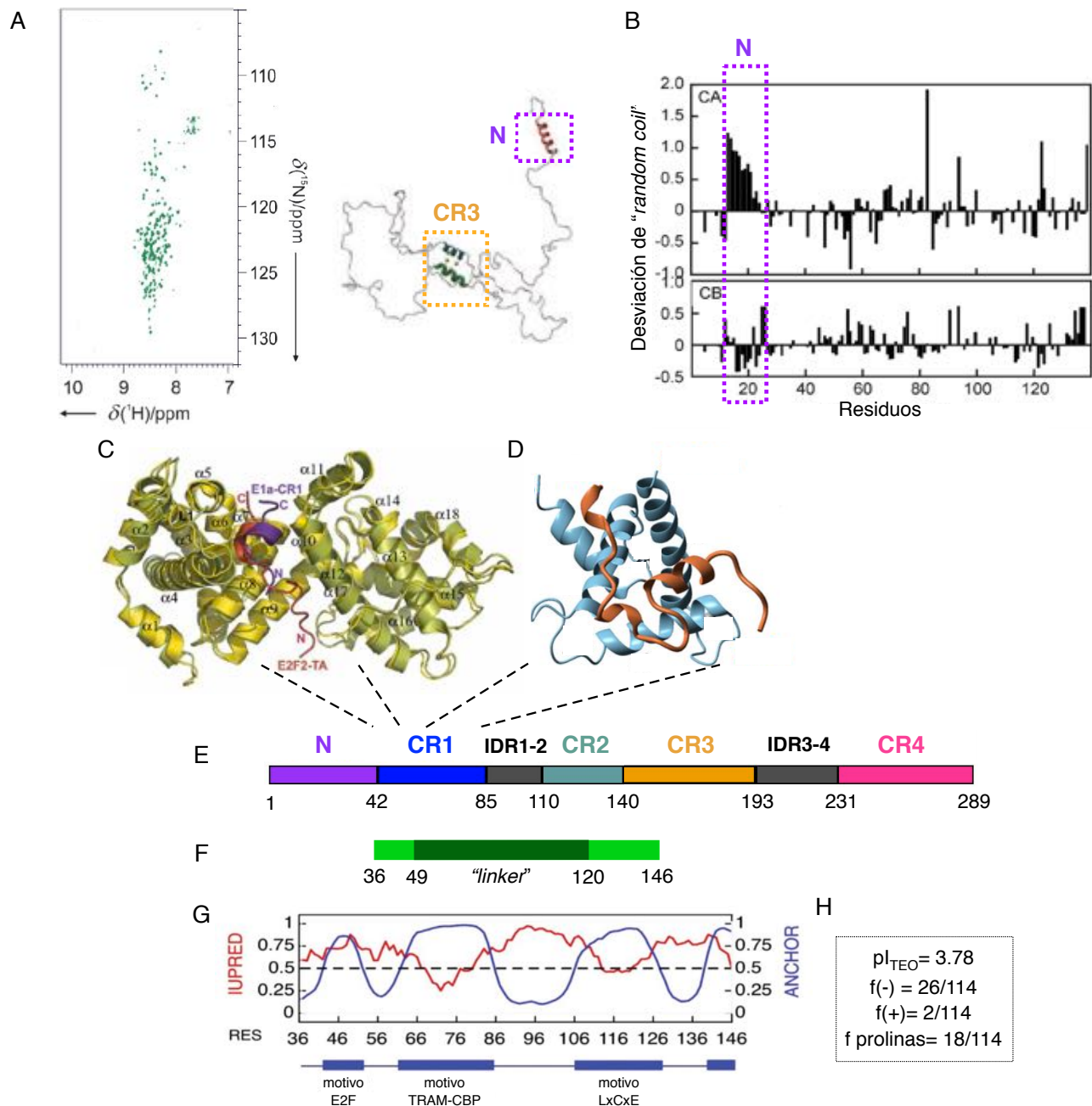


Figura 1.11. Evidencias de estructura, secuencia y predicción de desorden y sitios de contacto de AdE1A. A) Espectro ^1H - ^{15}N BEST-TROSY para AdE1A isoforma 13S (289 residuos) mostrando poca dispersión en la dimensión ^1H (alrededor de 1 ppm) y un ejemplo (de todos los del ensamble conformacional) de AdE1A, con la presencia de α -hélice en el dominio N-terminal (recuadro violeta) y las dos α -hélices del dominio CR3 (recuadro naranja) (Adaptado de [82]) B) Desviaciones de los desplazamientos químicos de $^{13}\text{C}\alpha$ y $^{13}\text{C}\beta$ respecto de valores de "random coil". Se observa la propensión a α -hélice en el segmento N-terminal (residuos 13 a 21) y como el resto es desestructurado en solución (Adaptado de [76]). C) Estructura cristalina del complejo RbAB con AdE1A-CR1 (residuos 41 a 49) –cinta violeta- alineado también con el motivo E2F2 celular –cinta roja-. Se observa como AdE1A está estructurada en α -hélice en la estructura (Adaptado de [84]). D) Representación de más baja energía de los presentes en el ensamble encontrado por RMN, para la estructura del dominio TAZ2 de CBP/p300 unido a AdE1A CR1 (residuos 59 a 78) –cinta marrón-. Se observa α -hélices estructuradas en AdE1A (Adaptado de [76]). E) Representación esquemática de los dominios y regiones interdominios presentes en AdE1A (HAdV5), los cuáles mostrando los límites de cada uno. F) Representación del fragmento de AdE1A (HAdV5), utilizada en la presente tesis. Los límites de la región "linker" de 71 residuos (49-120) –en verde oscuro- están definidos por el residuo final del motivo E2F (en CR1) y el residuo inicial del motivo LxCxE (en CR2). G) Predicción del desorden de la secuencia de AdE1A (HAdV5) utilizada en este trabajo, a partir del algoritmo IUPRED [85], dónde los valores por encima de 0.5 indican desorden y valores menores, estructuras ordenadas –línea roja- y predicción de presencia de parches hidrofóbicos en la secuencia utilizando el algoritmo ANCHOR [85], donde valores por encima de 0.5 indican la presencia de los mismos –línea azul-. Como se puede observar, coinciden con los motivos E2F, TRAM-CBP y LxCxE presentes en la secuencia de AdE1A los cuáles median interacciones con distintos blancos. H) propiedades de secuencia de AdE1A, la cual se

trata de una proteína ácida, con alto contenido de residuos cargados (+, -) y prolinas. Nótese que el número total de residuos de la secuencia (114 res), no coincide con los límites 36-146, esto se debe a cómo fue construido el fragmento (ver Capítulo I).

G. La relación entre AdE1A y las proteínas pRb, p107 y p130

La identificación del rol transformante de AdE1A y de su interacción con este grupo de proteínas están ligadas históricamente. Para la década de 1960 ya se había descrito la capacidad transformante *in vivo* e *in vitro* de Adenovirus humanos sobre células y animales de experimentación y alrededor de dos décadas después se asignó este rol oncogénico a AdE1A [68]. En una serie de estudios trabajando con AdE1A con residuos mutados entre las posiciones 120 y 150 -particularmente la mutación C124G- se encontró que su actividad transformante era defectiva [86]. Además, estudios de inmuno y co-precipitación llevaron a la observación de que AdE1A une a una proteína de 105kDa [87] y que los residuos en las posiciones 111 a 127 así como los residuos de 30 a 60 en AdE1A, son necesarios para esta interacción [88] lo que sugirió que los mecanismos transformantes se daban por alguna disfunción en las proteínas celulares y no necesariamente por una nueva función introducida por la proteína viral [89]. Contemporáneamente, pRb fue descrita como una proteína nuclear cuyo peso es 110 kDa [90] y se confirmó que esta se trataba de la proteína que co-precipitaba específicamente con AdE1A [91], siendo la primer demostración de la asociación entre estas proteínas. Posteriormente se logró identificar que AdE1A también interactuaba con las proteínas p107 [92] y p130 [93] las cuales exhiben características similares a pRb y forman parte de la misma familia. *Ahora bien ¿Por qué sería tan importante esta asociación de AdE1A con pRb y los otros miembros de su familia? ¿Qué funciones cumple esta familia de proteínas para hacerlas un blanco “atractivo” para la proteína viral?*

H. La proteína supresora de tumores Retinoblastoma y las proteínas p107/p130

La proteína supresora de tumores Retinoblastoma (pRb) es un regulador evolutivamente conservado del ciclo celular, de la diferenciación, apoptosis y estructura cromatínica en eucariotas [94, 95]. pRb y las proteínas relacionadas p107 y p130, forman parte de la familia de las proteínas “*pocket*”, que son conocidas por su control del ciclo celular principalmente a través de su interacción con la familia de factores de transcripción E2F, a los que bloquean su actividad transcripcional, reprimiendo así la expresión de genes involucrados en la progresión del ciclo celular, desde la fase G1 a la fase S [96, 97] (Figura I.12). pRb y p107/p130 presentan funciones solapadas, aunque pueden reprimir distintos grupos de genes debido a sus interacciones selectivas con distintos miembros de la familia E2F [98]. La represión de E2F ejercida por pRb (y por caso de p107/p130), es controlada a través de proteínas quinasas

dependientes de ciclinas (CDKs) las que median su fosforilación. Estudios estructurales revelaron que la fosforilación induce interacciones intramoleculares en pRb, que conllevan a la disociación del complejo que forma con E2F [99], lo que ocurre en respuesta a señales mitogénicas, y conlleva a la liberación del factor de transcripción, el cual induce los genes de la Fase S y la progresión del ciclo celular (Figura I.12). Las proteínas “*pocket*” también regulan la progresión del ciclo celular a través de mecanismos independientes de E2F como la vía pRb-SKP2-p27 la cual conduce a estabilización de p27, inhibición de quinasas dependientes de ciclinas y arresto del ciclo celular [97].

Este grupo de proteínas controla la especificación del linaje celular durante el desarrollo temprano, la diferenciación celular terminal, y el retiro del ciclo celular por ejemplo a través del ensamblado del complejo DREAM, el cual involucra a p130 y p107 en simultáneo y las proteínas MuvB. pRb además controla las respuestas apoptóticas frente a daños genotóxicos y la regulación de la estructura de la cromatina [95].

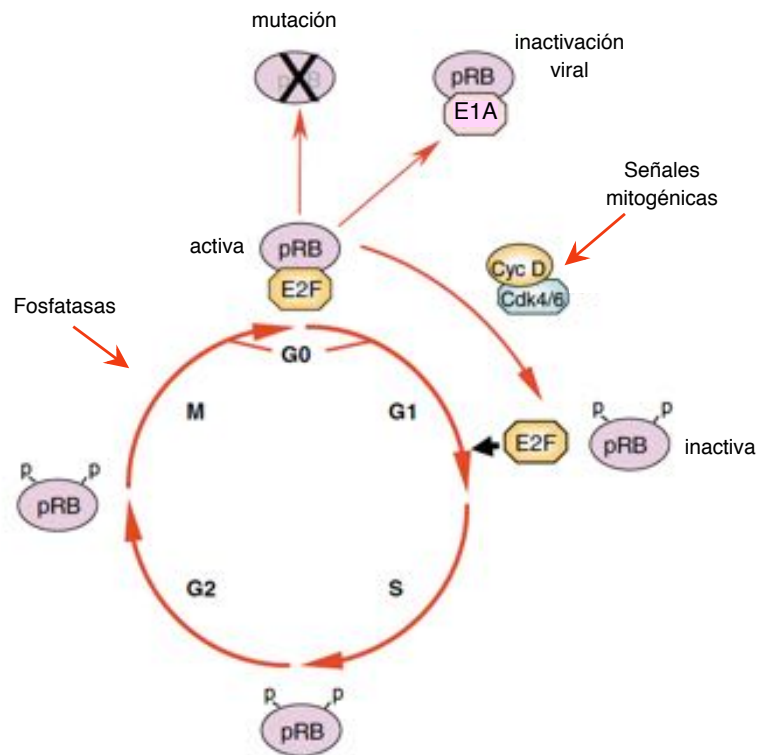


Figura I.12. Mecanismo de regulación del ciclo celular por acción de la pRb. Cuando pRb se encuentra desfosforilada (activa), reprime las funciones de E2F. Señales mitogénicas desencadenan cascadas de fosforilación que conducen a la disociación del complejo pRb-E2F y permiten el ingreso a la fase S, cuando E2F trasloca al núcleo y activa los genes de dicha fase. pRb se torna nuevamente activa, al desfosforilarse por acción de fosfatasas. En células tumorales, la interacción pRb-E2F se interrumpe por la mutación del gen RB (X) o por la unión de pRb a una oncoproteína viral AdE1A del Adenovirus E1A –o también E7 de HPV o LT de SV40- (Adaptado de [97]).

A nivel estructural pRb se compone de tres dominios definidos incluyendo dos dominios con doble motivo de ciclina (RbN y RbAB) y el dominio C-terminal desordenado (RbC) (Figura I.13 A).

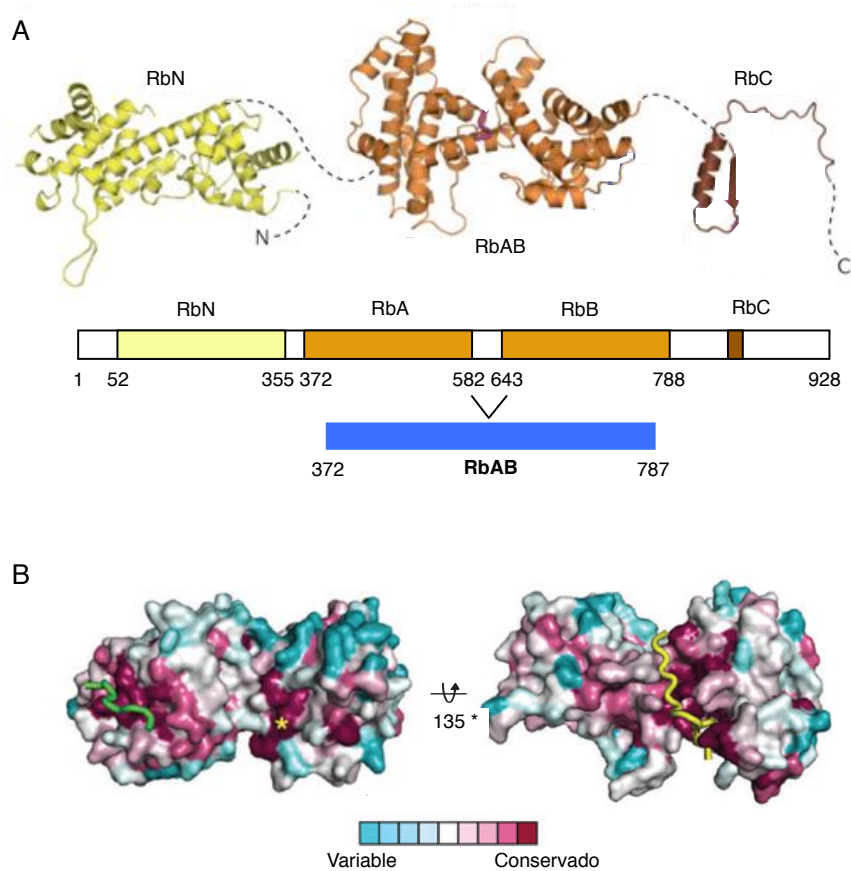


Figura I.13. Estructura, superficies de interacción y grado de conservación de superficies de pRb. A) Estructura de pRb con sus 3 dominios (RbN, RbAB, RbC) y las regiones linker (líneas punteadas) (Adaptado de [95]). Abajo, una representación esquemática de los dominios y sus límites, las partes que presentan estructura consolidada están coloreadas, las otras corresponderían a regiones menores desordenadas en la proteína. En la tercer línea, representación esquemática del dominio RbAB utilizada en la presente tesis; nótese que es una construcción a la que le falta la región desordenada central (Ver Capítulo 1). B) Grado de conservación de residuos en el dominio central RbAB. Las zonas más oscuras son las más conservadas y corresponden al *surco LxCxE* (izquierda) y al *sitio E2F* (derecha). La línea verde muestra al *motivo LxCxE* (HDAC, Ciclina D, E7 de HPV, AdE1A) que interactúa con el *surco LxCxE* y la línea amarilla muestra *motivo E2F* (factor de transcripción E2F, AdE1A) que interactúa con el *sitio E2F* de RbAB (Adaptado de [100]).

Además pRb contiene 16 sitios consenso de fosforilación para quinasas dependientes de ciclinas a lo largo de estos dominios [101]. El dominio N-terminal –RbN- (residuos 52 a 355) está implicado en funciones de desarrollo celular. El dominio intrínsecamente desordenado C-terminal –RbC- (residuos 788 a 928) tiene un papel esencial en el arresto del ciclo celular, mientras que el dominio central RbAB (residuos 372 a 787) es el dominio funcional más importante y es necesario para la mayoría de las funciones de pRb [102]. El dominio RbAB, formado por un doble motivo de ciclinas alberga dos superficies conservadas: una región llamada *surco LxCxE* en el sub-dominio B y una superficie de unión para el factor de

transcripción E2F, llamado *sitio E2F* o surco A/B complementario a E2F, ubicado en la región interdominio A-B de pRb [99] (Figura I.13 B). Estos dos sitios son esenciales en pRb porque se unen complementariamente a SLiMs presentes en diversos blancos para mediar PPIs (Figura 13 B) [37]. Los SLiMs que participan directamente en las interacciones con pRb en estos sitios son: el "*motivo LxCxE*" que está presente en alrededor de 20 blancos celulares de pRb como HDAC 1-2, ciclina D y RBP2 y blancos virales como E7-HPV, AdE1A, LT-SV40; y el "*motivo E2F*", existente en el factor de transcripción E2F y en el CR1 de AdE1A.

I. Desregulación de la función de pRb por acción de oncoproteínas virales: AdE1A

Las descripciones presentadas para pRb demuestran que la misma constituye una proteína "*hub*" o nodo de una extensa red de interacciones a nivel celular. Todas las funciones que pRb cumple están dadas por PPIs con más de 100 proteínas celulares [103], con las que interactúan a través de sus dominios (interacciones dominio-dominio) o mediadas por sus superficies conservadas que interactúan con SLiMs complementarios.

Por otro lado, se sabe actualmente que la inactivación funcional de pRb está implicada en la iniciación y progresión de muchos cánceres humanos [95, 104]. Estudios recientes revelaron que la pérdida de funcionalidad de pRb por mutaciones o interacciones con proteínas virales promueven el re-ingreso de las células post-mitóticas a la fase S, causando inestabilidad cromosómica y favoreciendo la aparición de mutaciones en las células afectadas [105]. Los eventos moleculares conducentes a la pérdida de funcionalidad de otros supresores de tumores clínicamente relevantes como p53 [95, 106], son más conocidas. Pero *¿Cuáles son los mecanismos moleculares por los que pRb pierde su función en el caso de las inactivaciones por oncoproteínas virales?*

Varias proteínas virales son capaces de interactuar e inhibir a pRb: AdE1A, E7 de HPV, LT de SV40, entre otras [106] en algunos casos con capacidad oncogénica y efectos transformantes, como en el cáncer cervical provocado por HPV. Se ha observado que esto se produce, como ya se mencionó, a través de la *mímica* de los SLiMs existentes en el hospedador, en este caso del "*motivo LxCxE*" y en algunos casos del sitio adicional "*motivo E2F*", los cuáles interactúan con el dominio central de pRb (RbAB) para lograr el desplazamiento efectivo de E2F y desregular la vía de control del ciclo celular [107] (Figura I. 14). Esto sería esencial pues es *uno de los mecanismos primarios por el cual estas proteínas secuestran la maquinaria celular a su favor*.

En el caso de E7-HPV, esta posee un CR1 con homología secuencial al CR1 de AdE1A, y se ha visto que interactúa con el dominio RbAB pero con afinidad baja, mucho menor a la que establece *motivo LxCxE* en CR2 con RbAB. Además, cuando E7-CR2 está unido a RbAB, la contribución del CR1 es casi nula, probablemente porque el “*linker*” de 8 residuos que separa CR1 de CR2 es muy corto, impidiendo interacción simultánea de estas regiones en una misma molécula con RbAB. Estas observaciones junto con el hecho de que E7 puede dimerizar, ha llevado a proponer que el CR2 de una molécula de E7 uniría por su *motivo LxCxE* al *surco LxCxE* complementario en RbAB mientras que la región CR1 de la segunda molécula E7 se uniría al *sitio E2F* de RbAB [100]. Este modo de interacción, un motivo (*LxCxE*) de alta afinidad a RbAB y un probable efecto cooperativo mediado por dimerización, llevaría al desplazamiento efectivo del factor de transcripción E2F (Figura I.14 A, Figura I.12). Además, CR3 también podría colaborar en este desplazamiento probablemente por un sitio adicional de unión o efecto de la dimerización.

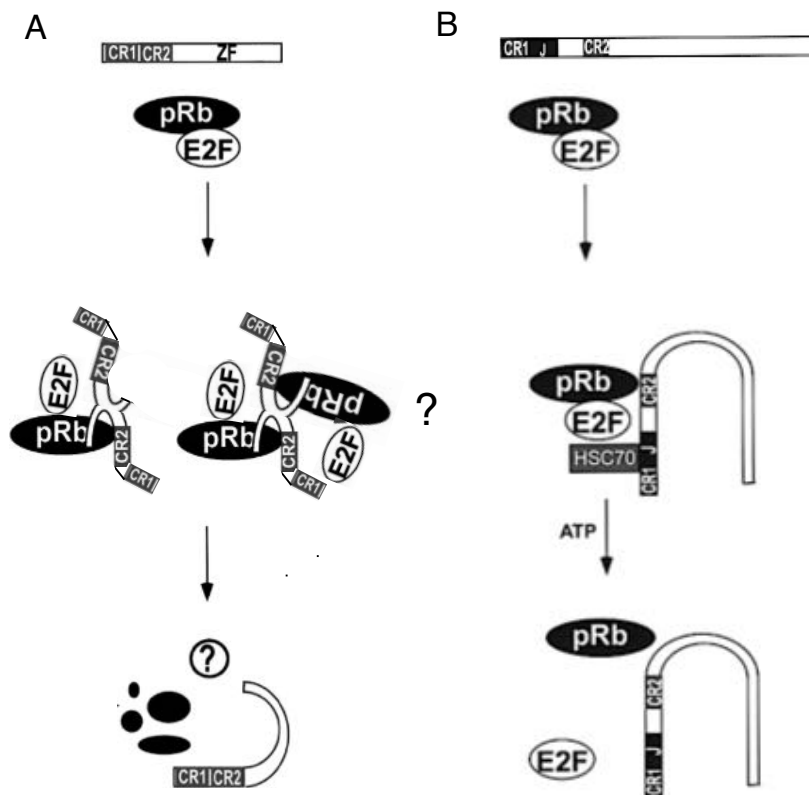


Figura I.14. Mecanismo de interacción con RbAB e interrupción del complejo RbAB-E2F por las oncoproteínas E7-HPV y LT-SV40. A) E7 interactúa por su motivo LxCxE de alta afinidad a RbAB y luego tras una probable dimerización, el CR2 de otra molécula de E7 colaboraría a disociar el complejo RbAB-E2F. B) LT-SV40 une a RbAB por el mismo motivo LxCxE, pero luego recluta HSC70 que hidroliza ATP, lo que genera el desensamblado del complejo RbAB-E2F (Adaptado de [107]).

El mecanismo por el cual LT-SV40 interrumpe el complejo E2F-RbAB para tomar control del ciclo celular, es otro. En este caso, LT-SV40 que presenta también un CR1 y CR2 con homología de secuencias a los de E7-HPV y AdE1A, utiliza el *motivo LxCxE* para unir a RbAB y un dominio J que interactúa con la co-chaperona HSC70. La interacción de E2F con RbAB se pierde, cuando HSC70 media una hidrólisis de ATP que dispara el desensamblado del complejo, subvirtiendo así el ciclo celular a su favor [107]. (Figura I.14 B, Figura I.12).

AdE1A interactúa con el dominio central RbAB a través de dos sitios de interacción independientes, los SLiMs ya identificados: *motivo E2F* ubicado en CR1 y el *motivo LxCxE* ubicado en el CR2 (Figura I.10) [108]. El *motivo E2F* interactúa con RbAB en el *sitio E2F* [108] -el mismo en el cual RbAB une el factor de transcripción E2F- y existe una estructura cristalográfica de dicha asociación (PDBID 2R7G) (Figura I. 11 C.). En tanto el *motivo LxCxE* interactúa con el *surco LxCxE* de RbAB [76].

Ambos motivos actuarían uniendo a RbAB de manera conjunta para desplazar efectivamente a E2F del complejo RbAB-E2F [108] (Figura I.15 A, Figura I.12). En esta inactivación de RbAB, también estaría involucrada la región “*linker*” -larga, flexible y desordenada- que separa ambos motivos en la cadena carbonada de AdE1A (Figura I.11 F), la cual sería requerida para permitir que ambos motivos actúen sinérgicamente para desplazar a E2F en mecanismos posibles como están planteados en la Figura I.15 B.

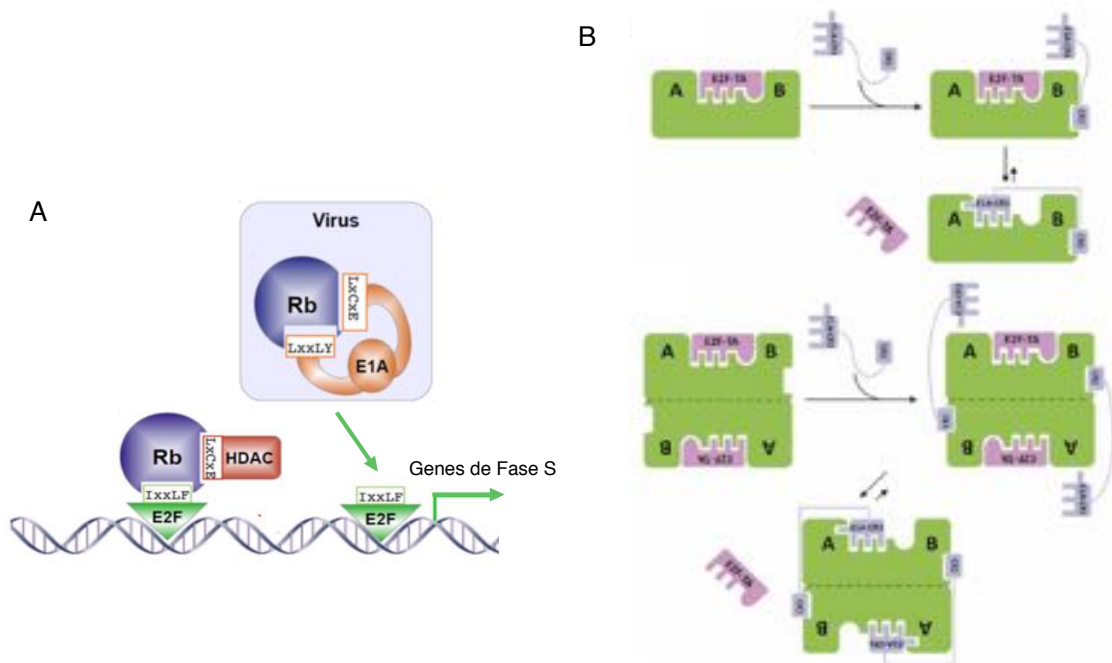


Figura I.15. Mecanismo de competencia y desplazamiento del complejo RbAB-E2F por acción de AdE1A. A) Representación esquemática de cómo la competencia con E2F e inactivación de RbAB, conduce a AdE1A tomar control del ciclo celular. B) Mecanismos probables de interacción y desplazamiento de E2F del complejo con RbAB en el que el “*linker*” juega un rol clave (Adaptado de [84]).

En este mecanismo el *motivo LxCxE* permitiría el “anclaje” a RbAB, mientras que el segundo sitio -el *motivo E2F*- competiría directamente con el factor celular E2F por la unión a RbAB desplazándolo, lo cual sería favorecido por el “*linker*” (Figura I.15 B, arriba). Alternativamente, investigadores como Liu y Marmorstein han planteado que este mecanismo de desplazamiento podría ocurrir tras producirse la interacción del motivo LxCxE de AdE1A con una molécula de pRb mientras que el motivo E2F uniría a una segunda molécula de pRb, mediante un “*linker*” flexible que permita esa interacción [84] (Figura I. 15 B, abajo). No obstante hasta el comienzo de esta tesis, no se contaba con información mecánica, cuantitativa ni datos estructurales que describan al complejo de estos dos SLiMs unidos por la región “*linker*” y cómo esto produce el desplazamiento del factor E2F.

Por otra parte, la región “*linker*” contiene otros motivos funcionales como el motivo TRAM-CBP que le permiten formar complejos ternarios, por ejemplo pRb-AdE1A-CBP/p300 [76, 82]. Esto llevó a autores como Ferreon y Wright [76] a plantear que existirían restricciones sobre el largo y flexibilidad del “*linker*” que permiten la interacción simultánea de ambos motivos y la unión de otras proteínas. Sin embargo, antes de esta tesis tampoco se contaba con un modelo que de cuenta del rol del “*linker*” en este complejo. Surgen preguntas como: *¿Cuáles son los requerimientos de este “linker” y cómo se comporta en la interacción conjunta de ambos sitios? Esta interacción ¿se encuentra optimizada evolutivamente? ¿Hay propiedades adicionales del “linker” que permitan la unión simultánea de otras proteínas, colaborando a la función “hub” de la proteína viral AdE1A?*

De forma más global, se plantea: *¿puede el estudio estructural y biofísico de este complejo echar luz sobre el aporte del desorden intrínseco en proteínas virales a su función “hub” o nodo?* Entender estos aportes es importante a fin de develar cómo los virus logran secuestrar y tomar control de la maquinaria celular a su favor. En esta tesis, se toma a la interacción AdE1A-RbAB como sistema modelo para el análisis de los mecanismos de subversión viral de la célula huésped.

HIPOTESIS Y OBJETIVOS

Los mecanismos moleculares de las interacciones virus-hospedador, que conducen a estos agentes patógenos a interferir efectivamente con la red celular, secuestrar y tomar el control de la maquinaria de la célula huésped a su favor, aún se encuentran lejos de ser completamente comprendidos.

Virus patogénicos como Adenovirus, Papilomavirus y Poliomavirus poseen proteínas que son capaces de interactuar con proteínas esenciales en el funcionamiento celular, como el supresor de tumores Retinoblastoma –pRb-, un regulador eucariótico central del ciclo, desarrollo y diferenciación celular [94, 95]. Dichas interacciones conducen a una desregulación en la célula infectada y en algunos casos esto está asociado a transformación celular y oncogénesis [104]. La oncoproteína viral AdE1A de Adenovirus se une a pRb a través del *motivo E2F* y el *motivo LxCxE* los cuáles se encuentran separados por una región de 71 residuos o “*linker*”. Estas interacciones que establece AdE1A permiten inactivar a pRb fundamentalmente por una competencia altamente eficiente con las interacciones celulares, como la del factor de transcripción E2F [76, 109]. No obstante, hay un limitado conocimiento mecanístico y estructural con respecto de esta interacción. Esto lleva a preguntarse: *¿qué propiedades estructurales-conformacionales posee AdE1A que le permiten interactuar eficientemente con pRb? ¿en qué medida contribuyen y cómo operan entre sí los motivos E2F y LxCxE? ¿cuál la contribución de la región “linker” en esta interacción? ¿Cuáles son los factores determinantes en la proteína que permiten desplazar de modo efectivo al factor E2F y así tomar control del ciclo celular?*

Evidencias experimentales recientes indican que la presencia de múltiples SLiMs contenidos en IDPs les confiere la capacidad de unirse a una plétora de blancos diferentes [76] y que varios estos SLiMs pueden trabajar sinérgicamente para unir a único blanco con alta afinidad [54]. El alto grado de desorden que puede prevalecer en la IDP aún cuando se forma el complejo, podría ser clave en sus roles regulatorios permitiéndoles mediar otros contactos simultáneamente [21]. Otras evidencias indican que habría gran información codificada en regiones intrínsecamente desordenadas –IDRs-, con influencia directa en sus propiedades conformacionales y que modularían las funciones y afinidad de SLiMs en ellas contenidas [30, 110]. También se ha observado que en sistemas bivalentes de interacción, ciertas características de las regiones “*linker*” [55, 111-113] serían factores clave que contribuyan a la interacción y afinidad global de estos sistemas.

En base a lo expuesto, la hipótesis de este trabajo es la siguiente: ***La oncoproteína viral AdE1A es capaz de establecer una interacción de alta afinidad con pRb, utilizando información codificada en su secuencia y en su estructura-***

conformación, que le permite competir efectivamente con el complejo endógeno pRb-E2F. AdE1A utilizaría una estrategia en la que sus SLiMs operan sinérgicamente en la unión a pRb, enlazados por un “linker” que mediante propiedades conservadas evolutivamente y relacionadas a su carácter “IDP” favorecen dicha interacción. En conjunto, estas propiedades permiten a AdE1A tomar control del estado replicativo de la célula huésped.

Para poder testear esta y otras hipótesis menores, se utilizará un abordaje experimental de carácter biofísico y estructural, mediante una combinación de múltiples experimentos *in vitro*. El estudio de estas interacciones incluirá un trabajo de caracterización cuantitativa de afinidades y experimentos estructurales de media y alta resolución. Si bien estos experimentos *in vitro* pueden no reflejar el exacto comportamiento *in vivo*, los mismos constituyen una herramienta poderosa al permitir condiciones controladas que brindan información precisa necesaria para esclarecer el funcionamiento de la proteína AdE1A, el rol de regiones desordenadas y su mecanismo de inactivación de pRb.

Objetivos generales. El objetivo general de este proyecto es develar a nivel mecanístico y estructural las bases moleculares subyacentes a las interacciones virus-hospedador. Para esto se utilizará como modelo la oncoproteína viral intrínsecamente desordenada AdE1A y su interacción con el supresor de tumores eucariótico pRb. La comprensión de estos mecanismos contribuirá a obtener conocimientos que permitan entender en mayor profundidad cómo los virus secuestran la maquinaria celular en su propio beneficio y allanará el camino para el desarrollo de estrategias terapéuticas contra infecciones y cáncer de etiología viral.

Objetivo específicos

Objetivo específico 1: Generación recombinante y caracterización bioquímica y biofísica de AdE1A y mutantes de motivos.

En primer lugar, dado que se quiere estudiar a AdE1A y su interacción con RbAB, se buscará generar un fragmento de la oncoproteína viral que contenga todos los sitios esenciales para dicha interacción, y producir 2 variantes mutantes de motivo, que permitan evaluar contribuciones diferenciales en la interacción. En segundo lugar, dado que no existe en la literatura un protocolo de expresión y purificación de AdE1A a escala que se adapte a las disponibilidades técnicas del laboratorio, se propone desarrollarlos a fin de obtener cantidades suficientes de proteínas para experimentos posteriores. Por otro lado, dado la falta de caracterizaciones biofísicas/bioquímicas en AdE1A, se evaluarán y caracterizarán sus

propiedades en estado libre y solución en condiciones controladas, para empezar a entender cómo funciona esta proteína.

Objetivo específico 2: Análisis de la interacción entre AdE1A y su blanco específico RbAB.

Aquí se propone en principio, entender cuál es el aporte de los SLiMs *E2F* y *LxCxE* en la interacción con pRb, a nivel de los residuos que los componen, realizando una revisión de datos experimentales disponibles en la literatura específica. Luego se analizará la interacción de la proteína AdE1A con RbAB, para definir estequiometrías y propiedades hidrodinámicas y conformacionales del complejo. Luego se procederá a hacer un análisis cuantitativo/termodinámico detallado de las interacciones entre AdE1A con RbAB por métodos que incluyen ensayos de Fluorescencia y Calorimetría de Titulación Isotérmica, para revelar contribuciones energéticas específicas a fin de construir un modelo para la interacción que integre los diferentes sitios de contacto entre ellos. Por último se evaluará la capacidad de AdE1A para desplazar E2F2 de su interacción con RbAB y develar los factores determinantes de esta competencia.

Objetivo específico 3: Estudio en alta resolución de AdE1A libre y en complejo con RbAB por Resonancia Magnética Nuclear (RMN).

Este objetivo incluye el análisis estructural de alta resolución de AdE1A y las mutantes de motivos por RMN, buscando analizar propensiones y dinámicas conformacionales y flexibilidad en sus estados nativos libres en solución, tendientes a comprender cómo influyen estas propiedades en la interacción con RbAB. Por otro lado, se analizará el complejo [RbAB:AdE1A], mapeando los residuos involucrados en la interacción y evaluando la contribución de la región “*linker*”, para validar a nivel estructural el modelo de interacción derivado de los datos termodinámicos.

Objetivo específico 4: Construcción de un modelo de interacción y análisis del complejo RbAB-AdE1A por Dispersión de Rayos X de ángulo pequeño (SAXS).

Por último, tomando en conjunto todos los datos previos obtenidos, se propone presentar un modelo de interacción entre AdE1A y RbAB que integre las contribuciones específicas energéticas y estructurales de los diferentes sitios de contacto y el rol de la región “*linker*” en la formación del complejo. Para ello, se planteará primero modelados teóricos de los complejos estudiados y evaluar la consistencia de los mismos a partir de experimentos SAXS sobre el complejo proteico pRb-AdE1A, para cerrar con un modelo final de interacción.

CAPITULO I

**Clonado, expresión y caracterización bioquímica y biofísica de AdE1A y
mutantes de motivo**

Con el fin de entender las funciones que media una proteína, primero es necesario comprender *cómo se comporta de manera aislada, cuáles son sus propiedades y su estructura y/o conformación*. Para ello, los estudios tendientes a estas caracterizaciones requieren de poder *disponer de cantidades suficientes de la proteína* de interés.

En primer lugar, dado que en lo sucesivo se quiere estudiar a AdE1A y su interacción con RbAB, se buscará generar un fragmento de la oncoproteína viral que contenga todos los sitios esenciales para dicha interacción: el motivo E2F, la región “*linker*” y el motivo LxCxE. También se buscará producir 2 variantes mutantes de motivo en la que se inactive uno de los dos motivos mientras el otro permanece intacto.

Para el caso de la proteína AdE1A, una macromolécula con un peso molecular considerable, su obtención de manera sintética resulta costosa y por tanto, inviable. Si bien ha habido grupos de investigación que han logrado expresar diferentes construcciones de AdE1A [76, 81, 114], *ninguno ofrece un protocolo de expresión y purificación detallado y completo que se adapte a las disponibilidades técnicas de este laboratorio*. Además, algunos de ellos encontraron dificultades en los procesos de purificación o en los experimentos posteriores trabajando con concentraciones altas de estos reactivos, por lo que aplicar sus procedimientos podría resultar desventajoso y perjudicial para algunos experimentos del presente trabajo. Con el objetivo de generar los reactivos adecuados se llevarán a cabo experimentos de biología molecular que implican la correcta manipulación del material genético, desde el sub-clonado en plásmidos de expresión hasta la mutagénesis para obtener variantes mutantes de interés. Posteriormente, se pondrán a punto sistemas de expresión recombinante claves para la obtención de cantidades necesarias y suficientes de AdE1A y mutantes de interés para estudios posteriores.

Por otro lado, para AdE1A existe en la literatura *escasos estudios conducentes a entender su comportamiento y propiedades en el estado libre* [83, 114]. Es evidente, por lo tanto que *hacen falta aportes en la caracterización de AdE1A, necesarias para luego poder profundizar el estudio en sus funcionalidades*. Para lograr caracterizar el comportamiento y propiedades de AdE1A libre en solución: propiedades hidrodinámicas, contenido de estructura secundaria, tendencias conformacionales, estabilidad *in vitro* y termoestabilidad, se llevarán a cabo una serie de experimentos de carácter biofísico-bioquímicos.

OBJETIVOS ESPECIFICOS:

- I. Generación de los reactivos de trabajo a ser utilizados en esta tesis, en particular la proteína viral AdE1A en su variante natural y dos mutantes de motivos. En particular:
 - a. Obtener por biología molecular el fragmento de interés de la proteína AdE1A, clonándolo en vectores de expresión procariota.
 - b. Generar mutantes sitio-específicas de AdE1A para inactivar uno de los dos motivos de interacción a RbAB mientras que el otro permanece intacto, a fin de evaluar la contribución de cada motivo y de la región “*linker*” en la interacción con RbAB en experimentos posteriores.
 - c. Desarrollar un protocolo de expresión recombinante y purificación para la obtención a escala de mg de AdE1A variante natural y mutantes de sitio para la realización de experimentos de caracterización biofísica y estructural.
- II. Evaluación y caracterización con ensayos bioquímicos y biofísicos de la estabilidad y propiedades estructurales-conformacionales de la proteína AdE1A en solución. Es decir:
 - a. Analizar la estabilidad de AdE1A en las muestras *stock*.
 - b. Estudiar las propiedades hidrodinámicas y estado de oligomerización de la proteína en solución.
 - c. Caracterizar el contenido y propensiones de estructura secundaria de AdE1A.
- III. Obtención de los reactivos con marca fluorescente necesarios para ensayos de espectroscopía de fluorescencia, desde su marcación y purificación a su identificación.

OBJETIVO I.1

Generación de AdE1A.WT y mutantes de motivo

En primer lugar, se generó un fragmento de AdE1A que contenga todos los sitios esenciales para su interacción con RbAB y 2 variantes mutantes de motivo, que en experimentos posteriores permitan evaluar contribuciones diferenciales en la interacción. Para todo esto se utilizaron técnicas de biología molecular para producir sub-clonados en plásmidos de interés y aplicación mutagénesis sitio-dirigida.

I.1.A. Biología molecular del gen Ad5E1A

Clonado de AdE1A en vectores de expresión

En este trabajo se utilizó un fragmento de la proteína E1A de Adenovirus humano C serotipo 5 (Ad5E1A) (Uniprot ID P03255) el cual fue provisto generosamente por el Dr. Ronen Marmorstein (del *Wistar Institute* y *University of Pennsylvania*, Filadelfia, EEUU) clonado en un vector pRSET (“pRSET- E1A_{CR1-3}”, aminoácidos 36 a 189) [114]. El mismo sirvió como molde para sub-clonar un fragmento de la secuencia natural de la proteína AdE1A (como se llamará en adelante) en dos vectores de expresión: pMalC2x y pRSETa.

El vector pMalC2x (New England Biolabs) posee un codón de iniciación (ATG) seguido por el gen *malE* que codifica para la Proteína de unión a Maltosa (MBP), un sitio de corte por el Factor Xa y el sitio de clonado múltiple. El vector pRSETa (Invitrogen), por su parte, posee un codón de iniciación (ATG) seguido por secuencia de poli-histidina (6xHis), un sitio de corte de Enterokinasa y el sitio de clonado múltiple. Para ambos casos se utilizaron plásmidos previamente modificados dónde los sitios de corte fueron sustituidos por un sitio de corte específico de proteasa Trombina (LVPRGS); este sitio deja como subproducto del corte los residuos GS hacia el extremo N-terminal de la proteína de interés, pero que de ninguna manera afectan sus propiedades estructurales y/o funcionales. Los vectores pMalC2x y pRSETa fueron elegidos porque:

- Proveen métodos de expresión del gen de interés inducido por el agregado de Isopropil- β -D-1-tiogalactopiranosido (IPTG) y de purificación que en principio ya estaban desarrolladas en términos generales en el laboratorio;
- Contienen un gen de *r^{amp}* de resistencia a la ampicilina, un antibiótico de fácil acceso que es usado para seleccionar los clones bacterianos transformados con el plásmido de expresión.

- En sus sitios múltiples de clonado, contiene los sitios de restricción *BamHI* y *HindIII*, enzimas que estaban adquiridas y las que era usual trabajar en el laboratorio.

El sub-fragmento de AdE1A que se decidió clonar en estos plásmidos se extiende desde el aminoácido 36 al 146 de la secuencia natural y comprende la región CR1 (conteniendo el *motivo E2F*) y la región CR2 (conteniendo el *motivo LxCxE*) separados por una región de 71 residuos, que llamamos de aquí en adelante: *linker*. En el Recuadro 1.1 se encuentra la secuencia FASTA y la descripción de los límites del sub-fragmento utilizado en este trabajo.

```
>sp|P03255|E1A_ADE05 Early E1A protein OS=Human adenovirus C serotype 5
OX=28285 PE=1 SV=1
MRHIICHGGVITEEMAASLLDQLIEEVVLADNLPPPSHFEPPTLHELYDLVDVTAPEDPNEE
AVSQIFPDSVMLAVQEGIDLLTFPPAPGSPEPPHLSRQPEQPEQRALGPVSMPLVPEVI
DLTCHEAGFPSSDDEDEEGEEFVLDYVEHPGHGCRSCHYHRRNTGDPDIMCSLCYMRTCG
MFVYSPVSEPEPEPEPEPEPARPTRRPKMAPAILRRPTSPVSRECNSSTDSCDSGPSNTP
PEIHPVVPLCPIKPVAVRVGGRRQAVECIEDLLNEPGQPLDLSCKRPRP
```

Recuadro 1.1. Secuencia FASTA de la proteína AdE1A utilizada en este trabajo. En negritas están señalados los aminoácidos del sub-fragmento seleccionado (36 a 146). En rojo, los aminoácidos centrales del motivo de interacción E2F, en azul, los del motivo de interacción LxCxE. Subrayados los 71 aminoácidos considerados dentro de la región "*linker*".

El esquema de trabajo seguido para sub-clonar este fragmento en los plásmidos de expresión se muestra en la Figura 1.1. Por un lado se procedió a generar los siguientes cebadores, los cuales fueron sintetizados por GenScript (Jinsite Science and Technology Co., Ltd., Nanjing, Japón):

1. E1A₃₆₋₁₄₆FrwBam

```

      G   S   M   S   H   F   E   P   P   T
5' CG GGA TCC ATG AGC CAT TTT GAA CCA CCT ACC 3'
      BamHI
```

2. E1A₃₆₋₁₄₆RevHind

```

      Stop Y   D   L   V   F   E   E   G
5' CC CAA GCT TCA ATA ATC TAA CAC AAA CTC CTC ACC 3'
      HindIII
```

El fragmento de AdE1A (aminoácidos 36 – 189) contenido en el plásmido de origen, ya se encontraba sub-clonado corriente arriba (sobre el aminoácido 36) a continuación de un sitio de restricción *BamHI* que fue naturalmente aprovechado al generar el cebador 1. Con este cebador 1 también se introdujo una Metionina anterior al aminoácido 36 de la secuencia natural de AdE1A.

Con el cebador 2, se buscó levantar la secuencia en forma reversa desde el aminoácido 146 y se introdujo un sitio de corte para *Hind III*. Estos cebadores fueron utilizados en una

Reacción en Cadena de la Polimerasa (PCR, por sus siglas en inglés) utilizando Pfu ADN Polimerasa (Thermo Scientific) sobre el molde pRSET- E1A_{CR1-3} y logrando amplificar el fragmento de interés, el cual fue evaluado analíticamente corriéndolo en gel de agarosa 1%. El producto amplificado de PCR: AdE1A (aminoácidos 36 – 146), fue luego purificado por gel y extraído con kit de purificación de bandas (GE Healthcare) (Figura 1.1).

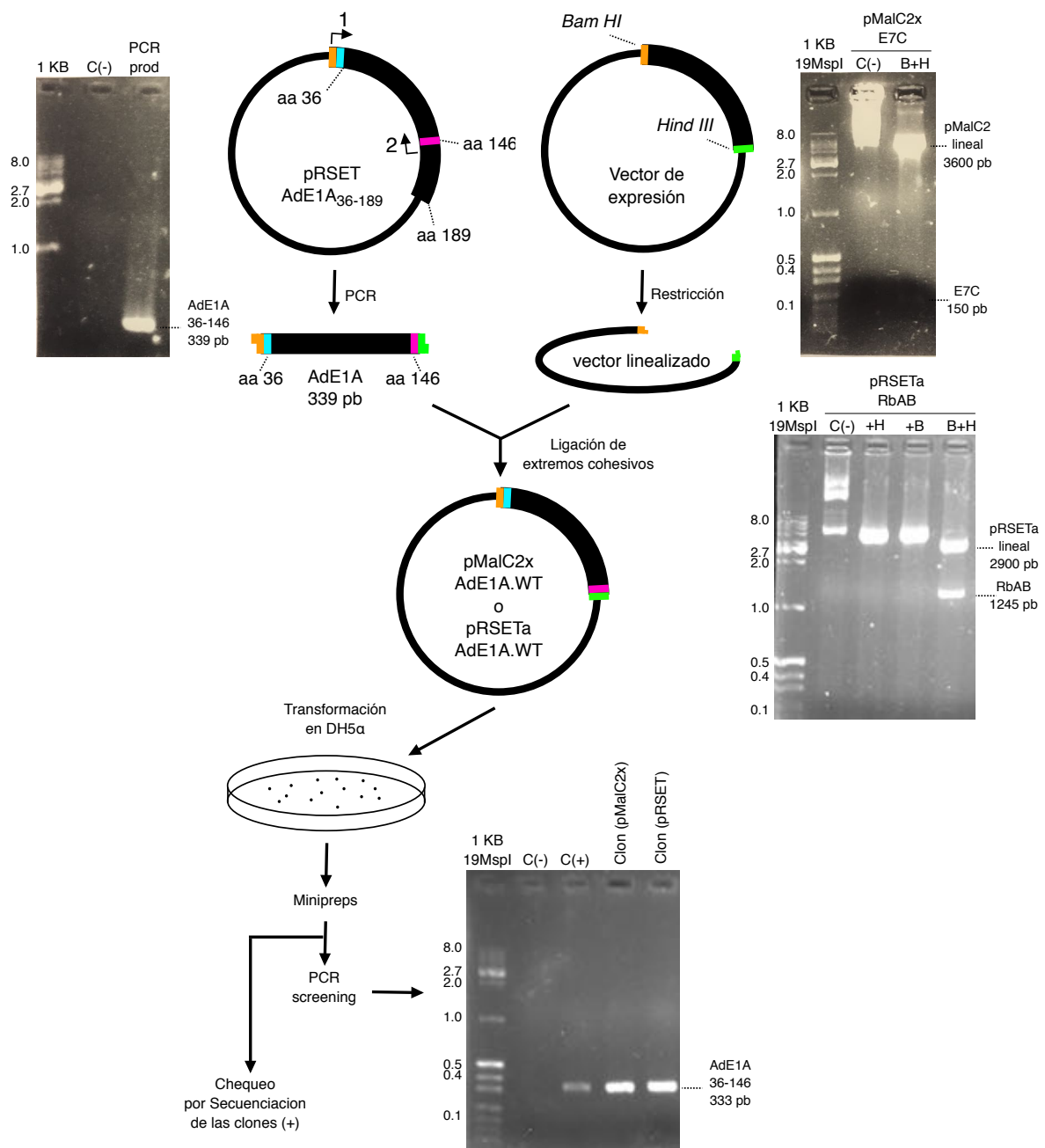


Figura 1.1: Esquema de trabajo para la obtención del pMalC2-AdE1A.WT. Los números 1 y 2 corresponden a los cebadores E1A₃₆₋₁₄₆FrwBam y E1A₃₆₋₁₄₆RevHind respectivamente. En código de colores, naranja corresponde al sitio de restricción *Bam*HI (enzima B) celeste al codón codificante del aminoácido 36, magenta al codón codificante del aminoácido 146 y verde al sitio de restricción *Hind*III (enzima H).

Por otro lado se escogieron un plásmido pMalC2x-E7C y un plásmido pRSETa-RbAB disponibles en el laboratorio y que contenían genes clonados entre los sitios de restricción *BamHI* y *HindIII*. Se digirieron con enzimas de restricción Bam HI (Promega) y Hind III (New England Biolabs), con el fin de recuperar sólo el plásmido lineal con extremos cohesivos. El producto de ambas digestiones fue evaluado en gel de agarosa 1%. La digestión se realizó en un modo preparativo y las bandas de ambos vectores linealizados fueron corridas en gel y cortadas y extraídas con el kit de purificación de bandas.

El fragmento amplificado AdE1A (aminoácidos 36 – 146) y los plásmidos linealizados se ligaron por sus extremos cohesivos utilizando T4 ADN Ligasa (Thermo Scientific) probando relaciones molares 1:1, 1:5 y 1:10 de vector : inserto. La incubación de esta reacción se realizó durante una noche a temperatura ambiente y el día posterior fueron transformados en la cepa DH5α. Sólo las bacterias transformadas con el producto de ligación en relación molar 1:5 crecieron. De múltiples clones transformados que se obtuvieron, se realizaron Minipreps para la extracción del material genético de interés (pMalC2x-AdE1A₃₆₋₁₄₆ o pRSETa-AdE1A₃₆₋₁₄₆), los cuales se evaluaron por PCR de cribado o “*screening*”, utilizando como cebadores universales pMalE (5' –TCAGACTGTCGATGAAGC–3') (Macrogen) para la secuencia clonada en pMalC2x y los cebadores T7 (5' –AATACGACTCACTATAG–3') y T7terminador (5' –GCTAGTTATTGCTCAGCGG–3') (Macrogen) para la secuencia clonada en pRSETa. Los clones identificados como positivos fueron confirmados por secuenciación. A los fines prácticos, este fragmento (AdE1A₃₆₋₁₄₆) de la variante natural de AdE1A se denominó en adelante AdE1A.WT. En todos los casos en los que se usaron enzimas o kits comerciales se siguieron instrucciones y protocolos del proveedor.

Como con el cebador 1 también se introdujo una Metionina anterior al aminoácido 36 de la secuencia natural de AdE1A, el **producto final** con el que se trabajó tras el corte con Trombina fue:

GSM SHFEPPTLHELVDLDVTAPEDPNEEAVSQIFPDSVMLAVQEGIDLLTFPPAPGSPEPPHLSRQP
EQPEQRALGPVSMPNLVPEVIDLTCHEAGFPPSDDDEEGEEFVLDY,

de 114 aminoácidos totales, donde GSM se consideraron los residuos -2, -1 y 0 respectivamente, anteriores a la secuencia 36 a 146. Los parámetros de este fragmento, fueron calculados por la herramienta web <https://web.expasy.org/protparam/>.

Desarrollo de las mutantes de la variante natural AdE1A

Para cumplir con los objetivos planteados en esta tesis, se desarrollaron simultáneamente dos variantes mutantes de AdE1A.WT donde se eliminan uno de los dos sitios de interacción con RbAB:

- Una primer variante llamada AdE1A- Δ E, que contiene el núcleo del *motivo E2F* mutado por un parche de tetra-alaninas (43-LHELY-47 Δ 43-AAAA-46) y el *motivo LxCxE* intacto.
- Una segunda variante llamada AdE1A- Δ L, que contiene el núcleo del *motivo LxCxE* mutado por un parche de tetra-alaninas (122-LTCHE-126 Δ 122-AAAA-125) y el *motivo E2F* intacto.

En la Figura 1.2 se presenta el esquema de trabajo que se siguió para generar estas variantes. En primer lugar se generaron los cebadores para llevar a cabo mutagénesis sitio-dirigida:

a. E2F_{site}mut-frw

5' GCA GCA GCA GAT TTA GAC GTG ACG GCC 3'

b. E2F_{site}mut-rev

5' TGC TGC GGT AGG TGG TTC AAA ATG GCT 3'

que permitirían la generación de la mutante del motivo E2F, y

c. LxCxE_{site}mut-frw

5' GCT GCT GCT GCT GGC TTT CCA CCC AGT 3'

d. LxCxE_{site}mut-rev

5' TGC GGC ATC GAT CAC CTC CGG TAC AAG 3'

que permitirían la generación de la mutante del motivo LxCxE. Para los cebadores –frw, los tres primeros codones no eran complementarios al gen de interés e introdujeron mutación por Alaninas, en tanto los otros seis codones si eran complementarios. Para los cebadores –rev, los últimos dos codones introdujeron mutación por Alaninas y los siete codones restantes eran complementarios a la secuencia.

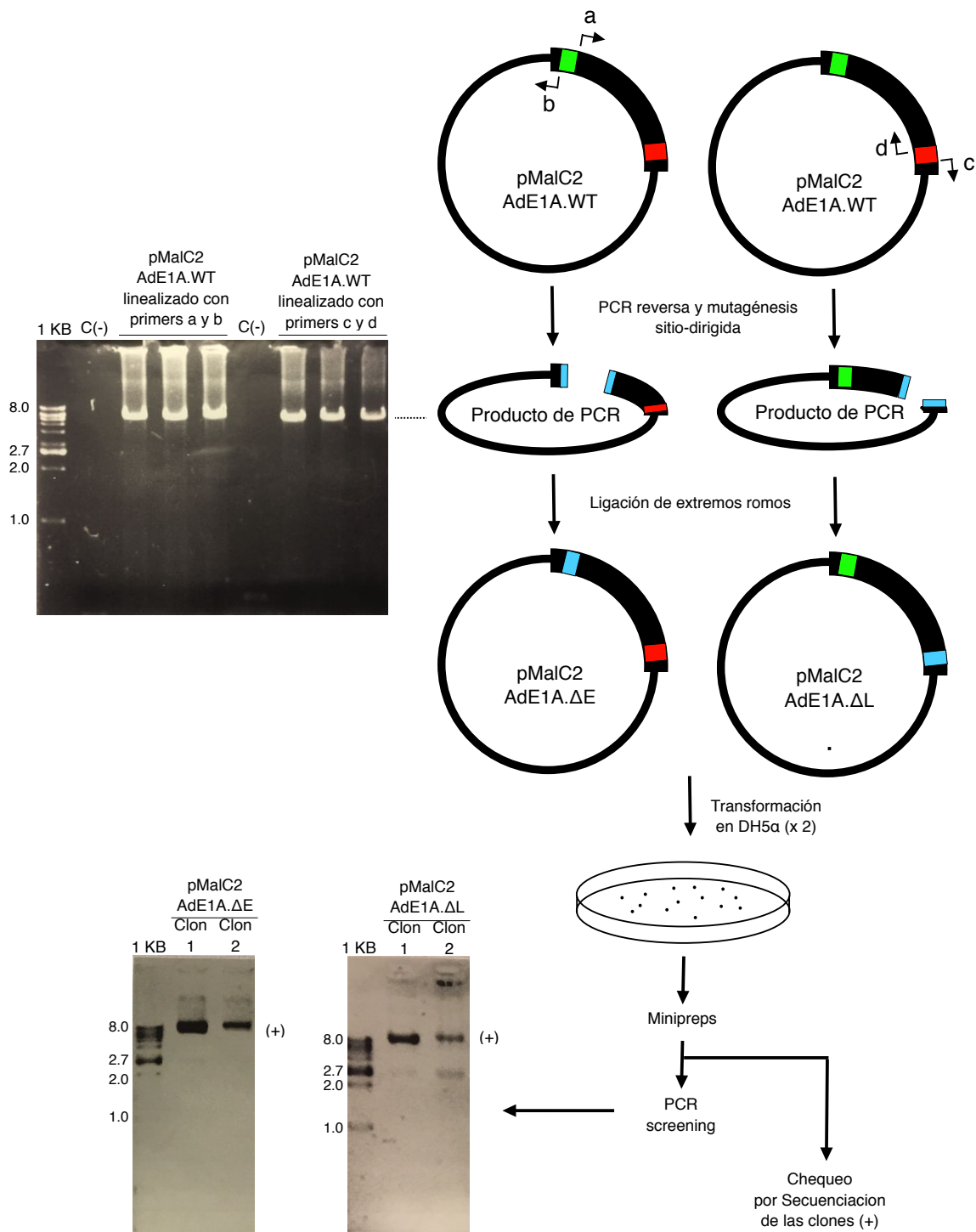


Figura 1.2: Esquema de la mutagénesis sitio-dirigida y ligación de extremos romos para la obtención de las mutantes AdE1A-ΔE y AdE1A-ΔL. Las letras a, b, c y d corresponden a los cebadores utilizados para guiar la mutagénesis y PCR. La banda negra corresponde a la secuencia de la proteína AdE1A entre los aminoácidos 36 y 146, en tanto que el recuadro verde representa el motivo E2F y el recuadro rojo, el núcleo del motivo LxCxE. Los recuadros celestes corresponden a las Alaninas introducidas para generar la mutación de los motivos.

Los cebadores se utilizaron independientemente en PCR reversa donde pMalC2x-AdE1A.WT sirvió como molde. Como en esta estrategia de PCR se iba a amplificar el inserto y

todo el plásmido pMalC2x, con una longitud total de ≈ 7000 pb, se requirió una enzima de alta fidelidad y procesividad, por lo que se usó una Q5 ADN Polimerasa de Alta Fidelidad (New England Biolabs). Además, como se producen fragmentos con extremos romos no cohesivos, fue necesario tratar previamente a los cebadores con T4 polinucleótido quinasa (PNK) (New England Biolabs) para fosforilarlos y facilitar así la posterior ligación. Los productos de las PCRs fueron evaluados analíticamente corriéndolos en gel de agarosa 1% y posteriormente purificados por gel y extraídos con kit de purificación de bandas. Los productos amplificados y linealizados de la PCR reversa fueron tratados con enzima *DpnI* (Thermo Scientific). Esta enzima corta selectivamente el enlace inmediatamente posterior a las Adeninas metiladas de la secuencia 5' -G_mATC-3', la que se encuentra naturalmente sólo en el ADN sintetizado por las bacterias, por lo que permite eliminar el vector molde residual post PCR.

Posteriormente se procedió a la re-ligación de los plásmidos amplificados y mutados a partir de sus extremos romos fosforilados, utilizando T4 ADN Ligasa (Thermo Scientific) durante una noche a temperatura ambiente. El día siguiente los productos de la ligación fueron utilizados para transformar bacterias de la cepa DH5 α . Los clones transformados se hicieron crecer para poder llevar a cabo Minipreps para la extracción del material genético de interés (pMalC2-AdE1A- Δ E y pMalC2-AdE1A- Δ L), los cuales se evaluaron por PCR de cribado o "screening" utilizando el cebador pMalE (5' -TCAGACTGTTCGATGAAGC-3') (Macrogen).

Originalmente se buscaba mutar cada motivo de 5 residuos por un parche de cinco alaninas consecutivas. Sin embargo, como las ligaciones de extremos romos fosforilados presentaron altas tasas de error, tras analizar por secuenciación más de 20 clones para cada mutante, se consideraron como positivos aquellos que presentaron 4 alaninas reemplazando al motivo original. Esta reducción en 1 aminoácido respecto de la longitud de la variante natural no produjo resultados inesperados ni impidió el análisis comparativo entre las secuencias. Así, los productos finales con los que se trabajaron tras el corte con Trombina fueron:

AdE1A- Δ E:

GSM SHFEPPTAAAADLDVTAPEDPNEEAVSQIFPDSVMLAVQEGIDLLTFPPAPGSPEPPHLSRQ
PEQPEQRALGPVSM PNLVPEVIDLTCHEAGFPPS DDEDEEGEEFVLDY,

AdE1A- Δ L:

GSM SHFEPPTLHELVDLDVTAPEDPNEEAVSQIFPDSVMLAVQEGIDLLTFPPAPGSPEPPHLSR
QPEQPEQRALGPVSM PNLVPEVIDAAAAAGFPPS DDEDEEGEEFVLDY,

ambos de 113 aminoácidos totales, donde GSM se consideraron los residuos -2. -1 y 0 respectivamente, anteriores a las secuencias entre 36 a 146. Los parámetros de cada fragmento fueron calculados por la herramienta web <https://web.expasy.org/protparam/>.

Por una cuestión en la simplicidad de las comparaciones, las secuencias en las mutantes también se enumeraron de 36 a 146, considerando una posición vacante entre el parche de tetra-alaninas. Los productos generados por biología molecular que fueron utilizados en esta tesis, fueron los siguientes:

AdE1A.WT

36_SHFEPPT**LHELY**DLDVTAPEDPN~~EE~~AVSQIFPDSVMLAVQEGIDLLTFPPAPGSPEPPHLSRQPEQPEQRALG
PVSMPNLVPEVID**LTCHE**AGFPPSDDEDEEGEEFVLDY_146

AdE1A-ΔE

36_SHFEPPT**AA**_AADLDVTAPEDPN~~EE~~AVSQIFPDSVMLAVQEGIDLLTFPPAPGSPEPPHLSRQPEQPEQRALG
PVSMPNLVPEVID**LTCHE**AGFPPSDDEDEEGEEFVLDY_146

AdE1A-ΔL

36_SHFEPPT**LHELY**DLDVTAPEDPN~~EE~~AVSQIFPDSVMLAVQEGIDLLTFPPAPGSPEPPHLSRQPEQPEQRALG
PVSMPNLVPEVID**AA**_AAAGFPPSDDEDEEGEEFVLDY_146

Recuadro 1.2. Secuencia de la proteína AdE1A.WT y mutantes utilizadas en este trabajo. Nótese que por simplicidad, las tetra-alaninas tienen una posición vacante en el medio. Esta consideración de ningún modo interfirió en los análisis o resultados presentados en la presente tesis.

I.1.B. Expresión y Purificación de Proteínas

Luego de tener clonados los proteínas de interés, se desarrollaron los protocolos de expresión en medio rico y medio mínimo a fin de obtener cantidades suficientes de proteínas para experimentos posteriores.

Ensayos iniciales de expresión de AdE1A.WT y mutantes en Medio Rico

En principio, se buscaron condiciones óptimas de expresión de AdE1A.WT clonado en plásmidos pMalC2x y pRSETa en medio rico, en ensayos de pocos ml, para luego escalar y aplicar también a la expresión de las variantes mutantes.

En primer lugar se ensayaron las expresiones en dos cepas bacterianas:

E. coli BL21 (DE3) genotipo F- *ompT hsdSB (r_B⁻ m_B⁻) gal dcm* (DE3)

E. coli C41 (DE3) genotipo F- *ompT hsdSB (r_B⁻ m_B⁻) gal dcm* (DE3),

Ambas cepas se encuentran diseñadas para expresar el gen de interés en un sistema regulado por el promotor T7. Contienen el lisógeno DE3 que porta el gen para la T7 ARN polimerasa bajo control del promotor *lacUV5*, de manera que la expresión de esta polimerasa -y en consecuencia del gen de interés- requiere IPTG y se puede inhibir por altas concentraciones de Glucosa. Además carecen de las proteasas *lon* y *OmpT*, reduciendo así las posibilidades de degradación de las proteínas expresadas en estas cepas. C41, es derivado de BL21, pero

con determinadas mutaciones fenotípicas que la confieren tolerancia a la expresión de proteínas tóxicas.

Estas bacterias fueron transformadas con los dos vectores y dejadas crecer ON a 37 °C. Se tomaron colonias y se estriaron dejándolas crecer por alrededor de 5hs más a 37 °C. De una misma estría se inocularon 3 ml de un cultivo control en medio rico LB suplementado con Ampicilina (1X) y glucosa 1% m/v (para mantener inhibida la expresión de la proteína de interés) y otros 3 ml de medio suplementados con Ampicilina 0.1 mg/ml (1X) y glucosa 0.2% m/v. Estos cultivos fueron crecidos a 37 °C agitando a 200 rpm, hasta que alcanzaron una DO_{600nm} de 0.7, momento en el que el cultivo de baja glucosa fue inducido con IPTG 0.8 mM. Post inducción siguieron creciendo a 37 °C y se cosecharon 17 horas después. Se tomaron alícuotas correspondientes a la misma cantidad de bacterias para todas las condiciones y se sembraron en geles de poliacrilamida desnaturalizantes (SDS-PAGE) al 15% para corroborar la expresión. En la Figura 1.3, que muestra los geles de estos experimentos, se puede observar que se obtuvieron niveles elevados de sobre-expresión de la proteína de fusión MBP-AdE1A.WT (de aproximadamente 57 kDa), equivalentes tanto en BL21 como en C41 transformadas con el plásmido pMalC2x-AdE1A.WT. Por otro lado, la banda de sobre-expresión en C41 transformado con pRSETa-AdE1A.WT fue mínima y en el caso de BL21 transformado con pRSETa-AdE1A.WT no se pudo siquiera identificar banda de expresión. Con estos resultados se decidió trabajar en adelante solamente con plásmido pMalC2x-AdE1A.WT.

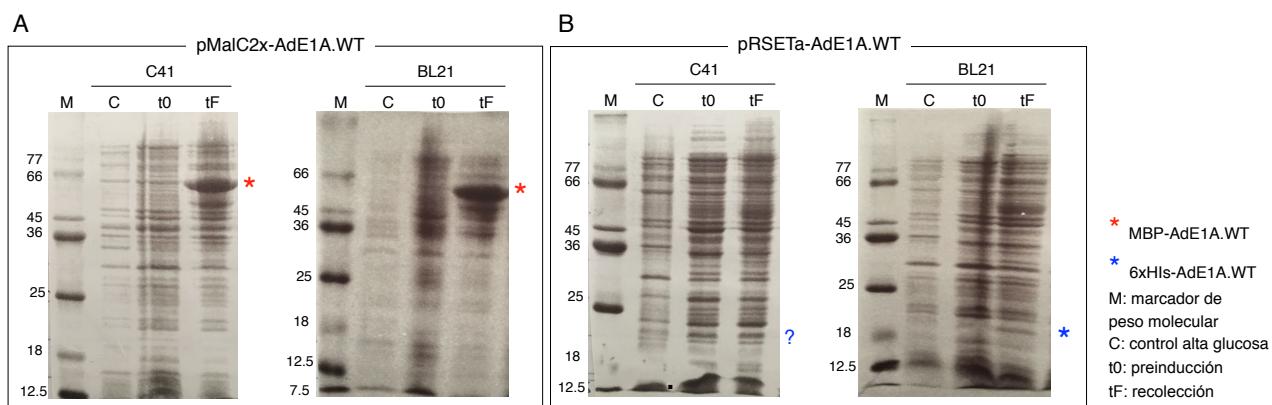


Figura 1.3: Análisis de la expresión de AdE1A utilizando los diferentes vectores construidos y cepas de expresión. A) Geles SDS-PAGE 15% evaluando la sobreexpresión de AdE1A.WT clonado en el plásmido pMalC2x en las cepas C41 y BL21. En ambos casos se observa buen nivel de expresión proteica. B) Geles SDS-PAGE 15% evaluando la sobreexpresión de AdE1A.WT clonado en el plásmido pRSETa en las cepas C41 y BL21. No se observó banda de sobre-expresión en C41, en tanto que en BL21, la misma fue mínima.

En un experimento posterior se testeó también si una temperatura de expresión de 28°C mejoraría los resultados provistos por la expresión a 37°C. Para ello se transformaron nuevamente las cepas C41 y BL21 con pMalC2x-AdE1A.WT y se siguió el mismo

procedimiento anterior, sólo que tras inducir los cultivos con IPTG 0.8 mM a llegar a DO_{600nm} de 0.7, los mismos fueron incubados a 28°C y 37°C para ser cosechados 16 hs después. Los geles SDS-PAGE 15% de este experimento están en la Figura 1.4. Como se puede ver para C41, ambas temperaturas permitieron la sobre-expresión de MBP-AdE1A.WT con la misma efectividad. Para BL21 los resultados fueron similares.

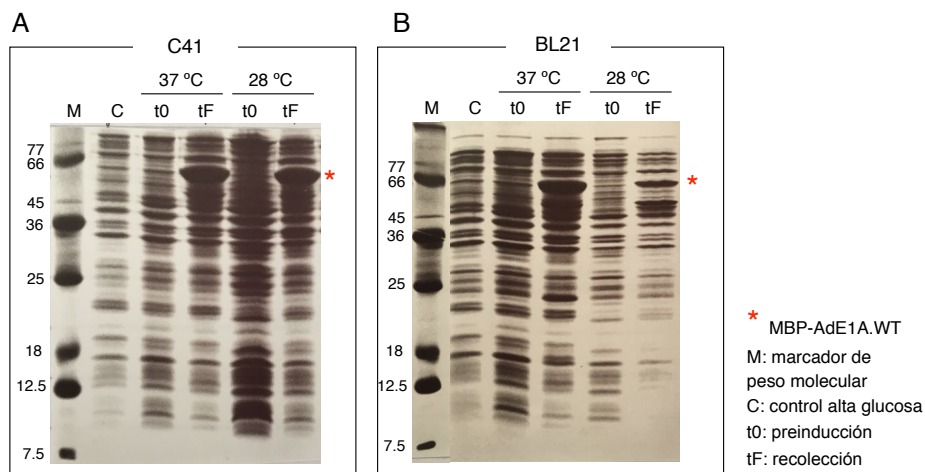


Figura 1.4: Análisis de la expresión de AdE1A a diferentes temperaturas de inducción. A) Gel SDS-PAGE 15% evaluando la sobre-expresión de AdE1A.WT en las cepas C41 con temperaturas de inducción a 37°C y 28 °C. En ambos casos se observa buen nivel de expresión proteica. B) La sobre-expresión de AdE1A.WT en BL21, mostró resultados similares. Nótese que para BL21 inducido a 28 °C la sobre-expresión parece menor, pero *no lo es* dado que a las calles del gel fueron cargadas con un tercio de la cantidad sembrada en las calles Control y para 37°C.

Elegidos el plásmido pMalC2x-AdE1A.WT, la cepa C41 y la temperatura de inducción a 37°C, el siguiente paso fue escalar la expresión de la proteína en cultivos de 500 ml y evaluar la conveniencia de inducir en fase *log* de crecimiento bacteriano o en fase saturada y la fracción en la que particiona la proteína de interés expresada en cada condición.

Tras una transformación en fresco, se inoculó 50 ml de medio rico 2X Triptona-Levadura (2TY) suplementado Ampicilina 1X y glucosa 0.2% y fueron crecidos a 37°C con agitación a 200 rpm, hasta que alcanzaron saturación (pre-cultivo). Con este pre-cultivo se inocularon 2 Erlenmeyer con 500 ml de medio 2TY y se hicieron crecer a 37°C a 200 rpm, uno hasta DO_{600nm} de 0.7 (fase logarítmica o log) y otro hasta DO_{600nm} de 4 (fase saturada) momentos en los que se indujeron con IPTG 0.8 mM. El que fue indujo en fase log se recolectó 16 hs después, una vez alcanzado DO_{600nm} de 4 y el inducido en fase saturada se recolectó 8 hs post inducción. Se tomaron alícuotas para todas las condiciones y se sembraron en geles SDS-PAGE al 15%. Luego los cultivos fueron centrifugados y el pellet bacteriano resuspendido y lisado (ver protocolo mas adelante) y sus fracciones se sembraron también en gel. Como se puede observar en la Figura 1.5, los niveles de expresión de AdE1A.WT y dónde particiona, resultaron

independientes del momento de la inducción. Aun así, se evidenció un mínimo nivel de expresión basal de la proteína en el cultivo que ya estaba saturado al momento de la inducción. Como se buscaban condiciones controladas de expresión, se eligió inducir los cultivos en fase log (Figura 1.5 A). La fracción donde se encontró AdE1A.WT expresada fue en la fracción soluble (Figura 1.5 B).

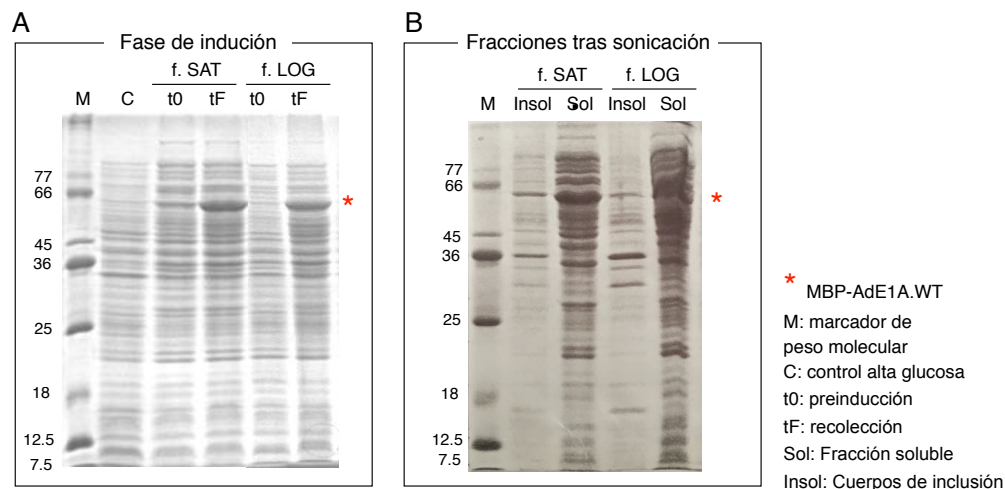


Figura 1.5: Análisis de la expresión de AdE1A escalado a 500 ml inducida en diferentes fases de crecimiento bacteriano. A) Gel SDS-PAGE 15% evaluando la sobre-expresión de AdE1A.WT tras la de inducción en fase saturada y en fase logarítmica de crecimiento de las bacterias C41. En ambos casos se observa buen nivel de expresión proteica. B) Gel SDS-PAGE 15% evaluando dónde particiona AdE1A.WT, en ambas calles sembradas cantidades equivalentes de las fracciones solubles e insolubles separadas tras sonicación y centrifugación. Independientemente del momento de inducción AdE1A.WT se encuentra en la fracción soluble.

Protocolo de expresión y purificación a escala de AdE1A.WT y mutantes en Medio Rico

Una vez comprobadas las mejores condiciones de expresión, se escalaron a 3 litros de cultivo. Las mismas condiciones luego se aplicaron para expresar AdE1A- Δ E y AdE1A- Δ L, resultando igualmente efectivas.

El protocolo que se muestra a continuación es para AdE1A.WT a modo ejemplificativo:

a. Transformación y Expresión: se transformaron bacterias *Escherichia coli* C41-(DE3) con $\approx 100 \mu\text{g}$ de los respectivos plásmidos (productos de Minipreps), y fueron dispuestas en placas LB-agar con Ampicilina 0.1 mg/ml (1X) y crecidas por una noche a 37 °C. Se seleccionó un único clon para inocular un Erlenmeyer con 150 ml de medio 2TY suplementado con ampicilina 1X y glucosa 1% (m/v), el cual se hizo crecer durante una noche a 37 °C con agitación constante de 220 revoluciones por minuto (rpm) para asegurar correcta oxigenación. El pre-cultivo saturado de 150 ml sirvió para inocular 3l de medio 2TY distribuidos en 4 Erlenmeyers de 750 ml cada uno y suplementados con ampicilina 1X y glucosa 0.2% m/v. Estos cultivos se hicieron crecer a 37 °C y 220 rpm, hasta alcanzar una Densidad Óptica

medida a 600 nm (DO_{600}) de 0.7, momento en el que las bacterias se encuentran en su fase logarítmica de crecimiento y se los indujo con IPTG a 0.8 mM. Los cultivos inducidos fueron recolectados post-inducción, tras un ON y una vez llegados a $DO_{600} = 4$. Los niveles de expresión fueron evaluados por geles SDS-PAGE. Las bacterias fueron centrifugadas a 12000 rpm por 20 minutos y el pellet bacteriano conservado en $-80\text{ }^{\circ}\text{C}$ hasta su lisis.

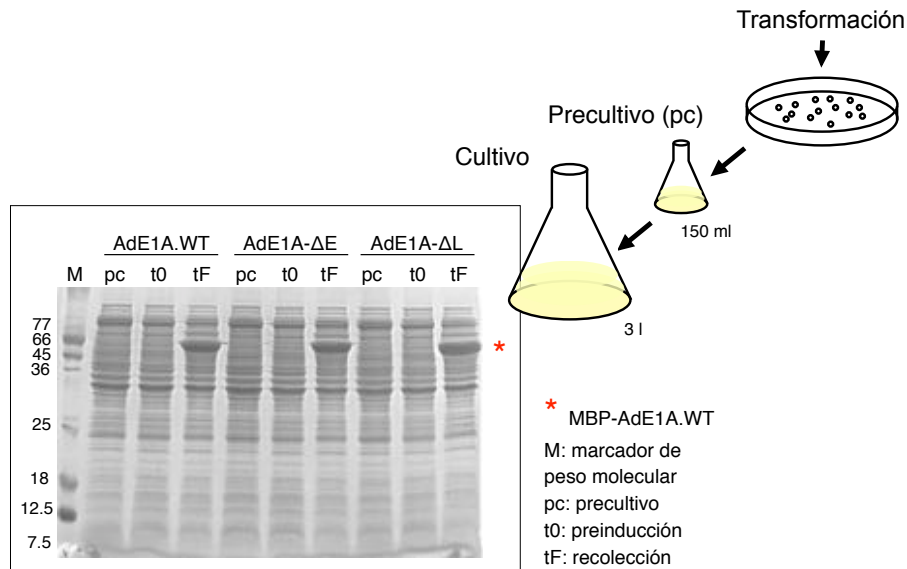


Figura 1.6: Expresión a escala de AdE1A.WT, AdE1A-ΔE y AdE1A-ΔL en 3 l de cultivo. Cantidades de bacterias equivalentes del pre-cultivo, del cultivo antes de inducir y del cultivo al momento de recolección fueron sembradas comparativamente en un gel SDS-PAGE 15% evaluando la sobre-expresión de AdE1A. Para las tres construcciones se observaron altos niveles de expresión.

b. Resuspensión, Lisis y precipitación con NH_4SO_4 : las bacterias fueron resuspendidas en 50 ml/litro de cultivo, de tampón de lisis conteniendo 50 mM Tris:Cl pH 7.5, 1mM ácido etilendiaminotetraacético (EDTA), 2mM 2-mercaptoethanol (β -Me) y 2mM fluoruro de fenilmetilsulfonilo (PMSF) y sometidas a un ciclo de congelado y descongelado (40 minutos a $-80\text{ }^{\circ}\text{C}$ y agitación a $4\text{ }^{\circ}\text{C}$ hasta descongelar). Luego fueron tratadas por cada 50 ml de material resuspendido con 900 μl de DNAsa1 (1 mg/ml) (Sigma), 1,5 ml MgCl_2 0.5 M y 0.45 ml CaCl_2 1M, incubándolas a $37\text{ }^{\circ}\text{C}$ por 40 minutos. Las células fueron lisadas por sonicación por duplicado y posteriormente centrifugadas a 18000 rpm por 20 minutos para aislar la fracción soluble que es donde particiona MBP-AdE1A. La recuperación de la proteína en la fracción soluble se evaluó por gel (Figura 1.7, izquierda) . Con posterioridad, y tras una primera puesta a punto (Figura 1.7, derecha), la fracción soluble se precipitó con NH_4SO_4 50% m/v para obtener un precipitado o pellet, el cual fue resuspendido con tampón A: 50 mM Tris-HCl, pH 7.5, 200 mM NaCl, 2 mM β -Me y 2 mM PMSF. Este producto resuspendido se dejó dializando un ON a $4\text{ }^{\circ}\text{C}$ para proceder con las técnicas de purificación ulteriores.

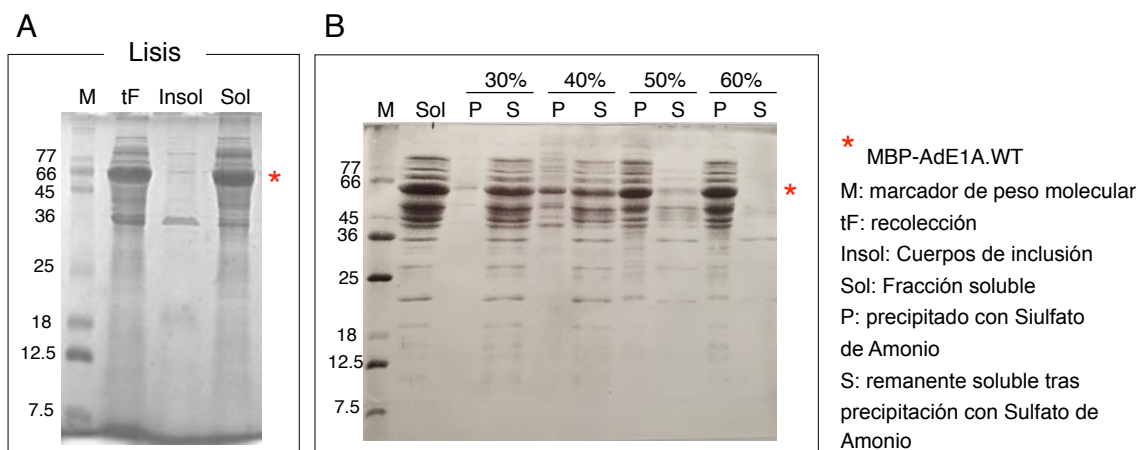


Figura 1.7: Análisis de la lisis y Precipitación con NH_4SO_4 . A) Gel SDS-PAGE 15% mostrando que AdE1A.WT particiona a la fracción soluble. Las tres calles sembradas con cantidades equivalentes de total bacteriano y fracciones soluble e insoluble separadas tras sonicación y centrifugación. B) Puesta a punto de la precipitación de la fracción soluble con NH_4SO_4 , dónde se trató por 4 horas a 4 °C y agitación constante con 30, 40, 50 y 60% m/v de sal. La condición elegida fue del 50% m/v, con la que se precipita máxima cantidad de MBP-AdE1A.WT y deja en el remanente no precipitado proteínas que no son de interés.

c. Purificación de MBP-AdE1A. Cromatografía de afinidad; se cargó la muestra resuspendida y dializada, en una columna conteniendo resina de Amilosa (NewEngland Biolabs), funcionando a flujo de gravedad y equilibrada en tampón A. Esta resina tiene afinidad por la MBP, lo que dejó retenida la construcción MBP-AdE1A. El percolado fue chequeado por Bradford hasta que no se vió coloración y MBP-AdE1A fue eluida con tampón A conteniendo 20 mM Maltosa. El eluato se dializó una noche a 4 °C contra tampón C: 50 mM Tris-HCl, pH 7.5, 200 mM NaCl, 2 mM β -Me y posteriormente se concentró con concentradores Amicon (Millipore) hasta concentración cercana a 2 mg/ml chequeada por Absorbancia a 280nm y usando como absortividad molar (ϵ) 48800 $\text{M}^{-1} \text{cm}^{-1}$ correspondiente a la especie MBP-AdE1A

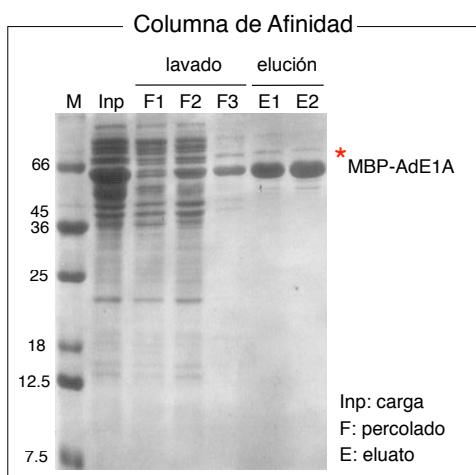


Figura 1.8: Cromatografía de Afinidad de Amilosa. Gel SDS-PAGE 15% mostrando las fracciones obtenidas tras pasar la muestra por una resina de amilosa que retiene por afinidad a MBP. Las fracciones finales del lavado y el eluido con tampón A más Maltosa 20mM libera la proteína de interés.

d. Corte de MBP-AdE1A con Trombina para obtener AdE1A; tras ensayos iniciales para poner a punto esta reacción, se eligió la condición en la que se requirió mínima cantidad de enzima para obtener el corte completo de la MBP-AdE1A, la cual fue la condición 1:2.5 (Figura 1.9 A). MBP-AdE1A se digirió entonces con Trombina (Sigma) a razón de 1U por cada 2.5 mg de proteína y concentración final de 2mg/ml de proteína, en tampón C durante 4 horas a 37 °C. La reacción se detuvo por el agregado de PMSF a 2 mM concentración final. El producto proteolizado se evaluó por gel (Figura 1.9, derecha) y se dializó contra tampón QA que contiene 30 mM Tris-HCl, pH 7.5, 50 mM NaCl, 2 mM β -Me durante una noche a 4 °C.

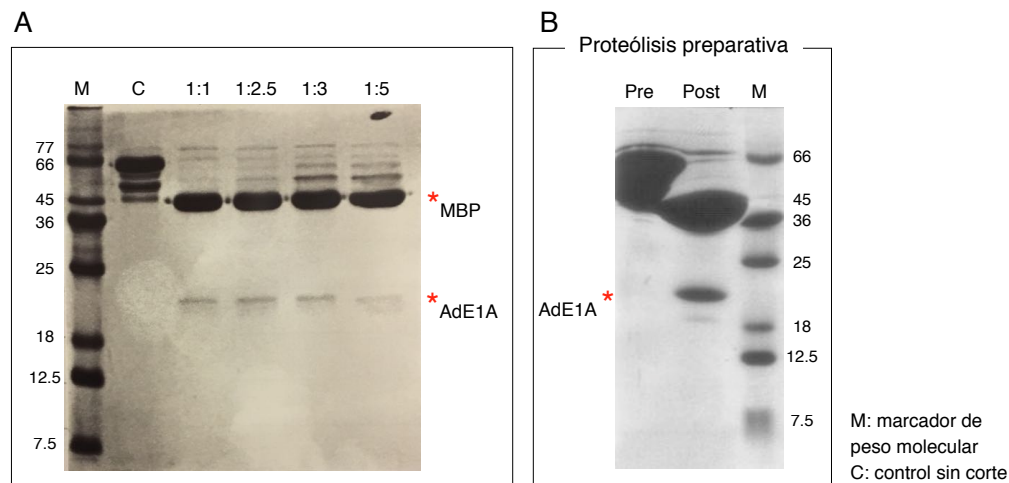


Figura 1.9: Proteólisis con Trombina. A) La puesta a punto de la proteólisis de MBP-AdE1A con Trombina. Se probaron relaciones 1U enzima por 1, 2.5, 3 y 5 mg de MBP-AdE1A con una incubación de 4.30 horas a 37°C en tampón de corte C. La condición elegida fue de 1U/2.5 mg de proteína, que muestra proteólisis completa y utiliza menor cantidad de enzima. B) Gel SDS-PAGE 15% mostrando el resultado de la proteólisis preparativa de MBP-AdE1A sembradas en ambas calles igual cantidad de muestra pre y post agregado de Trombina.

e. Cromatografía de Intercambio Iónico (aniónico); en este paso lo que se buscó fue separar selectivamente a AdE1A de la MBP tras el corte con trombina y otras impurezas que pudieron haber co-purificado en pasos anteriores y se aplicó una estrategia de separación basada en la carga neta de la superficie de la proteína de interés. Como AdE1A posee un pI 3.78, a pH 7.0 (pH de trabajo) la misma se encuentra en forma global cargada negativamente (especie aniónica), por lo que se eligió el uso de una matriz o fase estacionaria que se encuentre cargada positivamente a pH de trabajo Q-HyperD (Pall), la cual tiene como grupo funcional Amonio cuaternario, (NH_4^+ , especie catiónica). La separación en esta etapa puede ser dificultosa para IDPs porque podrían establecer interacciones promiscuas (con otras múltiples proteínas) arrastrando impurezas, lo que se observó en los primeros ensayos de esta etapa de purificación. Para solucionar este inconveniente, se recurrió a hacer el intercambio en presencia de Urea, un agente caotrópico que en altas concentraciones desnaturaliza la

estructura terciaria de las proteínas, lo que de modo consecuente podría evitar las interacciones inespecíficas con la IDP [115].

Previa a la carga de la columna, la muestra se le agregó Urea sólida hasta concentración final 8M, se agitó 40 minutos en frío y se filtró con filtros 0.45 μm (Millipore). La muestra se cargó en la columna con resina Q-Hyper-D acoplada a un equipo FPLC AKTA (GE Healthcare) y equilibrada con tampón QA y 5% QB (QA + 1.2 M NaCl) realizando un lavado extensivo, para luego aplicar con un gradiente salino de 5 a 100% de tampón QB. Con la aplicación de un gradiente salino creciente, aumenta la fuerza iónica del medio y se logra un apantallamiento de cargas de los $-\text{NH}_4^+$ de la matriz y de AdE1A (cargada negativamente), permitiendo su separación.

El percolado fue recolectado (hasta los 110 ml eluyentes, Figura 1.10 A), para luego aplicar un lavado extensivo (Vol: 110 a 135 ml, Figura 1.10 A) en el que no se recolectaron fracciones. Se aplicó el gradiente y se obtuvieron tres picos (gradiente: 25 al 80% de tampón QB, Figura 1.10 A) recolectados por separado. Estas muestras fueron sembradas en gel SDS-PAGE 15% (Figura 1.10 B) en el que se observó que en el percolado (F) eluyó fundamentalmente MBP la cual no se une a la columna, en el primer pico (P1) eluyó MBP (probable pegado inespecífico) y algo de MBP-AdE1A no digerido. En el segundo pico (P2) eluyó principalmente AdE1A y en baja proporción MBP, algo de MBP-AdE1A no digerido y un fragmento de menor peso molecular (que luego se identificó como producto de degradación de AdE1A -ver secciones siguientes-). En tanto el tercer pico (P3) estaba enriquecido en AdE1A. P2 y P3 fueron juntados para seguir con un paso de purificación final que permitió obtener AdE1A con alto grado de pureza. Esta muestra (P2+P3) fue dializada contra tampón 20 mM fosfato de sodio, 200 mM NaCl, 2 mM β -Me, 2 mM PMSF y fue concentrada en fracciones de hasta 2 ml (4ml totales en este punto por cada lote de producción de la proteína) en concentradores Amicon.

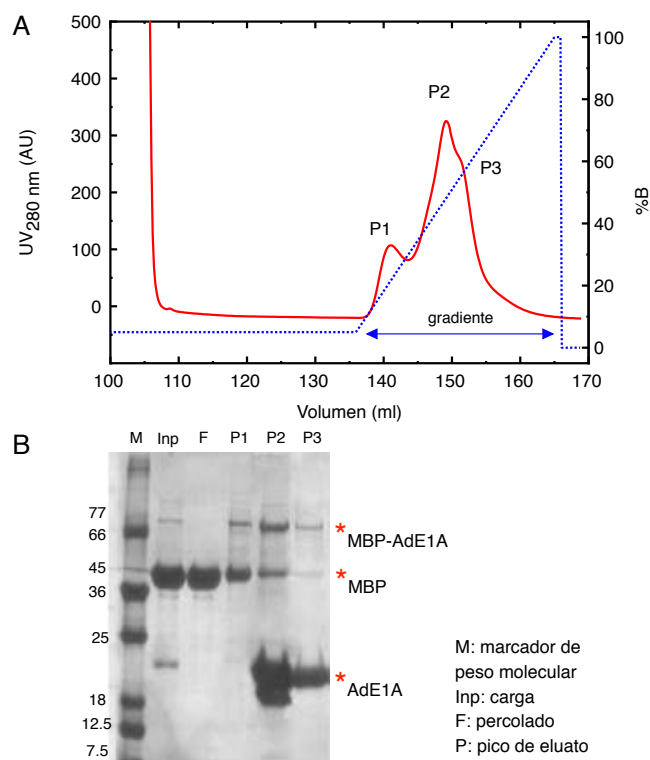


Figura 1.10: Cromatografía de Intercambio aniónico. A) El cromatograma de la corrida mostrando el lavado extensivo al 5% QB y el gradiente de 5 a 100% QB (línea azul punteada) y la señal a 280nm de la proteína (línea roja). B) Gel SDS-PAGE 15% mostrando las fracciones obtenidas en la cromatografía.

f. Cromatografía de filtración o exclusión molecular; la muestra de AdE1A obtenida luego de la purificación por Intercambio aniónico aún contenía impurezas: MBP-AdE1A (que corre entre las bandas de 66 y 77 kDa), MBP (a la par de la banda de 45kDa) y la banda de degradación (Figura 1.10 B). Para lograr un buen nivel de purificación de la proteína de interés, se procedió con una columna de exclusión para separarlas diferencialmente por sus R_H diferenciales.

La muestra fue inyectada a una columna preparativa Superdex 75 (Amersham-Pharmacia) acoplada a un equipo FPLC AKTA (GE Healthcare) y fue corrida con tampón 20 mM fosfato de sodio, 200 mM NaCl, 2 mM β -Me, pH 7.0, para finalmente obtener la proteína purificada de impurezas que no pudieron ser eliminadas en pasos anteriores. En este paso se inyectaron de a 2 ml por vez (4 ml en total) y se corrió a 2ml/min, obteniendo finalmente AdE1A.WT altamente purificado evaluado en gel SDS-PAGE 15%.

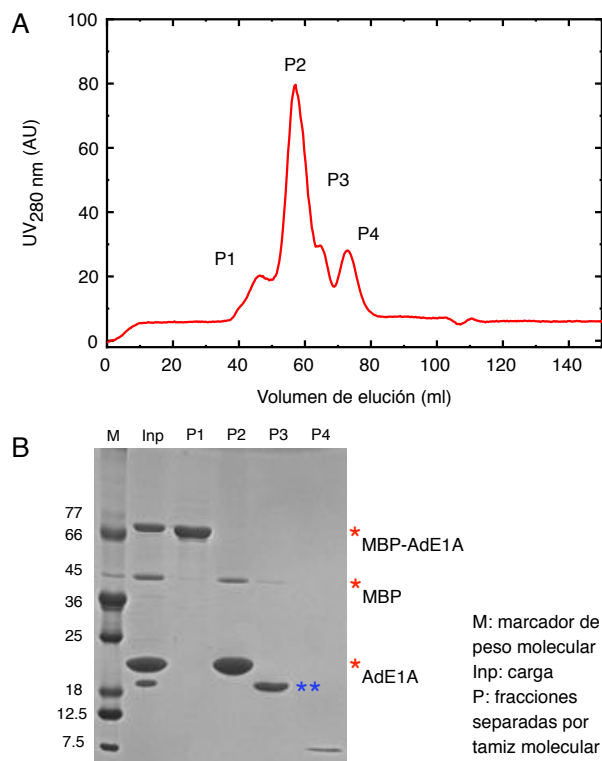


Figura 1.11: Cromatografía de Exclusión Molecular Superdex 75. A) El perfil de elución de AdE1A.WT y otras fracciones por una columna Superdex 75. El Pico P2 corresponde a AdE1A.WT. B) Gel SDS-PAGE 15% mostrando las fracciones separadas en la cromatografía. El doble asterisco azul, corresponde a una sub-especie producto de la degradación de AdE1A.

El perfil de elución de la S75 (Figura 1.11 A) muestra que fueron 4 picos distintos los que eluyeron y fueron recolectados por separado, siendo P2 en principal. Estos picos fueron sembrados individualmente en un gel SDS-PAGE 15%, en el que se observó que P1 correspondía a la especie MBP-AdE1A. P2, a la especie de interés AdE1A con una mínima impureza (evaluada en este y otros geles, como menor al 5%) correspondiente a MBP. P3 correspondió a una especie que luego se identificó como producto de la degradación de AdE1A y P4 a una impureza menor que corre a 7.5 kDa.

La fracción P2, conteniendo la proteína de interés se concentró en el mismo tampón 20 mM fosfato de sodio, 200 mM NaCl, pH7.0, hasta concentración deseada y se le agregó 2mM DTT y se conservó a en -80 °C hasta su uso. El rendimiento final máximo alcanzado siguiendo este protocolo fue de 9 mg de proteína por litro de cultivo con un grado de pureza superior al 95% evaluado por gel SDS-PAGE.

Ensayos iniciales de expresión de AdE1A.WT y mutantes en Medio Mínimo

Para la producción de proteína con simple (^{15}N) y doble ($^{15}\text{N}/^{13}\text{C}$) marca necesaria para los experimentos de Resonancia Magnética Nuclear, la expresión se llevó adelante en Medio Mínimo M9 (MM9) suplementado con $^{15}\text{NH}_4\text{Cl}$ y ^{13}C -glucosa como únicas fuentes de ^{15}N y ^{13}C respectivamente.

Los ensayos iniciales de expresión en medio mínimo se llevaron a cabo en cultivos de 3ml y se probó nuevamente si el plásmido pRSETa-AdE1A.WT podría ser útil. Se evaluaron cepas BL21, C41 y BL21 *rosetta* y temperaturas de inducción a 28 y 37°C, pero ninguna condición fue efectiva para la expresión de AdE1A.WT.

Por otro lado, se probó la expresión del plásmido pMalC2x-AdE1A.WT directamente en escala de 1l. Se evaluaron las cepas C41 y BL21 y temperaturas de inducción a 28 y 37°C. La Figura 1.12 A y B muestra los geles SDS-PAGE, con el resultado de estos experimentos. Todas las condiciones testeadas mostraron buenos niveles de expresión en medio mínimo M9. La condición de expresión en BL21 a 37°C fue la elegida para escalar la producción de la proteína en 3l.

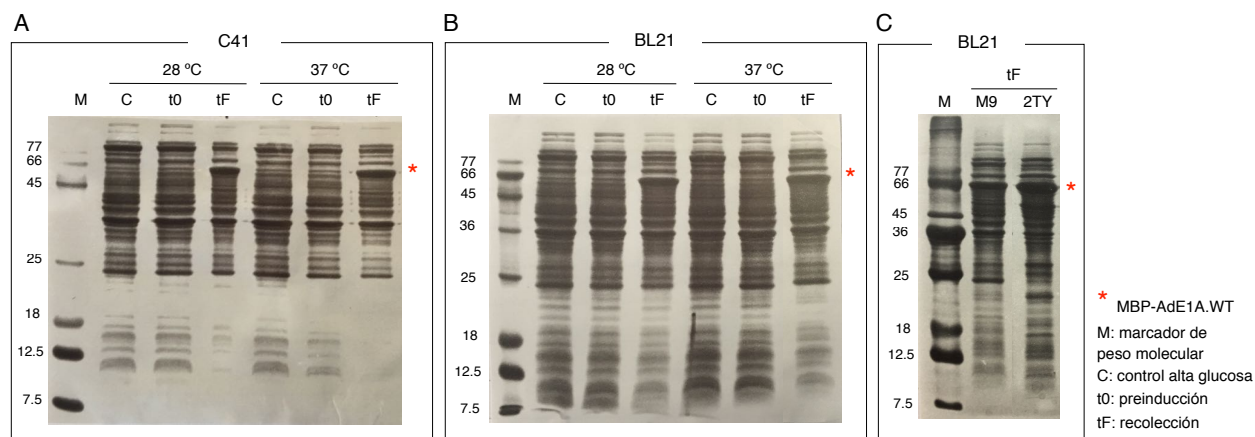


Figura 1.12: Análisis de la expresión de AdE1A en medio mínimo M9, escalado a 1l inducida en diferentes fases de crecimiento bacteriano y temperaturas de inducción. A) Gel SDS-PAGE 15% evaluando la sobre-expresión de AdE1A.WT en las cepas C41 inducidas en fase log a temperaturas de inducción a 37°C y 28 °C. B) La sobre-expresión de AdE1A.WT en BL21 bajo mismas condiciones de prueba. En todos los casos se observó resultados similares. C) Gel comparativo de los niveles de expresión en M9 y 2TY mostrando claramente que en medio mínimo se alcanzan menores niveles de expresión que en medio rico.

En la Figura 1.12 C, se muestra comparativamente los niveles de expresión de BL21 transformadas con pMalC2x-AdE1A.WT, inducidas en fase log ($\text{DO}_{600} = 0.7$) a 37°C, en 1l medio rico 2TY y medio mínimo M9 y recolectadas ambos en saturación. Como el crecimiento bacteriano es menor en medio mínimo respecto del medio rico y las $\text{DO}_{600\text{nm}}$ que se alcanzan difícilmente superan a 2, existe un compromiso mas delicado entre niveles de sobre-expresión y cantidad de bacterias expresando activamente la proteína de interés. Por ello, sobre un

cultivo de 1l M9 con BL21 transformadas con pMalC2x-AdE1A.WT e inducido en fase log y a 37°C, se siguió la curva de crecimiento y niveles de sobre-expresión post inducción. Se midieron las DO_{600nm} y se tomaron alícuotas de iguales cantidades de bacterias por cada punto de control, sembrando un gel SDS-PAGE 15%. Como se puede ver en la Figura 1.13, ni los niveles de expresión de MBP-AdE1A.WT ni la DO son óptimos tras un ON (16 horas post inducción) y la mejor relación entre ambos se alcanzó a las 4 horas post inducción, por lo que se lo eligió como tiempo de recolección del cultivo.

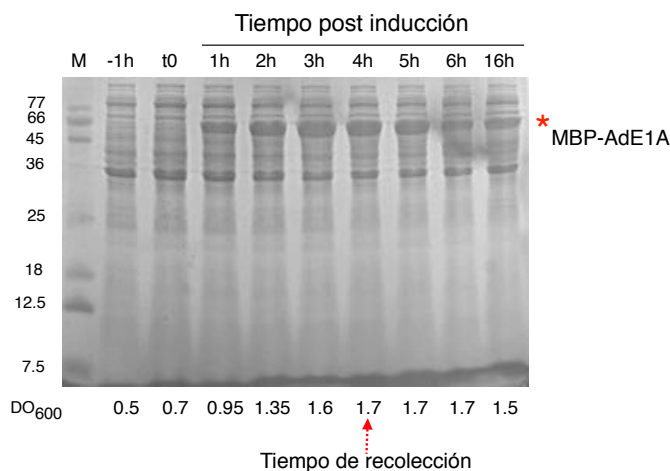


Figura 1.13: Análisis de los niveles de expresión de AdE1A en medio mínimo M9 y la DO bacteriana tras la inducción. Gel SDS-PAGE 15% evaluando niveles de sobre-expresión y DO bacteriana desde 1 hora pre- a 16 horas post-inducción. El tiempo de recolección está señalado y corresponde al de mayor DO y nivel de expresión de la proteína de interés.

Finalmente, con condiciones de expresión y tiempo de recolección puestos a punto, se escaló a cultivos de 2l finales.

Protocolo de purificación a escala de AdE1A.WT y mutantes en Medio Mínimo

El protocolo de expresión y purificación a escala en medio mínimo M9, fue el mismo que el descrito para medio rico, salvo el primer paso, que fue de la siguiente forma:

Preparación, inoculación y expresión: se preparó en esterilidad 2l de medio mínimo M9 sin la marca (¹⁵N o ¹⁵N/¹³C) el cual no se autoclavó si no que se filtró usando filtros estériles. Se dividieron en 1l de medio para 2 Erlenmeyer autoclavados, trabajando cerca de llama para esterilidad. Se tomaron 2ml de cada Erlenmeyer y se colocaron en tubos de ensayos que sirvieron de primer control. Los tubos se etiquetaron “-N/-amp/+bacterias” y se inocularon con bacterias, y se incubaron a 37°C ON, para evaluar que nada crezca debido a la ausencia de Nitrógeno. Para las muestras que requirieron doble marca, se preparó un tubo más “-N/-C/-amp/+bacterias”. para testear que no haya crecimiento en ausencia de ¹³C-Glucosa.

En el caso de marcación simple se agregó cuidadosamente 1g de $^{15}\text{N-NH}_4\text{Cl}$ (Cambridge Isotope Laboratories, Inc.) a cada Erlenmeyer conteniendo 1l de medio ya suplementado con glucosa común. De cada Erlenmeyer se tomaron 2ml para control “+N/-amp/+bacterias”, en el cuál las bacterias deben crecer. En el caso de doble marcación se agregó ^{13}C -Glucosa (Cambridge Isotope Laboratories, Inc.) en este punto y se tomaron 2ml para el control “+N/+C/-amp/+bacterias”, en el cuál las bacterias también deben crecer. Para ambos casos, luego se agregó antibiótico AMP 1X (final) a cada Erlenmeyer y se tomaron otros 2ml para control “+N/+amp/-bacterias”, en el cual nada debe crecer porque no hay bacterias inoculadas y constituye el control de esterilidad. Se tomaron 25ml de cada Erlenmeyer y se combinaron en un único Erlenmeyer de 250 ml para inocular las bacterias (pre-cultivo). Este último Erlenmeyer se dejó crecer a 37°C ON junto a los tubos controles. Los Erlenmeyers grandes se reservaron en frío hasta el otro día.

A la mañana siguiente, se chequearon los controles y se inoculó con 25 ml de pre-cultivo a cada Erlenmeyer grande atemperado a 37°C desde 1 hora antes. Se chequearon que las DO_{600} estén entre 0.04 y 0.06 en este paso y se dejaron crecer a 37°C. Una vez que alcanzaron una DO_{600} 0.7, se indujeron con IPTG 0.8 mM. Se volvieron a dejar crecer por 4.30 hs y se recolectaron las bacterias. A partir de aquí, el protocolo siguió como fue descrito anteriormente para medio rico.

Para todos los dos casos de expresión en medio mínimo M9 (simple y doble marcación) cada paso resultó equivalente al obtenido en medio rico, salvo por los rendimientos que son presentados en la Tabla 1.1. Los cromatogramas mostraron similares perfiles de purificación al igual que los geles. A modo representativo de la expresión en medio mínimo, se presenta la Figura 1.13, en la cual muestra los geles del primer (expresión) y último paso (post-filtración S75) del protocolo de expresión y purificación de AdE1A.WT con doble marcación.

Por otro lado, como los experimentos de RMN requirieron gran cantidad de muestra marcada (generalmente por ensayo: 600ul de AdE1A a concentración > 200 μM , lo que implica alrededor de 1.5 mg de proteínas por ensayo) se hicieron 1 ronda de doble marcación y dos rondas de simple marcación para el total de experimentos (ver capítulo III).

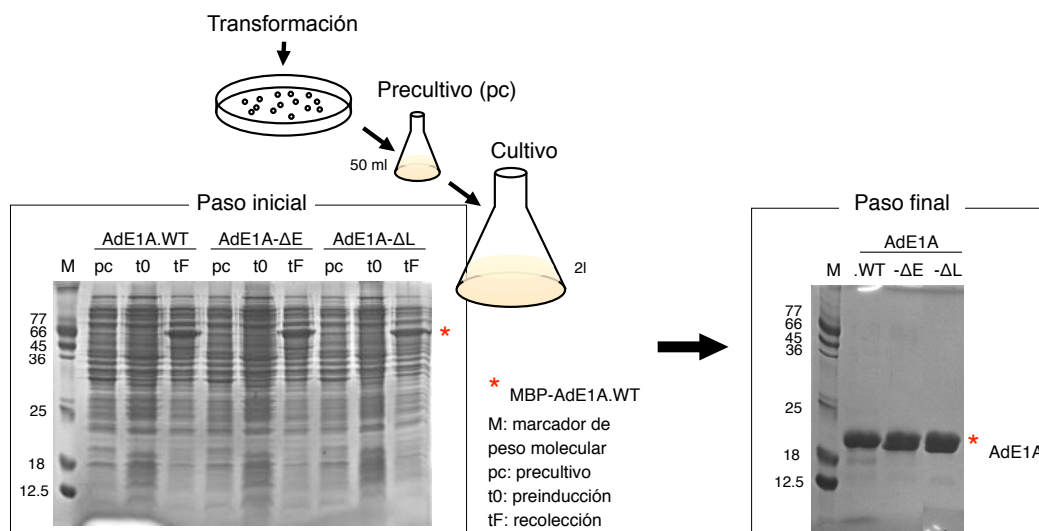


Figura 1.13: Geles representativos de la expresión y purificación de AdE1A doblemente marcada. A) Cantidades de bacterias equivalentes del precultivo, del cultivo antes de inducir y del cultivo al momento de recolección fueron sembradas comparativamente en un gel SDS-PAGE 15% evaluando la sobreexpresión de AdE1A en medio mínimo. Para las tres construcciones se observaron niveles de expresión equivalentes. B) Gel SDS-PAGE 15% que muestra las proteínas AdE1A.WT y mutantes purificadas como producto final obtenido post-filtración con S75. Para los tres casos la purificación fue superior al 95%.

Rendimiento de los procesos de purificación de AdE1A

La Tabla 1.1 muestra las diferentes condiciones de cultivo para las distintas variantes de AdE1A obtenidas y purificadas con el mismo protocolo. Los rendimientos que aquí se detallan, representan a los lotes o rondas de producción con mayor nivel de expresión final de proteína obtenida de todos los producidos durante el período de trabajo de tesis. El mismo, permite evaluar el efecto de las diferentes condiciones de expresión, calculado en miligramos finales por litro de cultivo obtenido para cada etapa implementada. También se presentan cuantos litros iniciales se trabajaron para la expresión de proteínas en los distintos medios, así como las máximos DO_{600} alcanzadas por los cultivos al momento de la recolección.

En líneas generales, las expresiones en medio rico 2TY dieron un rendimiento mayor por litro de cultivo (de 7 a 9 mg/l) respecto de los obtenidos en medio mínimo (de 2 a 4 mg/l). Curiosamente se observó que las expresiones en medio rico fueron mas efectivas para la variante natural AdE1A.WT, la cual presentó mayor rendimiento que las variantes mutantes, algo que se modificó en el caso de las expresiones en medio mínimo, donde las variantes mutantes tuvieron mayores rendimientos finales que AdE1A.WT.

Comparando las DO_{600} y los rendimientos de las obtenciones en medio mínimo para doble y simple marcación, se observa que la doble marcación fue mas efectiva. Esto

probablemente pueda deberse a que el *stock* sólido de Glucosa usada para las dobles marcaciones (^{13}C -Glucosa) haya estado en mejores condiciones que el *stock* disuelto al 40% de Glucosa sin marcar (Sigma) usado en los ensayos de simple marcación.

Teniendo en cuenta que AdE1A representa alrededor de la quinta parte de la proteína de fusión MBP-AdE1A obtenida de manera recombinante, el rendimiento final de la purificación representa un porcentaje en torno al 80% de la producción inicial de proteína en medio rico, y de un 50% de la producción inicial de proteína en medio mínimo. Para todos los casos las purezas finales fueron superiores al 95%

Tabla 1.1 Rendimientos máximos de las purificaciones a escala en las distintas condiciones ensayadas para AdE1A.WT, AdE1A- ΔE y AdE1A- ΔL

Variantes de AdE1A	Medio rico 2TY			Medio mínimo M9					
	.WT	- ΔE	- ΔL	Simple marca (^{15}N)			Doble marca ($^{15}\text{N}/^{13}\text{C}$)		
				.WT	- ΔE	- ΔL	.WT	- ΔE	- ΔL
Litros de cultivo por lote de producción	3	3	3	2	2	2	2	2	2
Máxima DO_{600} alcanzada	4.0	4.1	4.0	1.6	1.5	1.5	1.3	1.3	1.3
Rendimiento post-columna afinidad Amilosa (mg/l)	43.4	34.3	52	28.5	28.4	30.4	30	36.5	42.5
Rendimiento post-digestión y columna intercambio QHD (mg/l)	12.7	17	18	7.6	6.3	5.2	6.3	8.5	7.5
Rendimiento concentración pre-filtración S75 (mg/l)	11.7	13.5	13	5.4	5.2	4.5	5.4	7.7	4.8
Rendimiento post-filtración S75 y concentración	8.6	7.3	7.5	2.4	2.6	2.6	2.9	3.7	3.5
Rendimiento final (mg/l)									

Para todos los casos, las recetas de ambos medios de expresión y los tampones utilizados se encuentran descriptos en Materiales y Métodos.

OBJETIVO 1.2.

Caracterización de las propiedades estructurales-conformacionales de AdE1A

Una vez generados los reactivos y dado la falta de caracterizaciones biofísicas/bioquímicas en AdE1A, se evaluaron y caracterizaron sus propiedades en estado libre y solución en condiciones controladas, con el fin de entender cómo funciona esta proteína.

1.2.A. Análisis del comportamiento electroforético de AdE1A

La electroforesis en gel permite separar proteínas de mezclas complejas y purificarlas, investigar sub-unidades conformacionales y verificar homogeneidad, en función a la movilidad electroforética de dichos componentes en la muestra. En principio, una vez aplicado el campo eléctrico, la movilidad electroforética de las proteínas en el gel depende de su relación masa(tamaño)/carga (m/z). Con el uso de dodecil-sulfato de sodio (SDS), las proteínas se desnaturalizan y adquieren una carga negativa homogénea producto de la asociación con las moléculas de SDS. Además, en condiciones reductoras dadas por el uso de β -Me, los puentes disulfuro formados por la unión de dos Cisteínas y que estabilizan estructuras terciarias y /o cuaternarias, quedan reducidos. De esta manera la velocidad de migración se independiza de su la relación m/z y depende sólo del estado proteico post-desnaturalización (su tamaño), así como del tamaño de los poros de la matriz donde corre [116].

Esta técnica permite separar componentes proteicos en una muestra, referenciándolos a marcadores de peso molecular conocido. Como marcadores de peso se utilizan proteínas globulares, de manera que otras proteínas globulares en una muestra bajo análisis migran en concordancia a dichos marcadores. Para estas proteínas, se observa que su migración depende solamente de su peso molecular teórico (PM_{TEO}). Las IDPs en cambio, que interactúan con un menor número de moléculas de SDS en comparación con las proteínas globulares y presentan una conformación altamente extendida, pueden mostrar una migración electroforética anómala y corren con un peso molecular aparente (PM_{ap}) superior, alrededor de 1.2 a 1.8 veces su peso molecular teórico [21]. Además, las IDPs debido a su naturaleza unen más débilmente los colorantes como el Coomassie Blue, el cual es utilizado para revelar los patrones de migración proteica en estos geles, lo que dificulta su identificación.

La caracterización de AdE1A.WT comenzó por la observación de una banda en las electroforesis en gel de poliacrilamida al 15% en presencia de SDS (SDS-PAGE) que se llevaron a cabo en los últimos pasos de su purificación, la cual presentaba una migración

anómala respecto del PM esperado para dicha proteína: 12.5 kDa. Con el fin de caracterizar con qué PM_{ap} migraba AdE1A.WT en SDS-PAGE, se corrió la muestra purificada de la proteína y los marcadores de peso, se calcularon los Rf (razón de frente) y se construyeron curvas de calibración correspondientes.

De este análisis se encontró que AdE1A.WT tiene una migración electroforética anómala, corriendo como si tuviese un PM_{ap} de entre 18.5 a 19.5 kDa, unas 1.48 a 1.52 veces mayor que su PM_{TEO} de 12.5 kDa (Figura 1.14). Para el caso de las mutantes AdE1A- ΔE y AdE1A- ΔL , este análisis reveló que las mismas también corren de manera anómala, con un PM_{ap} 1.5-1.6 veces mayor que sus PM_{TEO} .

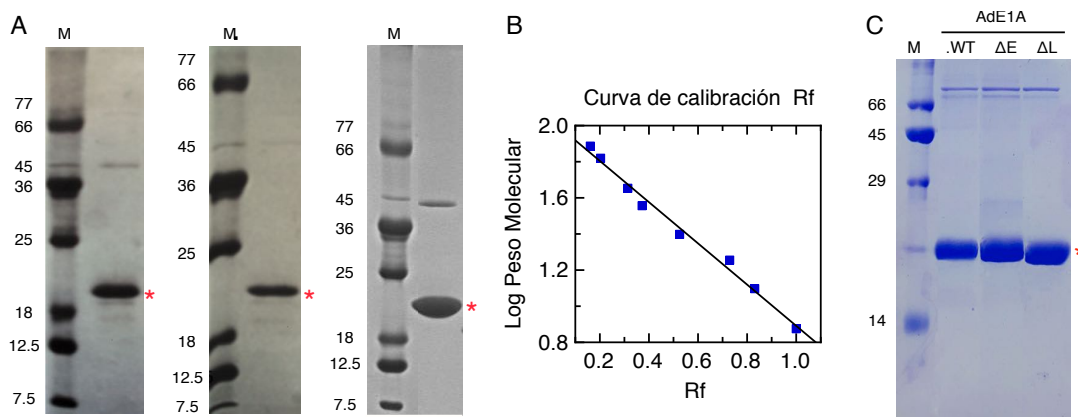


Figura 1.14: Análisis de la migración electroforética de AdE1A. A) Tres gels SDS-PAGE 15% representativos mostrando la migración de AdE1A.WT (asterisco rojo) junto a los marcadores de peso molecular a partir de los que se calcularon los Rf. B) Una curva de calibración de Peso Molecular en función de Rf (para el primero de los gels en A) a partir del que se calculó el PM_{ap} con el que migra AdE1A.WT. C) Gel SDS-PAGE 15% que muestra idéntica movilidad electroforética para la variante WT de AdE1A y las mutantes.

En conclusión, estos experimentos iniciales que describen las características de AdE1A en su variante natural y mutantes, muestran que se tratan de proteínas intrínsecamente desordenadas. Además, el patrón de migración electroforética similar entre las tres variantes, constituye un primer indicio de que las mutaciones, no introducirían mayores cambios conformacionales/estructurales en AdE1A.

I.2.B. Estudio de AdE1A por espectrometría de masa MALDI-TOF y análisis tróptico

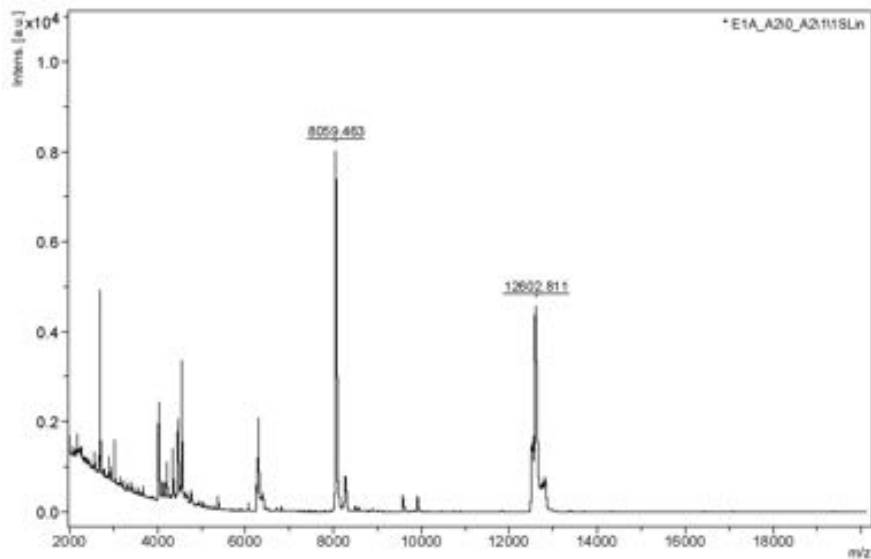
La espectrometría de masa MALDI-TOF es una técnica particularmente útil para la análisis de biomoléculas (péptidos, proteínas, oligosacáridos) que requieran medidas precisas de su pesos moleculares a fines confirmativos, así como para monitorear el resultado de reacciones enzimáticas, modificaciones químicas, etc. Aquí, las muestras conteniendo las biomoléculas (analitos) se pre-mezclan con una matriz altamente absorbente con la que

cristalizan y que convierte la energía del láser -con la que luego es bombardeada- en energía de excitación, conllevando a la desorción e ionización del analito. Estos analitos ionizados y desorbidos “vuelan” con un tiempo diferencial en función de su relación masa/carga (m/z). Este tiempo de vuelo es medido, evidenciando que los iones más pesados son más lentos que los livianos, lo que puede observarse en el espectro m/z que acumula los datos. Generalmente estas especies quedan monocargadas independientemente de su masa y resultan fáciles de identificar, con una precisión dentro del 0.01% de la masa molecular total de la molécula. No obstante, no todas las especies desorben con la misma facilidad y por ejemplo, en una muestra en la que se esperan ver distintas especies muchas de ellas pueden no llegar a identificarse.

Con el fin de confirmar que la proteína cuya banda migra anómalamente en la electroforesis y que fue aislada tras la purificación se tratase de AdE1A.WT, se procedió a la identificación por medio de espectrometría de masa MALDI-TOF. Para ello, fueron analizadas unas microgotas de diversas muestras *stock* de la proteína de interés. Como se trataba del análisis de una proteína, las muestras fueron mezcladas con ácido sinapínico, luego dispuestas en el equipo, impactadas por el láser y detectadas en modo lineal (Ver Materiales y Métodos).

Este análisis evidenció la presencia de dos especies principales desorbidas: una especie pesada de 12602.8 ± 1.26 m/z y otras más liviana de 8059.5 ± 0.81 m/z (Figura 1.15, arriba). Un acercamiento al espectro m/z en torno a la especie más pesada, mostró tres especies de pesos similares y crecientes: 12522.1 ± 1.25 , 12602.8 ± 1.26 y 12608.8 ± 1.26 m/z (Figura 1.15, abajo). En principio el PM_{TEO} esperado para AdE1A.WT es 12518.72 Da y si se compara con los valores de las especies obtenidas, a pesar de estar levemente por debajo, hay una buena concordancia con la especie de 12522.1 ± 1.25 m/z . Las otras especies podrían corresponder a la proteína formando aductos con iones Na^+ (22.99 Da) y/o K^+ (39.09 Da). Hasta este punto, se consideró que el análisis de masa confirmó que la especie aislada post-filtración en S75 se trataba de AdE1A.WT.

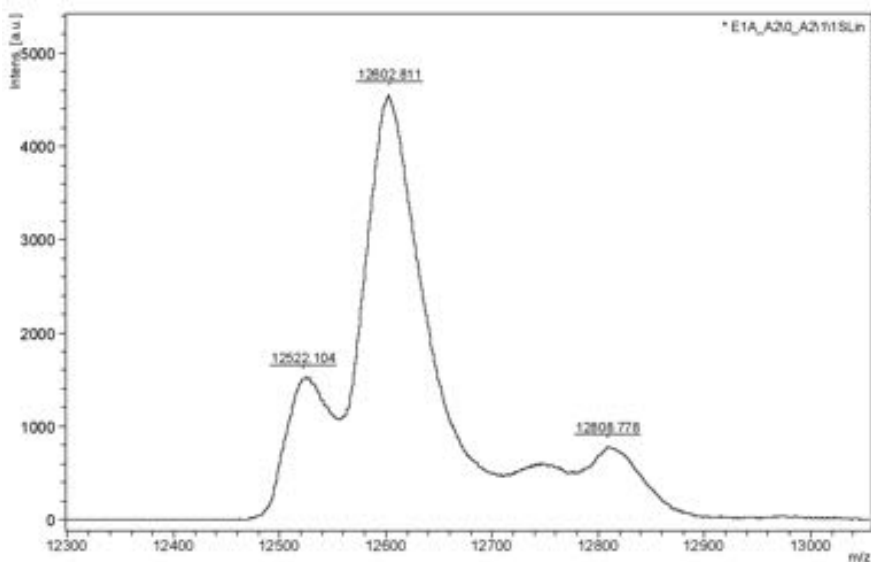
Por otro lado, al fragmento de 8059.5 m/z , se le asignó en principio responsable del pico y banda P3 (Figura 1.11) el cual co-purificaba mínimamente junto AdE1A.WT. Llamativamente, dicha banda aparecía gradualmente en mayor proporción cada vez que un *stock* de AdE1A.WT era descongelado y chequeado en gel SDS-PAGE previo a realizar algún experimento. Esto llevó a pensar que se trataba de un sub-fragmento producto de la degradación de AdE1A.WT, lo cual no era impensado pues las proteínas y regiones desordenadas están más expuestas y son más propensas a hidrólisis por proteasas que pudieran copurificar y las que son efectivas aun en muy bajas concentraciones.



Bruker Daltonics flexAnalysis

printed: 1/20/2016 11:48:16 AM

D:\Data\Lucia\20150904A1_E1A_36146_10uM_puro_sitacido_v20_A1\1



Bruker Daltonics flexAnalysis
BA070055

printed: 1/20/2016 11:49:31 AM

D:\Data\Lucia\20150904A2_E1A_36146_10uM_puro_acidw0_A2\1

Figura 1.15: Espectros MALDI-TOF de muestras "stocks" de AdE1A.WT. Arriba, los picos identificados en el análisis de masa. Abajo, señales correspondientes a AdE1A.WT (36-146) libre y con distintos grados de ionización.

Para saber si el fragmento de 8059.5 ± 0.81 m/z se trataba de un producto de degradación de AdE1A.WT, en primer lugar se analizó su secuencia a fin de ubicar sitios probables de corte por proteasas. La presencia de dos argininas R97 y R105 en la secuencia natural de AdE1A.WT, generó el siguiente mapa de corte:

AdE1A.WT 36-146 + GSM (N-terminal)	
GSM SHFEPPTLHELYDLDTAPEDPNEEAVSQIFPDSVMLAVQEGIDLLTFPPAPGSPEPPHLSRQPEQPEQRALGPVSMPLVPEVIDLTCHEAGFPSSDDEEGEEFVLDY	12518.7
AdE1A.WT 36-97 + GSM (N-terminal)	
GSM SHFEPPTLHELYDLDTAPEDPNEEAVSQIFPDSVMLAVQEGIDLLTFPPAPGSPEPPHLSR	7066.8
AdE1A.WT 36-105 + GSM (N-terminal)	
GSM SHFEPPTLHELYDLDTAPEDPNEEAVSQIFPDSVMLAVQEGIDLLTFPPAPGSPEPPHLSRQPEQPEQR	8059.9
AdE1A.WT 98-105	QPEQPEQR 1011.0
AdE1A.WT 98-146	QPEQPEQRALGPVSMPLVPEVIDLTCHEAGFPSSDDEEGEEFVLDY 5469.9
AdE1A.WT 106-146	ALGPVSMPLVPEVIDLTCHEAGFPSSDDEEGEEFVLDY 4476.8

Recuadro 1.3. Mapa de corte de AdE1A.WT. Secuencias probables de la fragmentación de la proteína AdE1A.WT por acción de la proteasa tripsina, la cual actúa sobre los 2 residuos Arginina que posee la secuencia. Las especies aquí presentadas poseen en el extremo N-terminal los residuos GSM que son extras a la secuencia natural de AdE1A, como fue explicado anteriormente. Por cuestiones de simplicidad, se obviará su presencia en la nomenclatura de los fragmentos, aunque se considerará su aporte al peso molecular del fragmento.

Como se puede ver en el Recuadro 1.3, uno de los fragmentos que se pueden generar por proteólisis es de 8059.9 Da y que podría corresponder al encontrado por análisis de masa en las muestras *stocks* de 8059.5 m/z.

Para llegar a su identificación final se llevó a cabo un estudio de digestión limitada utilizando proteasa tripsina (Sigma) la cual corta el enlace peptídico que sigue a un residuo Lisina (K) o Arginina (R) de una secuencia, siempre que el aminoácido que preceda a K o R no sea una Prolina (P). La enzima utilizada estaba tratada con Tosil-fenilalanil-clorometil cetona (TCPK) que inhibe irreversiblemente a otras serinproteasas como quimiotripsina, la cual puede producir cortes inespecíficos en la proteína. La reacción se llevó a cabo con una relación enzima: sustrato 1:200 m/m en tampón bicarbonato de amonio pH 8.0 por 20 minutos a 37 °C, hasta completar la proteólisis. Finalizada la reacción de digestión, la muestra fue sembrada en un gel SDS-PAGE y se evaluó que se haya completado chequeando la total desaparición de la banda de AdE1A.WT (36-146). Luego, unos 60ug de la muestra sin proteolizar y otros 60ug digeridos fueron corridos por cromatografía de alto rendimiento en fase reversa (RP-HPLC) utilizando una columna C8 y se aislaron los picos para ser analizados por espectrometría de masa.

La Figura 1.16 muestra los resultados de este análisis. Como se puede observar en el gel, la muestra *stock* ya se encontraba parcialmente degradada al momento de realizar la digestión y el análisis de masa del pico purificado por RP-HPLC (Figura 1.16 B, arriba) evidenció un resultado similar al encontrado en los análisis directos sobre muestras *stocks* (sin previa purificación por HPLC), identificándose una especie de 12519.6 ± 1.25 m/z y otra de 8053 ± 0.81 m/z. La primer especie se trataría de AdE1A.WT (36-146) que posee un PM_{TEO} de 12518.72 Da, con un muy buen acuerdo de valores –en el rango del error de la técnica-. La especie de 8053 ± 0.81 m/z correspondería a AdE1A.WT (36-105) la cual tiene un PM_{TEO} de 8059.90. Estas especies corresponderían a las bandas ♦1 y ♦2 en la primer calle del gel SDS-

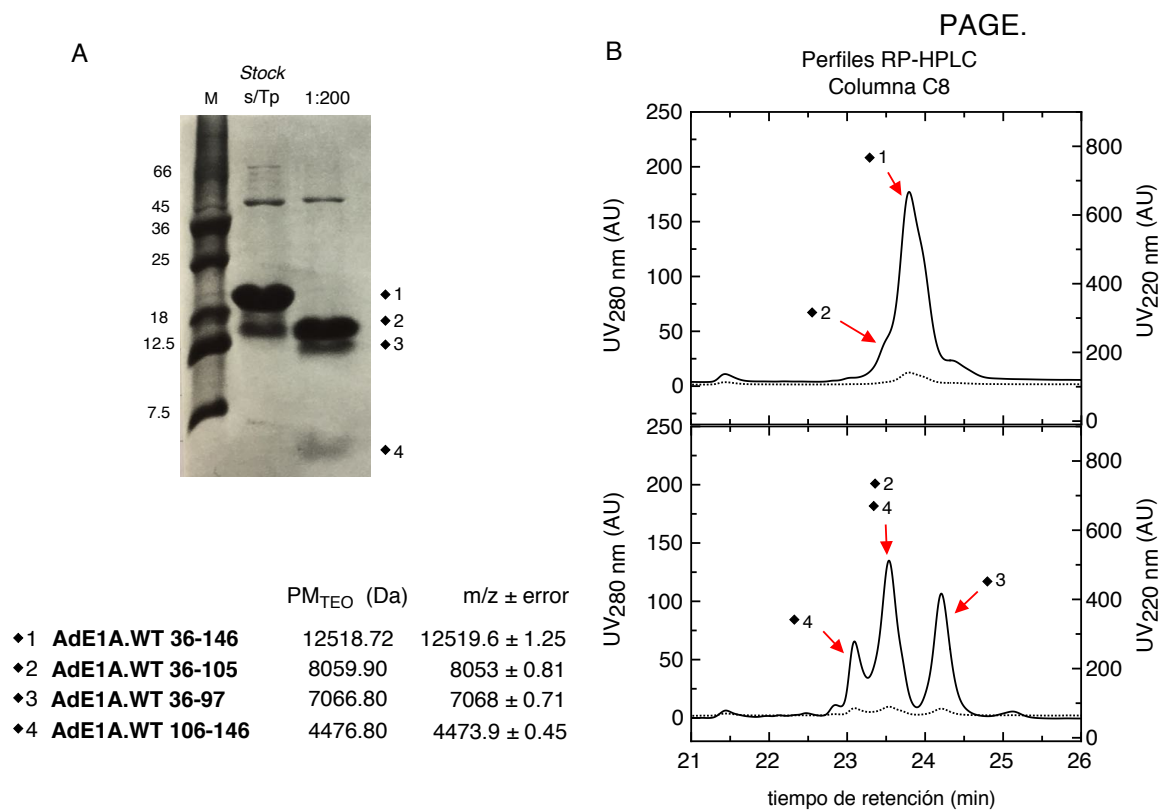


Figura 1.16: Análisis por espectrometría de masa de la digestión triptica de AdE1A.WT. A) gel SDS-PAGE 15% mostrando el resultado de la proteólisis de AdE1A.WT (36-146+GSM). Como se puede observar, la digestión fue completa. B) Los cromatogramas RP-HPLC seguidos a 280nm (línea punteada) y 220 nm (línea continua), de la muestra sin digerir (arriba) y la muestra totalmente digerida (abajo) y señaladas las especies identificadas por pico en el MALDI-TOF (m/z) junto a su PM_{TEO} (Da).

Por otro lado, para la muestra tripsinizada se confirmó completa digestión porque ningún pico eluyó en el RP-HPLC con el mismo tiempo de retención que AdE1A.WT (36-146) y porque no se identificó dicha especie en el análisis de masa de los productos digeridos separados por RP-HPLC (Figura 1.16 B, abajo). La digestión generó una serie de fragmentos, uno de los cuales fue de 8053 ± 0.81 m/z que corresponde a AdE1A.WT (36-105). Justamente esta

especie fue identificada en el pico con mayor área bajo la curva en el cromatograma RP-HPLC y correspondería del patrón de bandas post-digestión, a la banda más gruesa en el gel SDS-PAGE (♦2), la cual tiene idéntica movilidad electroforética a la banda producto de la degradación del *stock* de AdE1A.WT. También se identificó una especie de 7068 ± 0.71 m/z, que tiene buen acuerdo con el PM_{TEO} de 7066.80 Da de AdE1A.WT 36-97 y correspondería a la banda ♦3 del gel SDS-PAGE. Una última especie de 4473.9 ± 0.45 m/z fue identificada, siendo relacionada con AdE1A.WT 106-146 cuyo PM_{TEO} es de 4476.80 Da y correspondería a la banda ♦4 del gel SDS-PAGE.

Es de destacar que a pesar de haber sido esta una digestión extensiva con Tripsina, no se pudo identificar la especie AdE1A.WT 98-105, cuyo peso molecular (1011 Da) sería muy pequeño para ser detectado en el modo lineal de detección con el que se trabajó para este análisis. Por otro lado, la especie AdE1A.WT 98-146 de PM_{TEO} de 5469.9 Da, tampoco pudo ser detectada porque podría no estar desorbiendo bien. Otra cuestión notable es que la especie ♦4 se encuentre en mayor proporción que la especie ♦3 tras una digestión que fue extensiva, además sólo ♦4 se identifica en las muestras degradada naturalmente, lo que evidencia que R105 sería una arginina más lábil que R97.

En resumen, esta serie de resultados además de confirmar que se purificó la proteína de interés, también llevó a concluir que AdE1A.WT copurificaría con mínimas cantidades de proteasas que generan su fragmentación en el tiempo. Además, esto es esta otra característica que la identifican como una IDPs, siendo éstas extremadamente susceptibles a la proteólisis por su conformación extendida que expone sitios de modificación en los que actúan las proteasas. Para preservar AdE1A.WT de la degradación por estas enzimas, se decidió agregar PMSF como agente inhibidor de la proteólisis tanto en los pasos de purificación como en los stocks finales, además de trabajar evitando ciclos de congelado y descongelado múltiples.

I.2.C. Análisis de las propiedades hidrodinámicas y conformacionales de AdE1A

Ya se ha mencionado que las proteínas forman parte de un *continuum* conformacional donde hay una gradualidad de contenido estructural, en un rango que se extiende de aquellas completamente desordenadas y extendidas a proteínas completamente ordenadas y compactas [32], en donde las IDPs podían clasificarse en “RC-like”, “PMG” y “MG” al contener diversos grados de estructura residual (ver Introducción). Por lo tanto una manera de caracterizar una proteína es a partir de su grado de compactación estructural, que se logra estudiando su comportamiento hidrodinámico.

Para ello es útil la medición del radio hidrodinámico o de Stokes (R_H o R_S) de una proteína, que es el radio de una hipotética esfera sólida de peso molecular equivalente, que difunde con la misma velocidad que dicha proteína. Para una proteína globular su R_H medido es muy similar a este R_H teórico o hipotético, debido a su alto grado de compactación. En cambio, para una IDPs con el mismo PM que la proteína globular, el R_H medido es mayor al R_H teórico de la esfera, ya que se encuentra más extendido. Así, conforme aumenta su grado de desorden y extensión, su R_H medido será cada vez mayor (Figura 1.17).

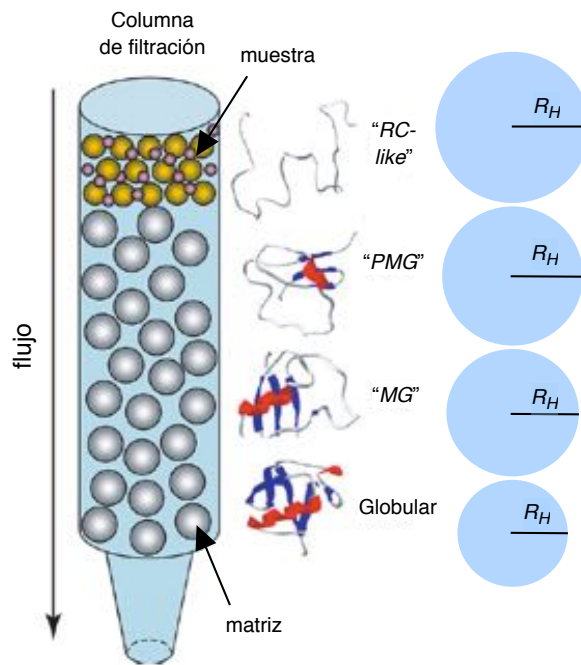


Figura 1.17: Comportamiento hidrodinámico de proteínas. Representación esquemática de los R_H asociado a proteínas en distintas conformaciones y los principios físicos de la separación molecular por SEC. La muestra está representada por moléculas de pequeño y gran tamaño. Las moléculas más grandes, migran más rápido que las pequeñas. A la derecha, los volúmenes hidrodinámicos ocupados una cadena polipeptídica larga en 4 conformaciones distintas: “RC-like”, “PMG”, “MG” y globular. Las esferas muestran un incremento en los R_H cuando una proteína del mismo largo se encuentra progresivamente más extendida (Adaptado de [21])

Las formas de medir R_H son varias: PFG-RMN (Resonancia Magnética Nuclear de Gradiente de Campo de Pulso), SAXS (Dispersión de Rayos-X de ángulo pequeño), DLS (Dispersión de luz dinámica), SEC o FG (Cromatografía de Exclusión Molecular o Filtración por gel). También se pueden predecir los R_H de las proteínas a partir de formulaciones desarrolladas en función de datos empíricos y que sólo consideran su longitud (número de residuos) y su naturaleza [117, 118].

En la Cromatografía de Exclusión Molecular, es posible asociar los volúmenes de elución (V_e) de las proteínas a sus PM aparentes (PM_{ap}) con los que eluyen, si la columna es

calibrada utilizando proteínas globulares de peso molecular conocido [21]. Cuando el PM_{ap} hallado por esta metodología es mayor al PM_{TEO} de una proteína, puede indicar un estado oligomérico de la misma o dar idea de su comportamiento hidrodinámico extendido. Justamente, los PM hallados por SEC son aparentes porque se relacionan con los R_H de las proteínas (Figura 1.17). Con la relación entre PM_{ap} y R_H estudiadas para de múltiples proteínas en diversas conformaciones y condiciones, ha sido posible desarrollar y validar formulaciones empíricas que permiten el cálculo del radio hidrodinámico [119, 120].

Por otro lado, el estudio conformacional de una proteína puede enriquecerse, si acoplado a la Cromatografía de Exclusión Molecular se hace un análisis de dispersión de luz láser estática (SLS) de la muestra conteniendo a la proteína. El SLS permite determinar -en forma independiente al comportamiento hidrodinámico- el peso molecular exacto de las poblaciones de especies aisladas en solución a partir de sus propiedades de dispersión de luz. De manera indirecta permite hallar el grado de oligomerización de la proteína de interés, discerniendo los casos entre estado oligomérico o R_H extendido debido a los cuáles el PM_{apSEC} pudo haber sido mayor al PM_{TEO} de una proteína en estudio [121].

La caracterización conjunta de SEC-SLS se llevó a cabo para AdE1A.WT y luego para las mutantes. Esto se llevó a cabo sembrando 100 μ l de una muestra (270 μ M para AdE1A.WT y AdE1A- Δ L y 540 μ M para AdE1A- Δ E) en una columna Superdex 75 analítica, que luego pasó por los equipos de detección dispuestos en tándem de dispersión de luz y de índice de refracción (Ver Materiales y Métodos).

Para AdE1A.WT este análisis mostró la elución de un único pico con una señal de índice de refracción homogénea que se correspondió con una especie de PM_{SLS} 13.93 kDa. Considerando el error experimental de la técnica (alrededor del 10-20% respecto del PM_{TEO}) este valor fue consistente con el PM_{TEO} monomérico de AdE1A.WT el cual es 12.5 kDa, de modo que la relación PM_{SLS}/PM_{TEO} asumió un valor 1.11 (Tabla 1.2). El valor de PM_{SLS} además dista mucho del PM_{TEO} que podría presentar especies diméricas o mayores de AdE1A.WT.

En el caso de las mutantes, el análisis también mostró señales de índice de refracción homogéneas correspondientes a especies únicas de PM_{SLS} 12.96kDa para AdE1A- Δ E y 11.58kDa para AdE1A- Δ L, ambos consistentes con sus valores de PM_{TEO} , indicando también estados monoméricos en solución para ambas mutantes (Tabla 1.2).

Los PM_{apSEC} de AdE1A.WT y mutantes fueron calculados a partir de la curva de calibración construída utilizando patrones de corrida: BSA, MBP y Lisozima (Figura 1.18 y ver

Materiales y Métodos). El valor PM_{apSEC} de AdE1A.WT fue 54.4 kDa, lo que revela -una vez confirmado su estado monomérico- que presenta un comportamiento hidrodinámico muy extendido. Tal es así, que la relación PM_{apSEC}/PM_{TEO} para la cual se espera un valor cercano a 1 en el caso de proteínas globulares monoméricas compactas, asume en este caso un valor 4.35 (Figura 1.18 y Tabla 1.2).

En el caso de AdE1A- ΔE , el análisis por SEC reveló que eluye con un PM_{apSEC} de 52.7 kDa y siendo su PM_{TEO} 12.1 kDa, la relación PM_{apSEC}/PM_{TEO} asumió un valor 4.36, indicando también que posee un comportamiento hidrodinámico altamente extendido. Para AdE1A- ΔL , los resultados fueron similares, el PM_{apSEC} fue de 53.7 kDa y con un PM_{TEO} de 12.2 kDa, la relación PM_{apSEC}/PM_{TEO} asumió un valor 4.40, revelando el mismo comportamiento hidrodinámico que para las otras dos variantes, altamente extendido (Figura 1.18 y Tabla 1.2).

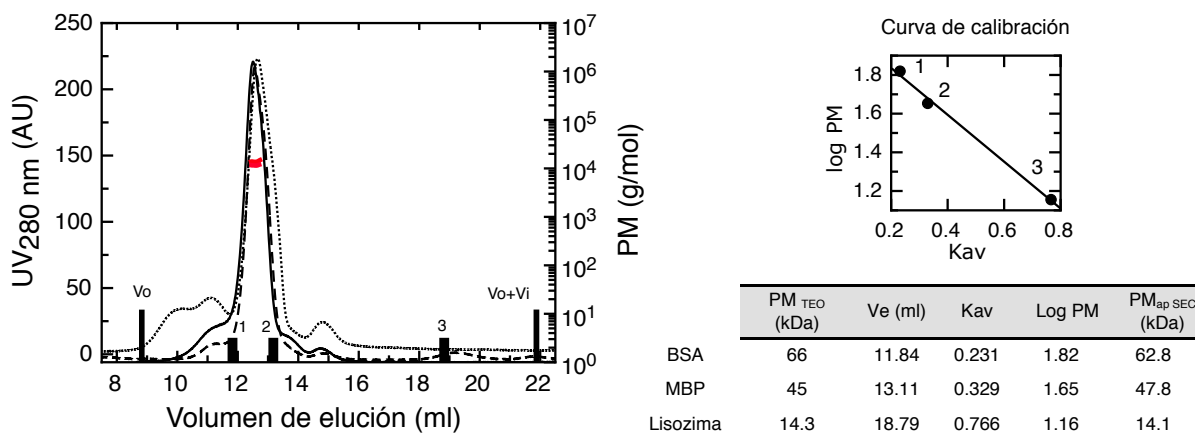


Figura 1.18: Comportamiento hidrodinámico y estado de oligomerización de AdE1A por SEC-SLS. A la izquierda, cromatograma de exclusión molecular de AdE1A.WT (línea continua), AdE1A- ΔE (línea punteada) y AdE1A- ΔL (línea de rayas) mostrando la elución de picos únicos seguidos a 280nm y sus correspondientes señales homogéneas de índice de refracción (líneas rojas) como medidas de los PM por SLS. Los volúmenes de la columna están señalados al igual que los V_e de las proteínas usadas para la calibración: BSA (1), MBP(2) y Lisozima (3). A la derecha, la curva de calibración (Logaritmo del Peso molecular vs coeficiente de partición -Kav-) construida para la determinación del PM aparente por SEC de AdE1A.WT y la tabla con los valores de dicha calibración.

Tabla 1.2 Propiedades hidrodinámicas de AdE1A.WT, AdE1A- ΔE y AdE1A- ΔL

	V_e (ml)	PM_{TEO} (kDa)	PM_{SLS} (kDa)	PM_{apSEC} (kDa)	Estado de oligomerización PM_{SLS}/PM_{TEO}	Comportamiento hidrodinámico PM_{apSEC}/PM_{TEO}
AdE1A.WT	12.51	12.5	13.93	54.4/ 59.68 \pm 0.03 ^a	1.11 (monómero)	4.35/ 4.77 ^a (extendido)
AdE1A- ΔE	12.66	12.1	12.96	52.7	1.07 (monómero)	4.36 (extendido)
AdE1A- ΔL	12.57	12.2	11.58	53.7	0.95 (monómero)	4.40 (extendido)

a. Resultado de promediar tres medidas independientes de DLS

Una vez hechos estos análisis se procedió a encontrar los R_H experimentales de las tres proteínas utilizando las relaciones empíricas desarrolladas por Uversky [119, 120] y para el caso de AdE1A.WT, utilizando DLS. También se procedió a hallar los R_H teóricos predichos, utilizando las formulaciones empíricas propuestas por Wilkins y por Marsh y Forman-Kay [117, 118]. En la Tabla 1.3 se resumen estos resultados.

Tabla 1.3 Cálculo de los R_H experimentales y R_H predichos para AdE1A.WT, AdE1A- Δ E y AdE1A- Δ L

	N^a	PM_{TEO} (kDa)	PM_{ap} (kDa)	R_H exp		R_H predichos										
				DLS ⁴	Uversky						Wilkins		Marsh & Forman-Kay			
					G	G ¹	D ¹⁻²	D ¹⁻³	MG ¹	PMG ¹	G	D	G	D	IDP	IDPc ⁵
AdE1A.WT	114	12.5	54.4	3.42± 0.01	3.07	1.81	3.17	3.06	2.07	2.49	1.88	3.29	1.90	3.14	2.77	3.49
AdE1A- Δ E	113	12.1	52.7	nd	3.03	1.79	3.12	3.01	2.04	2.46	1.87	3.27	1.89	3.12	2.76	3.46
AdE1A- Δ L	113	12.2	53.7	nd	3.05	1.80	3.13	3.02	2.05	2.47	1.87	3.27	1.89	3.12	2.76	3.46

Todas las medidas de R_H están expresadas en nm. Para proteínas G: globular, D: desnaturalizado, MG: "Molten-Globule", PMG: "Pre-Molten-Globule", IDP: Intrínsecamente desordenada.

^a N corresponde al número de aminoácidos en la secuencia.

¹ Calculado usando los valores de PM_{TEO}

² Calculado a partir la fórmula obtenida de condición desnaturalización en presencia de Gdm.Cl.

³ Calculado a partir la fórmula obtenida de condición desnaturalización en presencia de Urea.

⁴ Resultado de tres determinaciones independientes

⁵ Calculado usando la fórmula mejorada para el cálculo de R_H de IDPs teniendo en cuenta los componentes de secuencia.

nd no determinados

En su trabajo a partir de datos de literatura preexistentes Uversky encontró una relación entre los R_H y PM_{TEO} para proteínas en distintas condiciones: globular (G), desnaturalizados (D) por agentes caotrópicos como Cloruro de guanidinio (Gdm.Cl) y Urea; proteínas en estado "Molten-Globule" (MG) y "Pre-Molten-Globule" (PMG). Estas relaciones empíricas que tienen la forma $\log(R_H)$ vs $\log(PM_{TEO})$ (Ver Materiales y Métodos), fueron confirmadas luego por sus propios experimentos SEC bajo las mismas condiciones. De esta manera, si se utiliza SEC para hallar los PM_{apSEC} , como se hizo para AdE1A.WT y mutantes, estos pesos moleculares pueden ser usados para calcular sus R_H experimentales.

Los R_H que se calcularon según las formulaciones de Uversky son 6 a los fines comparativos. En el primero de los presentados en la Tabla 1.3 (G) se utilizaron los datos de PM_{apSEC} hallados y los R_H –experimentales- asumieron valores entre 3.07 nm (AdE1A.WT) y 3.05 nm (AdE1A- Δ E). Los siguientes R_H se calcularon utilizando el PM_{TEO} de las tres proteínas AdE1A, bajo las condiciones impuestas por cada formulación empírica: si fuesen compactas como proteínas globulares y eluyesen como tales de la columna SEC (G¹); si siendo globulares estuviesen eluyendo desnaturalizadas en presencia de Gdm.Cl (D^{1/2}); si siendo globulares estuviesen eluyendo desnaturalizadas en presencia de Urea (D^{1/3}); si estuviesen eluyendo

como “*Molten-Globule*” (MG) y si estuviesen eluyendo en conformación “*Pre-Molten-Globule*” (PMG). Como se puede observar la mayor concordancia para los valores R_H experimentales de los AdE1A (los deducidos a partir de PM_{apSEC}) se encuentra con aquellos R_H calculados considerándolas proteínas globulares totalmente desnaturalizadas ($D^{1/2}$, $D^{1/3}$), lo que revela su alto grado de desorden y comportamiento extendido en su forma nativa.

Para AdE1A.WT, también se llevó a cabo la determinación de R_H experimental por dispersión de luz dinámica (DLS). Este R_H asumió un valor de 3.42 ± 0.01 nm (Tabla 1.3) relativamente comparable con el valor de R_H experimental por SEC hallado por la formulación de Uversky. Este análisis además reveló la presencia de una población de partículas monodispersa -conformacionalmente homogénea- de la que se estimó un PM_{apDLS} de 59.68 ± 0.03 kDa (Tabla 1.2), concordante con el valor de PM hallado por SEC.

Por otro lado, se calcularon los R_H teóricos predichos para lo cual se utilizaron las formulaciones de Wilkins y por Marsh y Forman-Kay, quienes las elaboraron con una estrategia diferente a la de Uversky. Ambos se basaron en la Ley de escalamiento para dimensiones de homopolímeros, la cual relaciona bajo la fórmula $R_H = R_0 N^u$ (Ver Materiales y Métodos) el radio hidrodinámico del polímero con el número de sus monómeros (N), una constante (R_0) y un factor de escalamiento (u) que depende de las condiciones en las que se encuentra. Wilkins encontró valores R_0 y u , aplicable para proteínas globulares (G) y desnaturalizadas (D), trabajando con una serie de proteínas en dichas conformaciones a las que le determinó su R_H por PFG-RMN y que luego validó por comparación con datos de literatura [118]. Por su parte, Marsh y Forman-Kay, hallaron otros valores para estos parámetros aplicables a proteínas globulares (G) y desnaturalizadas (D) e intrínsecamente desordenadas (IDP) habiendo trabajado con datos de literatura de proteínas en dichas conformaciones. Además, lograron una mejora en la formulación empírica para R_H de IDP al considerar información de su secuencia [117] (Ver Materiales y Métodos).

Los resultados de los cálculos para R_H predichos de AdE1A.WT y mutantes utilizando estas últimas formulaciones están en la Tabla 1.3. Las predicciones de R_H calculadas usando las formulaciones de Wilkins, se acercan más a los valores experimentales de R_H hallados por SEC y DLS, si se las considera como IDPs. Cuando se utilizan las formulaciones de Marsh y Forman-Kay es llamativo como las predicciones de R_H se asemejan más a los R_H experimentales si se considera a estas proteínas como desnaturalizadas (D), en tanto que con la fórmula básica válida para IDP, los R_H se predicen mas compactos. No obstante, con la fórmula mejorada para IDPs ($IDPc^3$) al menos el valor predicho de R_H para AdE1A.WT es

sumamente concordante con el R_H experimental hallado por DLS. Esto probablemente se deba a que en dicha formulación se da un peso significativo al contenido de prolinas y de residuos cargados presentes en la secuencia, los cuáles son determinantes en el grado de compactación de una IDP [117] como sería el caso de AdE1A. AdE1A posee un total de 28 residuos cargados en su secuencia, de los cuáles 26 negativos, así como 18 prolinas a lo largo de una secuencia de 114 residuos. Las cargas negativas podrían estar generando un efecto repulsivo entre sí, lo que sumado a las Prolinas, un aminoácido disruptor de estructuras, podrían estar generando el alto grado de extensión en la cadena polipeptídica que se ve reflejado en sus valores R_H y la caracterizan como una IDP tipo “*random coil-like*”.

En conclusión, los resultados de SEC-SLS-DLS en conjunto confirmaron que AdE1A tiene un comportamiento hidrodinámico típico de IDP “random coil-like”, que son altamente extendidas, encontrándose además en estado nativo en solución como una proteína monomérica en una población conformacionalmente homogénea. Por otra parte también, estos resultados también confirman que ninguna de las mutaciones genera cambios sustanciales en dichas propiedades conformacionales en comparación con la variante natural.

I.2.D. Evaluación del estados de oxidación de AdE1A

En los experimentos SEC se encontró que, adicionalmente al problema de la degradación de AdE1A.WT tras ciclos de congelado y descongelado, también se formaba una especie de mayor PM_{ap} al caracterizado para la proteína monomérica (Figura 1.19 A). Pero cuando estos dos picos obtenidos por SEC se aislaban y eran analizados por SDS-PAGE se observaba que ambos corrían con idéntica movilidad electroforética (Ver “*Marcación de AdE1A.WT y rendimiento*”).

Como AdE1A.WT posee una única cisteína en la posición 124 (motivo LxCxE) de su secuencia natural y es altamente reactiva, una posibilidad es que la misma habría estado mediando una dimerización de esta proteína por formación de un puente disulfuro con una segunda cadena de AdE1A. Esto explicaría el por qué esta segunda especie poseía un mayor PM_{ap} y no se evidenciaba en los geles SDS-PAGE puesto que los mismos se hacen en condiciones reductoras.

Para comprobar si la segunda especie de mayor peso molecular era producto de la oxidación de cisteínas, una muestra con múltiples ciclos de congelado y descongelado fue dividida y sólo una parte fue tratada con 50 mM DTT a temperatura ambiente por 2 horas y la otra se dejó sin modificar, tras lo cual ambas muestras se corrieron en SEC analítica.

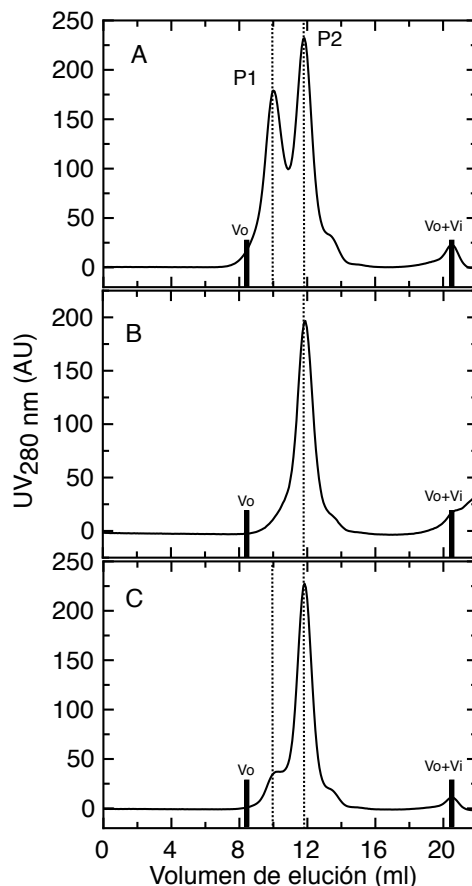


Figura 1.19: Análisis del estado de oxidación de stocks de AdE1A.WT por SEC. A) Cromatograma de la especie parcialmente oxidada con picos de AdE1A.WT reducida (P2) y la probable especie oxidada dimérica (P1). B) Cromatograma de la especie tratada con agente reductor. C) Cromatograma de la muestra control, la cual también evidencia un pequeño hombro correspondiente a la especie oxidada.

La Figura 1.19 muestra los resultados de este experimento. En el panel A, el cromatograma de 300ug de la proteína *stock* presenta dos picos: P2 correspondía sin dudas a la especie AdE1A.WT (cuyo $K_{av} = 0.283$, fue equivalente al $K_{av} = 0.282$ de la especie identificada como AdE1A.WT monomérica en los experimentos anteriores (Tabla 1.2)) y un P1, que eluye antes ($K_{av} = 0.127$) y que no se había identificado, pudiendo corresponder a una probable especie oxidada dimérica. En el panel B, correspondiente al cromatograma de 150ug del mismo *stock* tras el tratamiento con DTT, se puede ver que ya no se aprecia la especie de mayor peso molecular y solo eluye un único pico (con un $K_{av} = 0.284$), correspondiente a la especie reducida de AdE1A.WT idéntico al pico al que eluye un *stock* control de reciente obtención ($K_{av} = 0.282$), en el panel C.

Como era esperable, este comportamiento también se observó en la mutante AdE1A- ΔE , no así en la mutante AdE1A- ΔL , ya que la misma no posee la C124 reactiva responsable de la dimerización de AdE1A.

A partir de este resultado, se decidió conservar los stocks de proteínas en 2mM DTT y tratarlas con alta concentración de dicho agente reductor al menos 2 horas antes de realizar cualquier clase de experimento a fin de asegurar su correcto estado reducido.

I.2.E. Análisis del contenido de estructura secundaria de AdE1A.WT por Dicroísmo Circular

Una técnica que permite el análisis conformacional de proteínas es el Dicroísmo Circular (CD, por sus siglas en inglés). La misma se basa en la absorción diferencial de luz circularmente polarizada hacia la izquierda y la derecha a una dada longitud de onda, por parte de cromóforos asimétricos. En las proteínas, los enlaces amida y residuos aromáticos - naturalmente simétricos por su geometría planar- son los responsables de producir las señales medidas porque el complejo plegamiento proteico generan sobre ellos el ambiente asimétrico necesario para que experimenten dicho fenómeno [122].

Las proteínas presentan bandas de CD en el UV lejano (175-260 nm) y en el UV cercano (260-300nm) (Figura 1.20). El análisis en el UV lejano aporta información sobre el contenido de estructura secundaria basado en la asimetría generada por el plegamiento de la cadena polipeptídica [123]. Elementos de estructura secundaria canónica pura como α -hélice y lámina β presentan un patrón de bandas de CD característicos y también elementos como los giros β y Poliprolina tipo II, aunque mas variables y menos intensas. Estos tipos de estructuras secundarias presentan las siguientes características:

α -hélice. Es el mayor elemento estructural en una proteína y corresponde a una estructura helicoidal donde cada giro tiene una longitud de 5.4 Å y está formada por 3.6 residuos. Su espectro típico de CD presenta dos mínimos característicos a 208 nm y a 220 nm y una banda positiva a 195 nm (Figura 1.20, línea roja)

Lámina- β . Corresponde a dos o más cadenas- β unidas entre sí por puente de hidrógeno, pudiendo estar en posición paralelas o antiparalelas. Su espectro de CD típico es más variable que el de las α -hélices, con una banda positiva alrededor de 198 nm y una sola banda negativa que puede estar entre 214 y 218 nm (Figura 1.20 línea azul).

Hélice de poliprolinas tipo II (PPII). Es una una hélice levógira y extendida de 3 residuos por vuelta y una longitud por giro es de 9.3 Å. El espectro de CD típico de PII se caracteriza por presentar una banda positiva alrededor de 218 nm y un mínimo a 198 nm (Figura 1.20, línea roja)

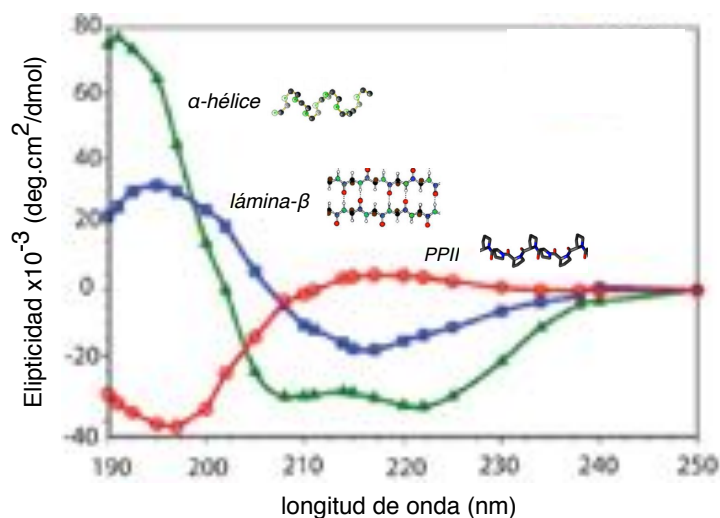


Figura 1.20: Espectros de CD en el UV lejano típicos de las estructuras secundarias canónicas y no canónicas. Se observan las bandas positivas y negativas características de cada estructura. Al lado de las identificaciones, representaciones esquemáticas de estas estructuras (Adaptado de [21])

Ahora bien, las proteínas pueden poseer distintos tipos de estructura secundaria y en cantidad variables, y por tanto tienen espectros de CD que resultan de la combinación de bandas características. Las IDPs pueden tener espectros similares a los esperados para polipéptidos tipo “*random coil*”, aunque esto varía con el grado de contenido de estructura secundaria que puedan poseer.

Para AdE1A.WT, AdE1A-ΔE y AdE1A-ΔL, las medidas de sus espectros de CD se realizaron como se indican en la sección Materiales y Métodos y la elipticidad molar residual media ($[\theta]_{MRW}$) para cada una fue calculada y representada como lo muestra la Figura 1.21. Como se puede observar las tres proteínas presentaron espectros superponibles entre sí. Para los tres casos, sus espectros presentaron una banda negativa con un mínimo a 200nm y ausencia de una banda positiva a 220 nm.

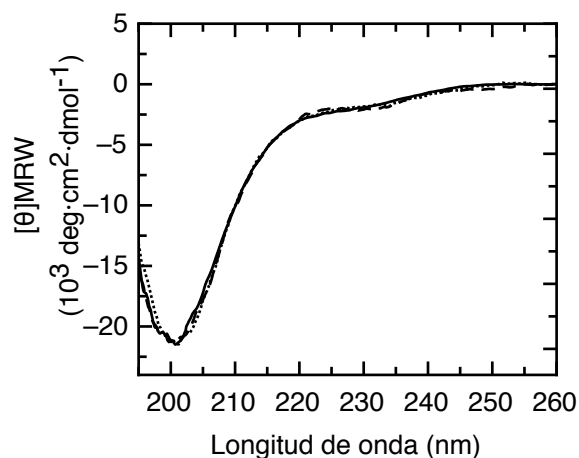


Figura 1.21: Espectros de CD en el UV lejano para AdEA1.WT y mutantes. Se observa una clara superposición de los espectros de CD de AdE1A.WT (línea continua) AdE1A-ΔE (línea punteada) y AdE1A-ΔL (línea de rayas), característicos de IDPs.

En resumen, el tipo de espectro descrito es justamente el que se refiere regularmente como uno típico para “random coils”, lo que indica que AdE1A se trataría de una proteína carente de estructuras secundarias a lo largo de su secuencia, clasificándola como una IDPs tipo “random coil-like”.

Estudio de elementos de estructura secundaria residual en AdE1A

Dado que las IDPs presentan gran heterogeneidad conformacional y que a diferencia de polipéptidos tipo “*random coil*”, determinadas regiones pueden presentar preferencias conformacionales específicas, el contenido de estructuras secundarias que exhiben es dinámico [120]. Como estos elementos estructurales secundarios pueden ser modificados, estabilizados o perturbados por agentes físicos o químicos, interacciones o modificaciones post-traduccionales, es útil llevar a cabo medidas de dicroísmo circular en distintas condiciones y observar su comportamiento como parte de la descripción de su naturaleza desordenada [122]. Por estas razones sobre AdE1A.WT se llevó a cabo una serie de análisis de transiciones conformacionales por acción del solvente TFE y en distintas condiciones de pH y temperatura.

Estabilización del contenido de estructura α -hélice en AdE1A.WT por acción del TFE

El 2,2,2-trifluoroetanol (TFE) es un solvente, que como algunos otros (ej. 2-clorobutanol) es utilizado para estabilizar estructuras α -hélice preexistentes en regiones en los que esta conformación se encuentra en equilibrio con su forma desordenada [124, 125].

Para evaluar la presencia de α -hélices locales y transitorias en AdE1A.WT se realizó un experimento de titulación midiendo los espectros de CD en el UV lejano de una muestra con contenido creciente de TFE. Se adquirieron 14 espectros en una muestra de 0 a 65% TFE v/v, con incrementos de a 5% y los datos fueron corregidos por dilución y convertidos a elipticidad molar. Como puede ser observado en la Figura 1.22, el TFE produjo un incremento en el valor absoluto de la banda a 222nm y un desplazamiento en la posición del mínimo de 200 a 208 nm y descenso de su intensidad, indicando pérdida de desorden y estabilización de estructura α -hélice. Cabe destacar que a la señal de elipticidad 222nm también contribuyen levemente el contenido de estructuras tipo PPII y residuos aromáticos. No obstante esta contribución debería ser menor, particularmente por el hecho de que AdE1A.WT posee solo dos tirosina como aminoácidos aromáticos en su secuencia de 114 residuos.

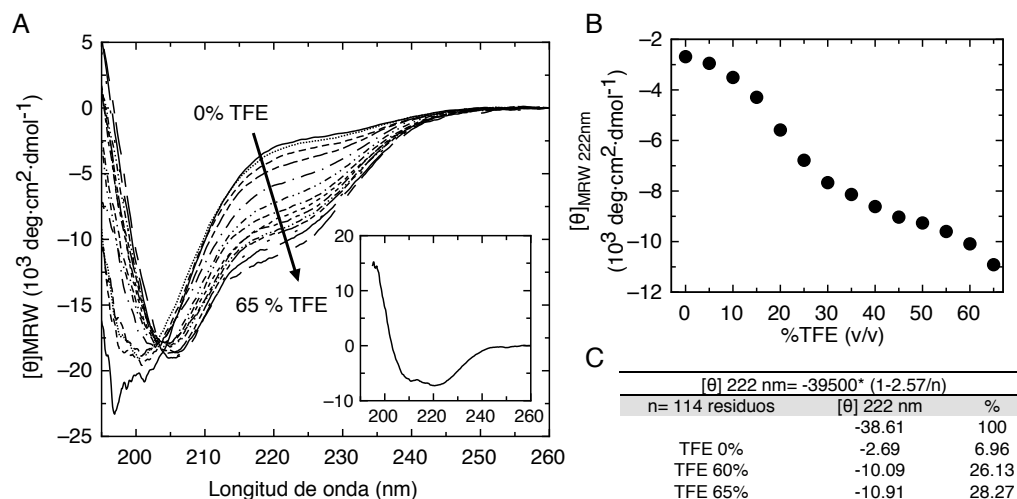


Figura 1.22: Caracterización de la propensión α -hélice de AdE1A.WT por titulación con TFE seguida por dicroísmo circular. A) Espectros de CD en el UV lejano para AdE1A.WT en concentraciones crecientes de TFE, expresados en % v/v. En el gráfico interno, espectro de CD característico de α -hélice resultado de la diferencia entre los valores de CD a 60 y 0% TFE. B) Curva de titulación siguiendo los valores de elipticidad molar a 222nm en función de la concentración creciente de TFE. C) Porcentajes de residuos en α -hélice calculados a partir de los valores de elipticidad molar a 222nm en distintas condiciones.

Como se puede observar en la Figura 1.22 B, la transición parece mostrar un efecto cooperativo mínimo, pero al no observarse punto isodicroico (Figura 1.21 A) no fue posible aplicar el modelo de equilibrio de dos estados “coil”-hélice. Por lo tanto, el cambio en la elipticidad molar a 222nm en función de TFE% no pudo ser ajustado a dicho modelo para obtener parámetros termodinámicos de la transición. Esto indica que probablemente en el proceso de estabilización estarían involucrados dos o más transiciones conformacionales a α -hélice.

El porcentaje de residuos en AdE1A.WT estabilizados en conformación α -hélice antes y después del agregado de TFE fue obtenido comparando el valor de elipticidad molar en 0, 60 y 65% TFE con el valor de elipticidad molar a 222nm de AdE1A.WT con hipotética conformación 100% en α -hélice. Estos resultados están presentados en la Tabla de la Figura 1.22 y muestran que al estado nativo solo un escaso 7% de residuos en AdE1A.WT se encuentra en dicha conformación. Esto puede implicar que al poseer 114 residuos totales, unos 8 residuos contiguos en AdE1A.WT se encuentren en conformación 100% hélice en el estado nativo, o que tal conformación esté distribuida en un mayor número de residuos de la cadena carbonada. El porcentaje de residuos en α -hélice crece al 28% en condiciones saturantes de TFE, implicando también que cerca de 32 residuos se encuentren en conformación 100% hélice bajo dicha condición o que más bien se encuentren distribuidos en la cadena polipeptídica.

Predicción de α -hélice en la secuencia de AdE1A

A fin de evaluar cuáles podrían ser las regiones en AdE1A.WT que son más propensas a adquirir estructuras tipo α -hélice, y que podrían estar experimentando estas transiciones en presencia de TFE, se realizó una predicción de tendencia α -hélice con el algoritmo AGADIR [126]. Este algoritmo predice la tendencia local de una secuencia a adoptar estructura α -hélice en ausencia de plegamiento terciario. La Figura 1.23 muestra el resultado de esta predicción.

Como se puede observar en la Figura 1.23 se pueden identificar 4 regiones con tendencia que podrían estar poblando estructuras α -hélice, como fue demostrado en estos experimentos en presencia de TFE. Según estas predicciones la mayoría de estas regiones corresponden a los motivos de interacción que presenta AdE1A.WT y que se sabe que median interacciones con RbAB (motivos E2F y LxCxE) [84, 108] y CBP/p300 (motivo TRAM-CBP) [76, 82].

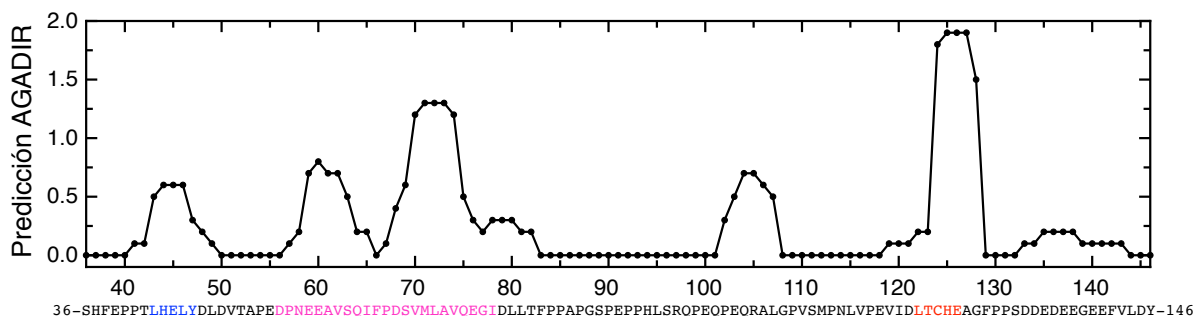


Figura 1.23: Predicción de tendencia local a estructurar α -hélice en la secuencia de AdE1A.WT usando el programa AGADIR. A) La medida presentada se realizó a pH7.0. En la secuencia de AdE1A.WT (abajo) se señalan con distintos colores los motivos lineales de interacción, los cuáles parecen presentar la mayor tendencia a estructurar α -hélice. En azul, el motivo E2F, en magenta, el motivo TRAM-CBP y en rojo, el motivo LxCxE.

Efecto del pH en la inducción de estructura α -hélice en AdE1A.WT

Dado que la fracción de cargas negativas en AdE1A.WT es significativa, ya que posee 26 residuos ácidos (Asp y Glu) sobre un total de 114 ($f^- = 0.228$), las mismas podrían participar en la modulación de las conformaciones α -hélice en esta proteína con el cambio de pH. La disminución de pH, genera la protonación de los carboxilatos de los cadenas laterales de los residuos ácidos neutralizando su carga ($pK_a^{\beta\text{-COOH}} \text{ Asp} = 3.9$ y $pK_a^{\gamma\text{-COOH}} \text{ Asp} = 4.07$). Esto permite la propagación de la conformación en α -hélice de las regiones que ya estructuraron en presencia de TFE, hacia regiones acídicas contiguas. Este mismo efecto podría darse por la protonación del grupo Imidazol de los residuos His ($pK_a^{\text{Imidazol}} \text{ His} = 6.04$).

Para evaluar la influencia del pH en las transiciones a α -hélice con TFE, se hicieron medidas de CD utilizando tampones a pH 7.0 y pH 5.0. Los espectros de CD de soluciones de

AdE1A.WT a 0 y 50% TFE en tampón fosfato de sodio pH 7.5 y formiato de sodio pH 5.0 fueron adquiridos, corregidos por dilución y convertidos a elipticidad molar.

En la Figura 1.24 se puede observar como la estructura α -hélice es inducida a 50% TFE a ambos pH, aunque hay un ligero incremento a pH 5.0 (26.70% vs 23.98%). No obstante, la neutralización de cargas a bajo pH no parece producir grandes incrementos en la propensión a α -hélice en AdE1A.WT.

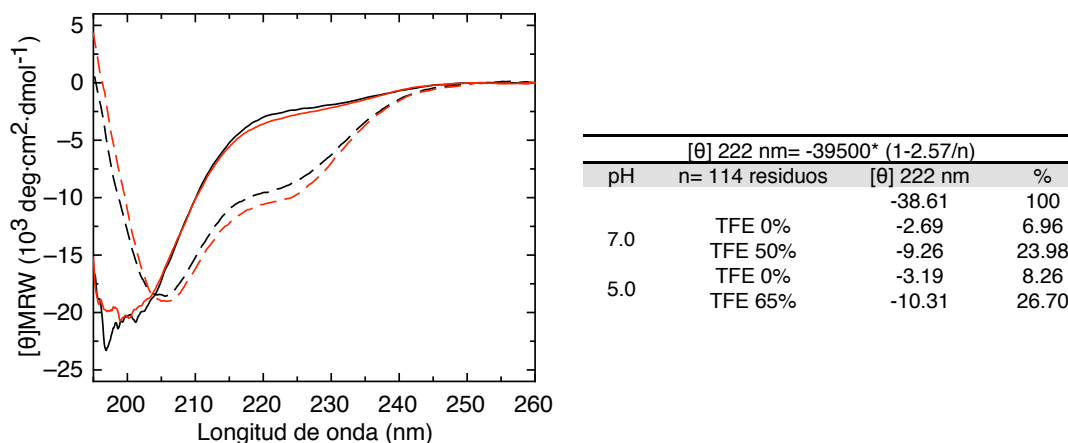


Figura 1.24: Dependencia de la propensión α -hélice de AdE1A.WT con el pH. A la derecha, espectros de CD en el UV lejano para AdE1A.WT en concentraciones de 0% TFE a pH 7.0 (línea continua negra), 50% TFE a pH 7.0 (línea de rayas negra), 0% TFE a pH 5.0 (línea continua roja) y 50% TFE a pH 5.0 (línea de rayas roja). A la izquierda, porcentajes de residuos en α -hélice calculados a partir de los valores de elipticidad molar a 222nm en las distintas condiciones.

También se probaron condiciones de menor pH (4.0 y 3.0) pero se observó que a medida que la muestra se acercaba al pI teórico de AdE1A.WT (pI= 3.78) la misma experimentaba agregación macroscópica, porque probablemente con la neutralización total de sus residuos ácidos se produzca el colapso de la cadena polipeptídica.

Propensión a estructura de Poliprolina tipo II en AdE1A.WT

Las hélices de Poliprolina tipo (PPII) constituyen una conformación no canónica que es predominante en proteínas desordenadas [127] y polipéptidos desplegados por acción de agentes caotrópicos (como Urea y Cloruro de guanidinio) [128, 129]. Esta conformación puede ser adoptada por polipéptidos que pueden o no contener prolinas en su secuencia así como en aquellos con alto porcentaje de residuos cargados en su secuencia. Esta estructura no forma patrones regulares de puentes intra o intercatenarios y por lo tanto es marcadamente más flexible que las α -hélice y lámina- β [130]. Como AdE1A.WT posee un alto contenido de Prolinas, unas 18 sobre un total de 114 residuos, la mayoría de ellas en una *región rica en prolinas* del residuo 84 al 117 (11 prolinas), es probable que esta proteína tenga alta propensión a esta estructura secundaria no canónica.

Dado que las hélices PPII se estabilizan con bajas temperaturas [131], se procedió a adquirir los espectros de CD de AdE1A.WT a 85, 20 y 5 °C, con el fin de estudiar su propensión conformacional a esta estructura. Los datos fueron corregidos convertidos a elipticidad molar y representados en la Figura 1.25.

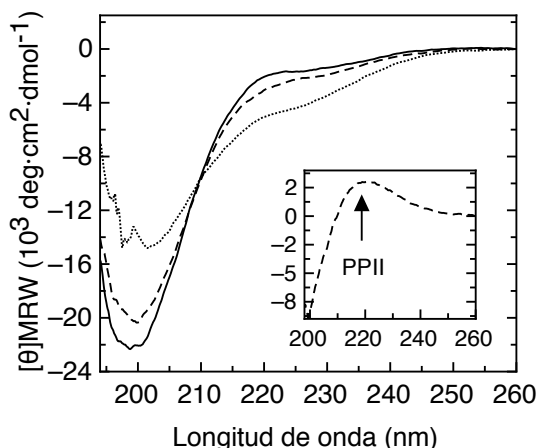


Figura 1.25: Estabilización de estructura tipo PPII en AdE1A.WT por bajas temperaturas. Espectros de CD en el UV lejano para AdE1A.WT adquiridos a 5 °C (línea continua), 20 °C (línea de rayas) y a 85 °C (línea de puntos). En el gráfico interno, espectro de CD característico con la banda a 218nm de PPII resultado de la diferencia entre los valores de CD a 5 y 85 °C.

Como se puede observar, con el descenso de la temperatura (Figura 1.25, línea continua), hubo una mayor señal positiva en la banda a 218nm y un aumento de la señal negativa en la posición del mínimo, además de presentarse un desplazamiento del mínimo hacia menores longitudes de onda (de 200 a 198 nm), lo que implica estabilización de estructuras tipo PPII. Estos cambios ocurridos también se pueden evidenciar en el espectro resta 5 – 80 °C, donde se ve claramente la formación de la banda a 218nm.

Concluyendo, esta serie de resultados indican que aun cuando en condiciones nativas AdE1A.WT posea un escaso contenido de conformaciones canónicas y no canónicas, tiene capacidad de estabilizar en forma transitoria elementos de estructura α -hélice y PPII, lo que revela su gran plasticidad conformacional.

I.2.F. Desnaturalización térmica de AdE1A.WT

La estabilidad de una proteína puede ser estudiada mediante el análisis de su proceso de plegamiento-desplegamiento en función de la temperatura. Para una proteína globular, que presenta estructura terciaria y contenido de estructura secundaria canónica, un aumento en la temperatura puede producir desplegamiento de la cadena polipeptídica en procesos cooperativos, reversibles o no, de los que se pueden obtener descriptores cuantitativos como la

T_m (Temperatura de desnaturalización). Para el caso de las IDPs, el comportamiento térmico difiere del caso de las proteínas globulares en tanto son carentes de estructuras terciarias consolidadas en solución que pueden plegar cooperativamente o poseen comparativamente menor contenido de regiones estructuradas; así, estas transiciones térmicas son generalmente reversibles y poco cooperativas [131].

La estabilidad térmica de AdE1A se estudió mediante curvas de termodesnaturalización, que se obtuvieron siguiendo la señal de elipticidad a 200nm al aplicar un gradiente de temperatura de 20 a 80 °C y otro de 80 a 20 °C, a una velocidad de barrido de 3°C/min. Como se puede observar en la Figura 1.26, el proceso mostró una variación lineal de la elipticidad (a 200 nm) con la temperatura, completamente reversible, no mostrando evidencias de desplegamiento cooperativo. Además, al finalizar el gradiente, no se observó agregación macroscópica alguna en la muestra y el espectro de CD fue superponible al adquirido antes de aplicar dicho gradiente.

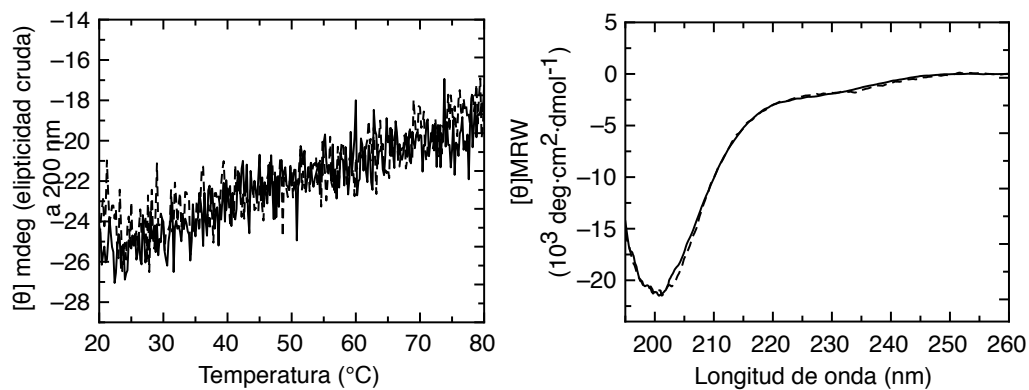


Figura 1.26: Desnaturalización térmica de AdE1A.WT. A la izquierda, elipticidad cruda a 200nm en función de la temperatura en barridos desde 20 a 80 °C (línea continua) y de 80 a 20 °C (línea a rayas). A la derecha, espectros de CD en el UV lejano para AdE1A.WT adquiridos a 20 °C antes (línea continua) y después (línea de rayas) de los barridos térmicos.

Resumiendo, estos resultados demuestran por un lado ausencia de estructuras (secundarias y terciarias) consolidadas con plegamientos cooperativos en su estado nativo, así como una alta estabilidad térmica al no poseer dichas estructuras.

OBJETIVO I.3

Obtención de reactivos para espectroscopía de fluorescencia.

I.3.A. Péptidos y proteínas marcados con FITC. Protocolo de marcación y rendimiento para AdE1A.WT

Para poder llevar a cabo los ensayos espectroscópicos de fluorescencia que permitieron la determinación de parámetros termodinámicos de la interacción con RbAB (Ver Capítulo II), se trabajó con los siguientes 5 péptidos y 3 proteínas:

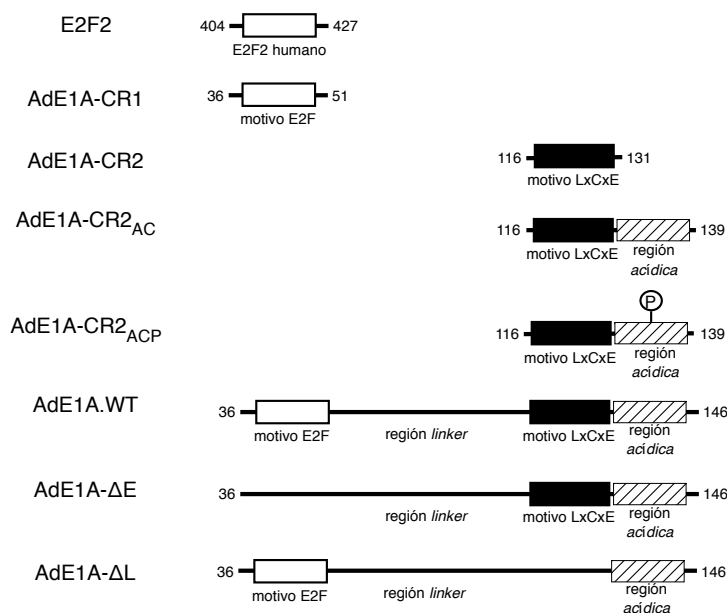


Figura 1.27: Esquemas de péptidos y proteínas utilizados en este trabajo y que fueron marcados con FITC. Para todos los reactivos se señalan la nomenclatura utilizada, los límites de los fragmentos con la posición de los residuos en su secuencia natural y las regiones en ellos comprendidos.

Todas estas especies fueron marcadas con isotiocianato de fluoresceína (FITC). La metodología de marcación que se siguió para péptidos -que fueron obtenidos de manera sintética y cuyo grado de pureza era superior al 98%- ya se encontraba puesto a punto en el laboratorio. Para el caso de la marcación de proteínas con FITC, el proceso requirió mayor análisis y un paso extra de purificación.

El FITC es una molécula fluorocromo capaz de emitir fluorescencia (con su máximos a λ_{em} a 520 nm) cuando es excitado a λ_{ex} a 495nm. Su pequeño tamaño y alta reactividad la hacen útil para marcar macromoléculas (como péptidos y proteínas) a fin de estudiar propiedades de éstos últimos por métodos de fluorescencia. FITC reacciona con grupos aminos libres. A pH utilizado (pH 8.0/9.0) los grupos α -amino libres (amino-terminal) se encuentran desprotonados pero no así grupos ϵ -amino que tienen un pKa más alto (pKa α -NH₃⁺ = 8.9 vs pKa ϵ -NH₃⁺ = 10.3), de modo que el FITC reacciona preferencialmente con el grupo amino-

terminal y menos probablemente con grupos amino de las cadenas laterales de los residuos lisina y arginina. Los péptidos utilizados en este trabajo no contienen ni lisinas ni argininas, por lo que el grupo amino terminal fue el único sitio posible para la incorporación de FITC. En el caso de las proteínas, éstas contienen 2 argininas, por lo que su marcación demandó chequear el producto final para evaluar la incorporación de la sonda en el amino terminal.

En el caso de la marcación de AdE1A.WT, ésta requirió la utilización de 3.3 mg proteína iniciales que fueron marcadas según protocolo y purificadas por columnas de desalado PD10 (Ver Materiales y Métodos). El eluido obtenido fue cuantificado por Absorbancia a 220 nm (para cuantificar proteína total) y a 495 nm (para cuantificar FITC). A este punto la cantidad total de proteína recuperada fue del 70% (2.3mg de proteína total) y el rendimiento de marcación (*razón F/P*) fue del 64%.

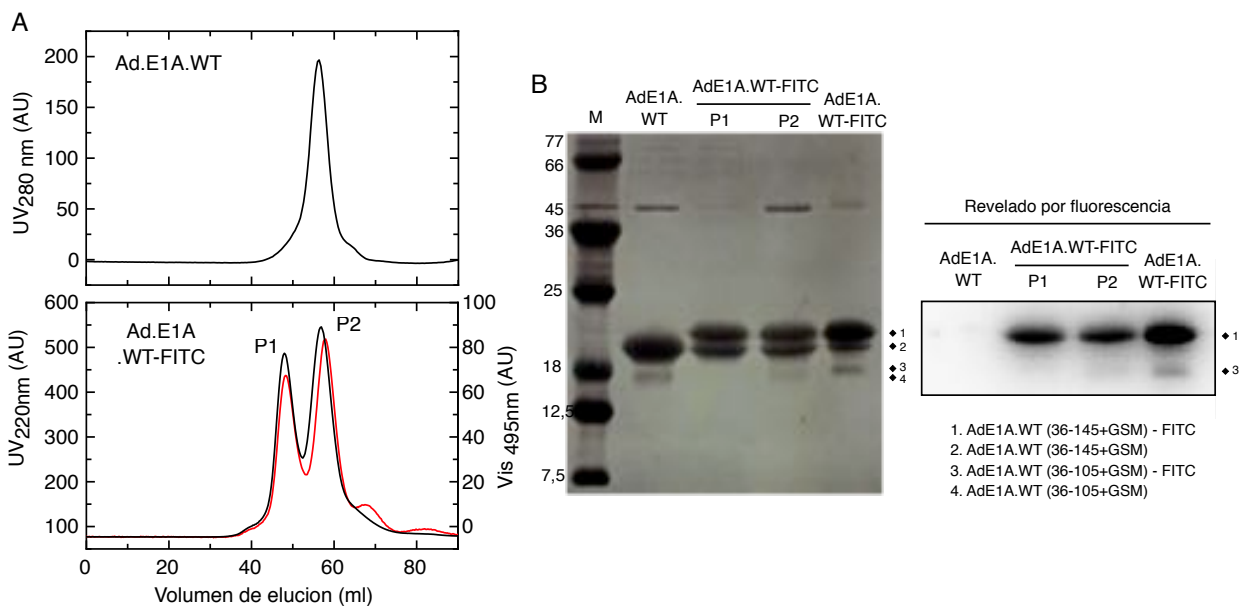


Figura 1.28: Cromatografía de exclusión molecular de AdE1A.WT marcada con FITC. A) Cromatograma de exclusión molecular control de AdE1A.WT reducida (arriba) y de AdE1A.WT post marcación con FITC seguida a 220nm (absorción de enlaces amida) y a 495nm (absorción de FITC). Como se observa, se separan dos picos correspondientes a la especie monomérica y dimérica de AdE1A.WT. B) A la derecha, gel SDS-PAGE 15% donde se observan el patrón de bandas de la muestra control, de los dos picos de AdE1A.WT post marcación y SEC y de AdE1A.WT post marcación y previa a la cromatografía. A la derecha, el revelado del mismo gel por fluorescencia. Las bandas están señaladas y numeradas ♦1, ♦2, ♦3 y ♦4 y se indican a qué correspondería cada una.

A continuación se realizó una cromatografía de exclusión molecular en columna Superdex 75 para evaluar si por este método era posible separar especies marcada de no marcada. Como se observa en la Figura 1.28 A, el perfil de corrida de la muestra marcada (abajo) evidenció la presencia de dos picos (P1 y P2), los cuales fueron aislados. El análisis por SDS-PAGE (Figura 1.28 B) reveló que en cada pico elúan principalmente dos especies con

distinta movilidad electroforética (♦1 y ♦2) y que tenían prácticamente idéntico perfil de bandas. Como ya fue discutido anteriormente, la existencia de dos picos eluyentes en SEC pero de patrón similar en SDS-PAGE fue atribuido a la presencia de especies monomérica y dimérica de AdE1A.WT (esta última mediada por la Cisteína 124 en condiciones oxidantes). El revelado por fluorescencia del mismo gel, demostró que sólo las bandas ♦1 y ♦3 correspondían a especies fluoresceinadas, en tanto que ♦2 y ♦4 no estaban marcadas.

I.3.B. Purificación e identificación de AdE1A.WT-FITC

Dado que no fue posible aislar la especie fluoresceinada por SEC, se procedió a correr la muestra por RP-HPLC utilizando una columna C8. Los picos y hombros obtenidos se aislaron para luego ser analizados por espectrometría de masa y lograr la identificación unívoca de la especie con marca FITC.

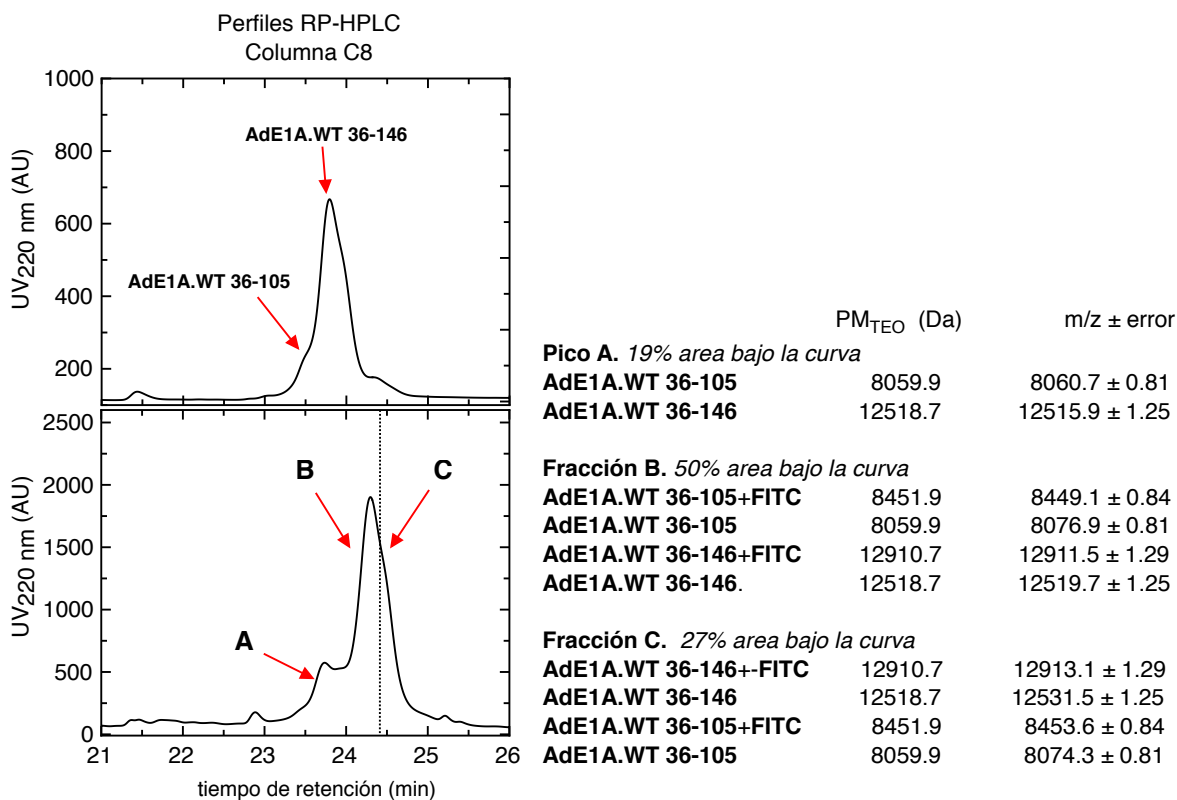


Figura 1.29: Análisis por espectrometría de masa de la muestra de AdE1A.WT marcada con FITC. Los cromatogramas RP-HPLC seguidos a 220nm de la muestra control (A) y la muestra marcada con FITC (B). En la tabla se encuentran las especies identificadas por pico en la espectrometría de masa MALDI-TOF (m/z) y su PM_{TEO} (Da). La línea punteada señala el tiempo de retención en el que se dejó de recolectar la fracción B y empezó a recolección de la fracción C.

En la Figura 1.29 A se muestra el cromatograma control de la corrida en el que se utilizó AdE1A.WT sin marca, la cual ya había sido analizada previamente en los estudios de tripsinización de la proteína. En el perfil de la corrida de la muestra marcada con FITC (Figura

1.29 B), se puede observar un pico de baja intensidad (A, abarcando un 19% del área de la curva) presenta el mismo tiempo de retención que el pico principal de la corrida control, en tanto que el pico B/C (77% del área bajo la curva) se encuentra desplazado hacia un tiempo de retención mayor, sugerente de la adición del isotiocianato de fluoresceína a la fracción principal de la muestra. El pico A y las fracciones B y C del pico principal fueron recolectados por separado.

Como en el HPLC-RP las fracciones eluyen en medio ácido (Acetonitrilo–TFA) y el FITC en estas condiciones tiene un pico de absorción menos intenso y corrido a longitudes de onda menores que 495nm, en primer lugar los picos recolectados fueron liofilizados y resuspendidos en tampón 20 mM fosfato de sodio pH 7.0 antes de ser analizados por espectrometría de masa.

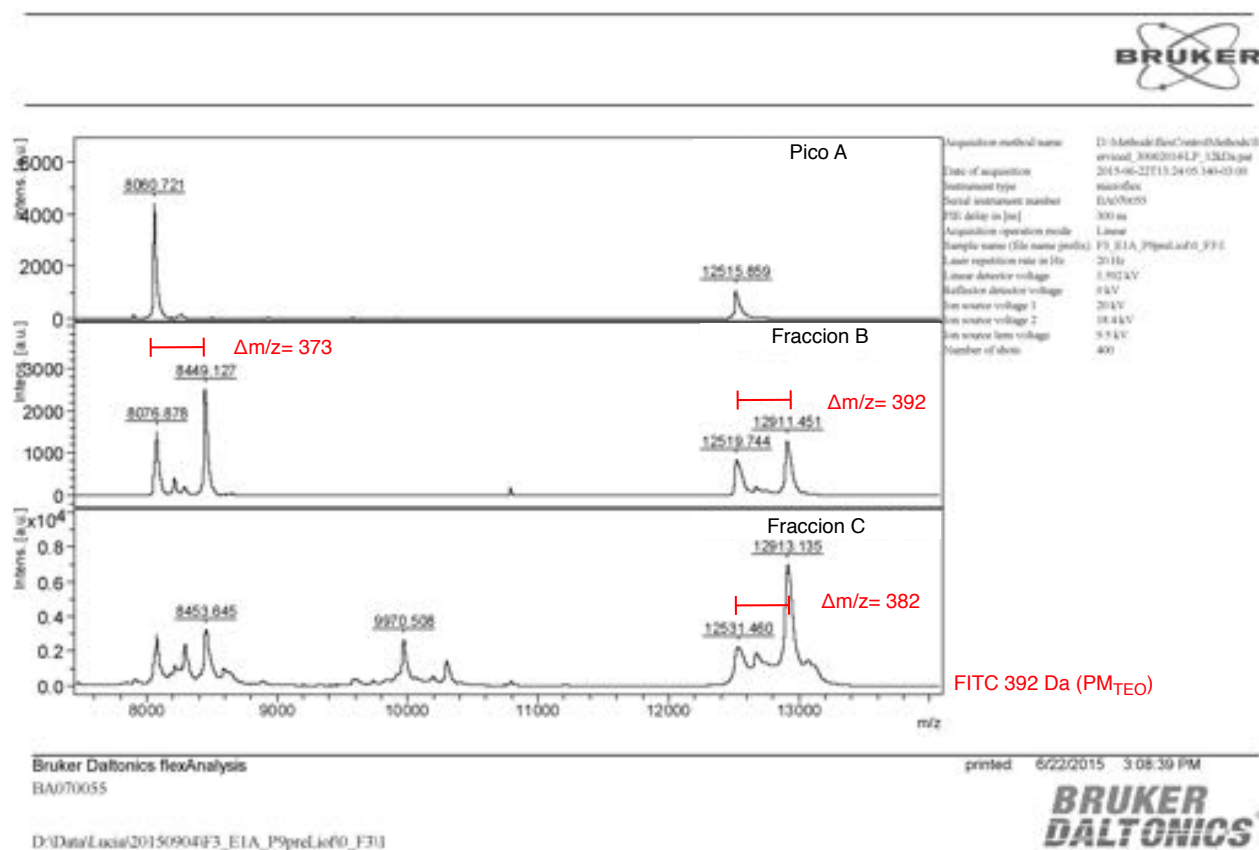


Figura 1.30: Espectros MALDI-TOF de muestras de AdE1A.WT marcada con FITC y purificada por RP-HPLC. Los picos identificados en el análisis de masa en cada una de las tres fracciones obtenidas de la purificación.

El análisis de masa MALDI-TOF (Figura 1.30), reveló que el pico A contenía especies de 12515.9 ± 1.25 m/z y 8060.7 ± 1.25 m/z que son correspondientes a AdE1A.WT (36-146) y AdE1A.WT (36-105) respectivamente y además que en este pico no estaban presentes especies con marca fluorescente. Para la fracción B en tanto, se identificaron especies de

8076.9 \pm 0.81 m/z probablemente tratándose de AdE1A.WT (36-105) (PM_{TEO} 8059.9 Da) asociada a iones y 12519.7 \pm 1.25 m/z asumiendo que se trataba de AdE1A.WT (36-136) (PM_{TEO} 12518.7 Da) por la gran concordancia de los valores. Además, fueron identificadas especies de 8449.1 \pm 0.85 m/z y 12911.5 \pm 1.29 m/z. Considerando que FITC tiene un PM_{TEO} 392 Da, estas últimas especies mencionadas se consideraron AdE1A.WT (36-105)+FITC (PM_{TEO} 8451.9 Da) y AdE1A.WT (36-136)+FITC (PM_{TEO} 12910.7 Da), respectivamente. En el caso de la fracción C, los resultados fueron similares a los de la fracción B, aunque se encontraron mayores discordancias entre los PM de las especies identificadas 8074.3 \pm 0.81 m/z y 12531.5 \pm 1.25 m/z, respecto de las que podrían corresponder: AdE1A.WT (36-105) y AdE1A.WT (36-136), porque probablemente se hayan tratado de especies asociadas a iones. No obstante, la especie que de interés: AdE1A.WT (36-136)+FITC, fue identificada con éxito.

Como se puede ver en la Figura 1.30 en los espectros de MALDI-TOF de estas muestras, teniendo en cuenta las relaciones de intensidades de las señales, la fracción C contenía en mayor proporción la especie de interés AdE1A.WT (36-146)+FITC. Por otro lado, la cuantificación de los *razones F/P* de las fracciones B y C post liofilizadas y resuspendidas, arrojó valores de 0.67 y 0.8 respectivamente.

Como las especies identificadas sólo tenían asociada una sola molécula de FITC, y habiendo controlado pH de la reacción de marcación -en la que se produce la derivatización de α -aminos- se asumió que las especies estaban marcadas en el grupo amino-terminal. El conjunto de estos resultados permitió además elegir la fracción C (post RP-HPLC, liofilización y resuspensión en tampón a pH 7.0) como la muestra de AdE1A.WT (36-136)+FITC con la que se ensayaron los experimentos de espectroscopía de fluorescencia.

CONCLUSIONES I

Como primer objetivo de este capítulo, se planteó la generación de los reactivos con los que se trabajó en lo sucesivo. Esto incluyó la generación de AdE1A.WT y dos variantes mutantes: AdE1A- Δ E (carente del motivo E2F: 43-LHELY-47) y AdE1A- Δ L (al que le falta el motivo LxCxE: 122-LTCHE-126) reemplazándose por parches de tetra-alaninas. A continuación se desarrollaron protocolos de alto rendimiento para la expresión recombinante y purificación de estas proteínas en medio rico 2TY (con un rendimiento máximo de \approx 9 mg/l de cultivo) y en medio mínimo M9 (con un rendimiento de hasta \approx 4mg/l de cultivo). En ambos casos se obtuvieron proteínas cuya pureza fue superior al 95% y cuya identidad fue confirmada luego por espectrometría de masa. Estas proteínas obtenidas pudieron ser concentradas incluso hasta concentraciones cercanas a 900 μ M (11.2 mg/ml) sin que se observase agregación alguna.

Una vez que se dispuso de cantidades adecuadas de proteína AdE1A se inició su estudio centrado en la caracterización de su estabilidad y propiedades estructurales-conformacionales en solución. AdE1A posee una migración electroforética anómala respecto de lo esperado para su peso molecular teórico, probablemente por el sesgo en su composición de secuencia (enriquecida en residuos ácidos) y una conformación sumamente extendida [132]. Además, esta proteína evidencia alta sensibilidad a la proteólisis, característica asociada frecuentemente a regiones desordenadas (Figura 1.30). La proteólisis limitada además demostró que las dos únicas argininas de la secuencia de AdE1A (R97, R105) no serían igualmente sensibles a ataque proteolítico. A nivel técnico, la identificación de estos sitios proteolíticos, significó adicionar inhibidores de proteasas en todos los pasos de purificación, así como extremar los cuidados en el tratamiento de estas proteínas a fin de mantener su integridad para ensayos ulteriores.

Los experimentos de SEC-SLS-DLS demostraron que AdE1A es una proteína intrínsecamente desordenada que en solución se encuentra en una población conformacionalmente homogénea y monomérica y que presenta un comportamiento hidrodinámico altamente extendido, lo que la clasifica como una proteína “*random coil-like*”. La naturaleza de su composición aminoacídica es determinante en este comportamiento. AdE1A.WT particularmente posee una alta fracción de residuos ácidos (26 entre Asp y Glu) de un total de 114, y consecuentemente posee un pI bajo= 3.78. Probablemente el efecto repulsivo de las fuerzas electrostáticas negativas mantengan su conformación expandida. La neutralización progresiva de esas cargas negativas con la disminución de pH por ejemplo, podría conllevar a estructuras cada vez más compactas incluso a una final agregación, como se ha visto que ocurre a pH 4 y 3. Por otro lado, el alto contenido de prolina de AdE1A.WT (18 en

total, 11 de las cuáles se encuentran en una *región rica en prolinas*), también podrían colaborar a su comportamiento hidrodinámico extendido. Por ejemplo, los residuos prolina podrían experimentar isomerización *cis-trans* pudiendo inhibir conformaciones mas compactas [117], además de tener la tendencia natural a estructurar conformación PPII, la cual se trata de hélices extendidas.

Los experimentos de CD han permitido caracterizar a AdE1A como una proteína desordenada que carece de cantidades significativas de estructuras secundarias canónicas y no canónicas, pero que sin embargo es altamente dinámica y tiene capacidad de estabilizar en forma transitoria estructura α -hélice y PPII (por efectos de solvente y temperatura), lo que revela su gran plasticidad conformacional (Figura 1.30). Estos elementos de estructura secundaria transitorios podrían estabilizarse en la interacción con distintos blancos celulares. Justamente una predicción de tendencias α -hélice -a nivel de residuos- hecha sobre la secuencia de AdE1A.WT revela mayores probabilidades de estructurar dicha conformación sobre los motivos de interacción E2F, TRAM-CBP y LxCxE. Por otro lado, el escaneo térmico seguido por CD, demostró también la ausencia en AdE1A.WT de estructuras consolidadas con plegamientos cooperativos en su estado nativo, así como su alta estabilidad térmica, lo que es característico de IDP con alto grado de desorden (“*random coil-like*”) (Figura 1.30).

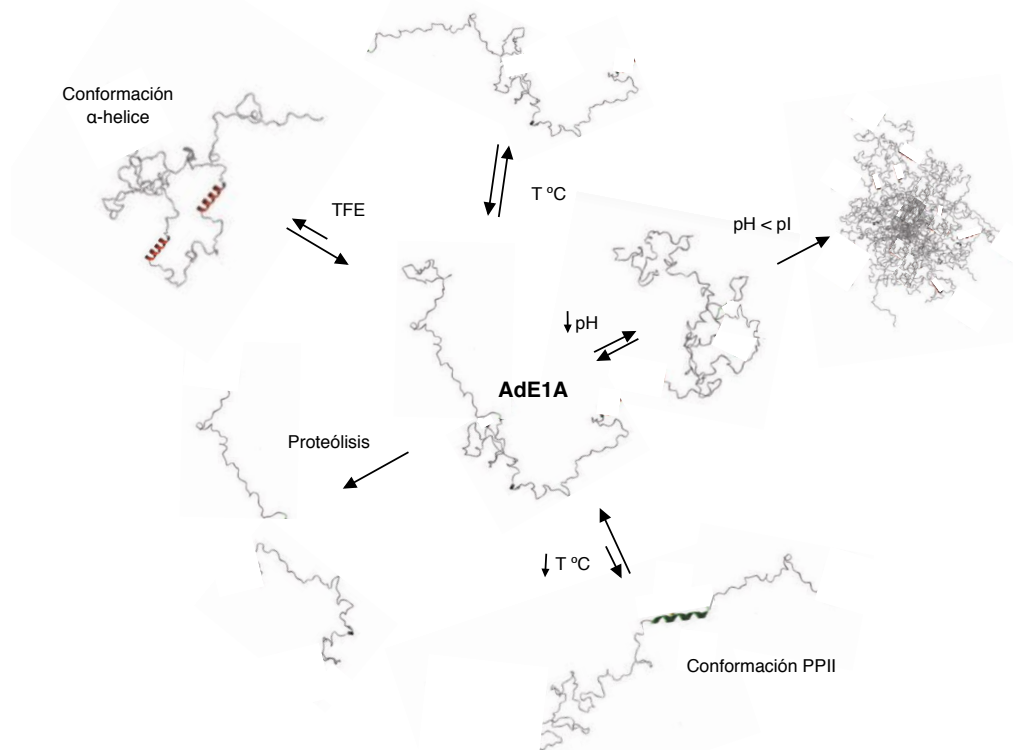


Figura 1.30: AdE1A es una IDP desordenada y altamente plástica. AdE1A es una IDP tipo “*random coil-like*” con un radio hidrodinámico altamente extendido, probablemente por su secuencia rica en prolinas y residuos cargados. No presenta estructura secundaria típicas en su estado nativo, pero tiene tendencias conformacionales a poblar α -hélices y PPII. Es una proteína con alta estabilidad térmica y no presenta estructuras con plegamientos cooperativos. AdE1A es sensible a la proteólisis.

CAPITULO II

Análisis del mecanismo de interacción entre AdE1A y su blanco RbAB

Hasta el momento la caracterización de la interacción entre AdE1A y RbAB permanece incompleta, tanto que *aún no resulta claro el modelo de interacción global de estas proteínas* ni hay estudios que *permitan comprender cómo la oncoproteína viral es capaz de desplazar el factor celular E2F de su complejo endógeno con RbAB*. Para otras proteínas virales, como en E7 de HPV, se ha demostrado que un motivo de interacción de alta afinidad es clave para unir e inactivar a RbAB [133], pero no sería el caso de AdE1A. Para esta última se ha planteado que para unir e inactivar a RbAB, ambos motivos E2F y LxCxE requieren interactuar en simultáneo con la proteína globular [76, 84, 108], pero aun resta saber cómo lo harían. Surgen así las siguientes preguntas *¿cómo logra AdE1A unir efectivamente a RbAB para tornarla inactiva, desplazando al factor de transcripción celular E2F2 y subvertir así el ciclo celular a su favor? ¿Cuál es el rol de los motivos de interacción E2F y LxCxE y la región “linker” en esta interacción?*

Ya presentadas las propiedades de AdE1A en forma aislada, en este capítulo se buscará *desarrollar un modelo cuantitativo* de la interacción en estudio y del desplazamiento efectivo de E2F2 celular. Esto se hará particularmente trabajando sobre 3 hipótesis planteadas en la siguiente figura:

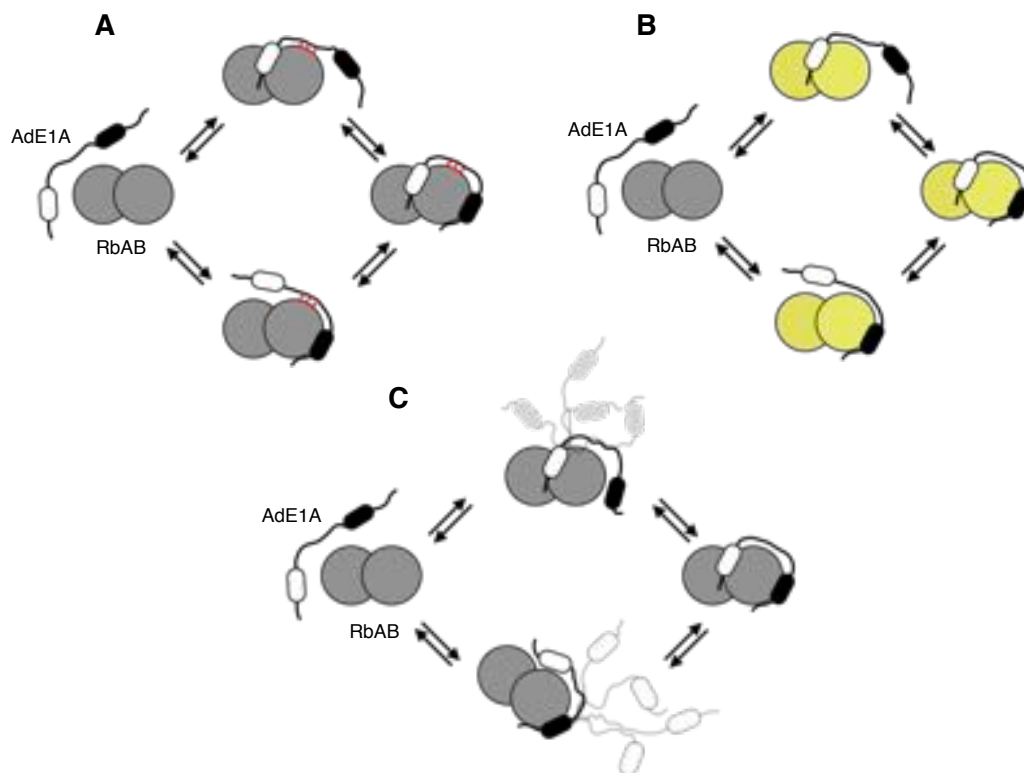


Figura 2.1. Modelos de interacción de alta afinidad entre AdE1A y RbAB a ser testeados.

En el Modelo A, se considera que uno de los dos motivos de interacción en AdE1A se une a RbAB a través de su sitio complementario, seguido por una contribución de la región “*linker*” que estabiliza el complejo por contactos directos, para que luego se produzca la unión del segundo motivo a RbAB (Figura 2.1 A). En el Modelo B, la interacción del primer motivo en AdE1A a su sitio complementario en RbAB, modula alostéricamente la afinidad del segundo sitio favoreciendo su interacción, a partir de cambios conformacionales inducidos en RbAB, mientras que el “*linker*” cumple un rol pasivo (Figura 2.1 B). En la tercera, Modelo C, una vez unido el primer motivo de AdE1A a RbAB, el “*linker*” produce un aumento efectivo de la concentración del segundo motivo en la región cercana a su sitio de interacción, favoreciendo su interacción (Figura 2.1 C).

La caracterización buscada a partir de estos estudios cuali-cuantitativos de la interacción de AdE1A con RbAB en este capítulo, profundizará el conocimiento de las interacciones virus-hospedador y los mecanismos subyacentes a la subversión de la maquinaria celular.

OBJETIVOS ESPECIFICOS:

- I. Presentar una revisión de datos experimentales de interacciones mediadas por los motivos E2F y LxCxE a miembros de la familia “*pocket*”, evaluando el rol de residuos en la modulación de la afinidad y especificidad de dichas interacciones.
- II. Analizar la interacción de la proteína AdE1A y fragmentos, con RbAB. Específicamente:
 - a. Hallar su estequiometría de interacción y caracterizar propiedades hidrodinámicas y conformacionales del complejo [RbAB:AdE1A], a partir de su análisis biofísico -dicroísmo circular y SEC-SLS-.
 - b. Estudiar las contribuciones energéticas asociadas a las interacción entre AdE1A y RbAB, a nivel de fragmentos, a partir de la determinación de sus parámetros termodinámicos, por espectroscopía de fluorescencia y por calorimetría de titulación isotérmica.
 - c. Encontrar correlatos estructurales a partir de las descripciones termodinámicas de la interacción en estudio
 - d. Evaluar el efecto de la fosforilación de la región ácida del CR2 en AdE1A en la interacción con RbAB.
 - e. Evaluar la capacidad de los distintos fragmentos para desplazar E2F2 de su interacción con RbAB y evaluar el rol de la región *linker* en este desplazamiento.

OBJETIVO II.1.

Revisión de datos experimentales de interacciones a miembros de la familia de pRb mediadas por los motivos lineales E2F y LxCxE.

Como fue mencionado en la introducción, las proteínas pRb, p107 y p130 son miembros de una familia multifuncional de proteínas, las proteínas “*pocket*”, las cuales son reguladores centrales del ciclo celular eucariota y están involucradas en el desarrollo y diferenciación celular, estabilidad genómica, regulación de la apoptosis, respuesta al daño a ADN y se encuentran inactivadas –particularmente pRb- en un amplio rango de cánceres en humanos de etiología genética y viral [97]. Estas múltiples funciones están mediadas por interacciones proteína-proteína entre dominios globulares o a través de motivos lineales de interacción (SLiMs) presentes en sus proteínas celulares y virales interactuantes. Los SLiMs, que son secuencias cortas de aminoácidos ubicados en regiones desordenadas de una proteína [37] median no solo interacciones sino también procesos de señalización, modificaciones post-traduccionales y degradación [37, 134]. Los SLiMs que median las interacciones de múltiples blancos interactuantes con el dominio central AB de las proteínas “*pocket*”, son el *motivo LxCxE*, que se une al *surco LxCxE* en el subdominio B [135] y el *motivo E2F*, que se une en la región interdominio A-B, el sitio complementario a E2F o *sitio E2F* [136].

Con el fin de describir el aporte de los motivos de E2F y LxCxE en las interacciones con los miembros de la familia “*pocket*”, pero a nivel de los residuos que componen, evaluando sus roles determinantes o moduladores de afinidad y especificidad, se realizó una revisión de evidencias experimentales disponibles en bibliografía. El análisis presentado a continuación resultó de un trabajo en coautoría publicado en Palopoli *et al* PEDS, 2018.

En esta revisión de literatura, se buscaron todas las menciones a ambos motivos E2F y LxCxE y se analizó la secuencia del entorno en el que se encontraban dichos motivos y las evidencias experimentales de su interacción con las proteínas “*pocket*”. De una primer revisión se encontró por ejemplo, que alrededor de 30 de mas de 100 interactores celulares y virales de pRb poseen el motivo LxCxE [37, 102, 137]. Luego esto fue refinado de modo que fueron descartados aquellos casos en los que estos motivos estaban embebidos en secuencias predichas (o conocidas) ordenadas o estructuradas comprometiendo su accesibilidad y funcionalidad [37] y/o no tenían o eran insuficientes las evidencias bioquímicas directas de interacción (Inmunoprecipitación, Co-Immunoprecipitación, ensayos de interacción directa como ITC, etc.). Se elaboró así una lista que comprende 19 proteínas de mamíferos y plantas y 14 de virus de mamíferos y plantas que contienen el motivo LxCxE capaz de unir a algún miembro

“pocket”. Por otro lado, el motivo E2F se ha identificado por ahora solamente en el dominio de transactivación de los miembros de la familia de transcripción E2F 1 a 5 (E2F-TD) y en la región CR1 de la proteína AdE1A de Adenovirus. Todas estas instancias están presentadas en la Tabla 2.1 y disponibles en la base de datos ELM (<http://elm.eu.org/>).

Tabla 2.1: Motivos LxCxE y E2F de unión a pRb validados experimentalmente

PROTEINA	código	ORGANISMO	código PDB	Referencia
Motivos LxCxE celulares				
Demetilasa 5A lisina-especifica	KDM5A_HUMAN	<i>Homo sapiens</i>		[138]
proteína 4A con dominio interactivo rico en AT	ARI4A_HUMAN	<i>Homo sapiens</i>		[138]; [139]
Histona Deacetilasa 1	HDAC1_HUMAN	<i>Homo sapiens</i>		[140]; [141]
Histona Deacetilasa 2	HDAC2_HUMAN	<i>Homo sapiens</i>		[142]; [144]
homólogo de proteína del cinetocoro NDC80	NDC80_HUMAN	<i>Homo sapiens</i>		[143]; [144]
proteína 1 con HMG-box	HBP1_RAT	<i>Rattus norvegicus</i>		[145]; [146]
inhibidor de diferenciación 1 interactuante c/EP300	EID1_HUMAN	<i>Homo sapiens</i>		[147]; [148]
Proteína 2 con dominio PR y dedos de zinc (RIZ)	PRDM2_HUMAN	<i>Homo sapiens</i>		[149]; [150]
subunidad 26 regulatoria de proteína fosfatasa 1	PPR26_HUMAN	<i>Homo sapiens</i>		[151]
Activador de transcripción SNF2L2 (Brm)	SMCA2_HUMAN	<i>Homo sapiens</i>		[152]; [153]
Activador de transcripción BRG1	SMCA4_HUMAN	<i>Homo sapiens</i>		[152]
Brg1	Q63928_9MURI	<i>Mus musculus</i>		[154]
Ciclina D1 específica de G1/S	CCND1_HUMAN	<i>Homo sapiens</i>		[155]; [156]
Ciclina D2 específica de G1/S	CCND2_HUMAN	<i>Homo sapiens</i>		[157]; [158]
Ciclina D3 específica de G1/S	CCND3_HUMAN	<i>Homo sapiens</i>		[157]; [158]
Ciclina D1-1	CCD11_ARATH	<i>A. Thaliana</i>		[159]
Ciclina D2-1	CCD21_ARATH	<i>A. Thaliana</i>		[159]
Ciclina D3-1	CCD31_ARATH	<i>A. Thaliana</i>		[159]
Factor de transcripción eritroide humano	GATA1_HUMAN	<i>Homo sapiens</i>		[160]
Motivos LxCxE virales				
Antígeno T grande	LT_SV40	<i>Simian Virus 40</i>	1GH6 ¹	[161]; [162]
Antígeno T grande	B8ZX42_9POLY	<i>Merkel cell polyomavirus</i>		[163]; [164]
E7	VE7_HPV16	<i>Human papillomavirus 16</i>	1GUX ¹ ; 4YOZ ²	[165]; [135]
E1A	E1A_ADE05	<i>Human adenovirus 5</i>		[166]; [108]
UL97 Serin-Treonin Kinasa	GCVK_HCMVA	<i>Human cytomegalovirus</i>		[167]
Wsv069	Q77J89_WSSVS	<i>Shrimp white spot syndrome</i>		[168]
Wsv056	Q77J94_WSSVS	<i>Shrimp white spot syndrome</i>		[168]
MC007 *	Q98178_MCV1	<i>Molluscum contagiosum</i>		[169]
RepA	REPA_BEYDV	<i>Bean yellow dwarf virus</i>		[170]
RepA	REPA_WDVS	<i>Wheat dwarf virus</i>		[171]
RepA	REPA_MSVS	<i>Maize streak virus</i>		[172]
Clink	CLINK_FBNY1	<i>Faba bean necrotic virus</i>		[173]
UNK	Q9WKM8_BBTV	<i>Banana bunchy top virus</i>		[174]
Motivos LxSxE celulares				
LIN-52**	LIN52_HUMAN	<i>Homo sapiens</i>	4YOS ²	[175]
Factor de transcripción eritroide murino	GATA1_MURINE	<i>Mus musculus</i>		[160]
Motivos E2F celulares				
E2F1	E2F1_HUMAN	<i>Homo sapiens</i>	1O9K ¹	[176]; [177]
E2F2	E2F2_HUMAN	<i>Homo sapiens</i>	1N4M ¹	[136]
E2F3	E2F3_HUMAN	<i>Homo sapiens</i>		[98]
E2F4	E2F4_HUMAN	<i>Homo sapiens</i>		[98]
E2F5	E2F5_HUMAN	<i>Homo sapiens</i>		[136]
bucle interdominio AB de Rb	RB_HUMAN	<i>Homo sapiens</i>	4ELL ¹	[178]; [179]
Motivos E2F virales				
AdE1A	E1A_ADE05	<i>Human adenovirus 5</i>	2R7G ¹	[84]; [108]

Todos los motivos reportados están apoyados en múltiples evidencias bioquímicas (IP-Western, mutagénesis, interacción in vitro) y son conocidos o se predicen estar localizados en regiones desestructuradas.

* Esta proteína tiene 2 motivos LxCxE

** Esta proteína se une a p107, no a pRb

¹ código PDB para la interacción con Human pRb humana.

² código PDB para la interacción con Human p107 humana.

Residuos centrales en el motivo LxCxE como determinantes de afinidad

De todas las instancias mencionadas en la Tabla 2.1, se conocen muy pocas afinidades de interacción -entre el motivo y la proteína “*pocket*”- que estén cuantificadas. Esto incluye a 3 instancias celulares y 3 virales del motivo LxCxE y 5 instancias celulares y sólo 1 viral del motivo E2F. En la Tabla 2.2 están compendiados estos datos.

Tabla 2.2. Afinidades reportadas para los motivos LxCxE y E2F unidos a proteínas de la familia de pRb.

Proteína	Unido	Secuencia	Técnica	K_D (M)	Referencia
Motivos LxCxE virales					
HPV16 E7 (16–31)	pRb	QPETTDLYCYEQLNDS	FD	$(2.10 \pm 0.03) \cdot 10^{-9}$	[109]
HPV16 E7 (16–40)	pRb	QPETTDLYCYEQLNDSSEEEDEIDG	FD	$(3.10 \pm 0.05) \cdot 10^{-9}$	[109]
HPV16 E7 (16-40) pS31pS32	pRb	QPETTDLYCYEQLNDSSEEEDEIDG	FD	$(0.36 \pm 0.08) \cdot 10^{-9}$	[109]
HPV16 E7 (21–29)	pRb	DLICYEQLN	EF	$(4.7 \pm 1.7) \cdot 10^{-9}$	[100]
HPV16 E7 (16–31)	pRb	QPETTDLYCYEQLNDS	EF	$(5.1 \pm 1.3) \cdot 10^{-9}$	[100]
HPV16 E7 (16–40)	pRb	QPETTDLYCYEQLNDSSEEEDEIDG	EF	$(6.5 \pm 1.0) \cdot 10^{-9}$	[100]
HPV16 E7 (16-40) pS31pS32	pRb	QPETTDLYCYEQLNDSSEEEDEIDG	EF	$(1.8 \pm 0.4) \cdot 10^{-9}$	[100]
HPV16 E7 (21–29)	pRb	DLICYEQLN	ITC	$(0.11 \pm 0.03) \cdot 10^{-6}$	[135]
HPV16 E7 (21–29)	pRb	DLICYEQLN	ITC	$(0.19 \pm 0.07) \cdot 10^{-6}$	[141]
HPV16 E7 (21-29) L22I	pRb	DIICYEQLN	ITC	$(0.32 \pm 0.07) \cdot 10^{-6}$	[141]
HPV16 E7 (21-29) L28F	pRb	DLICYEQFN	ITC	$(0.14 \pm 0.03) \cdot 10^{-6}$	[141]
HPV16 E7 (21-29) D21R	pRb	RLICYEQLN	ITC	$(1.05 \pm 0.04) \cdot 10^{-6}$	[141]
HPV16 E7 (21-29) D21R/Y25E	pRb	RLYCEEQLN	ITC	$(4.79 \pm 0.20) \cdot 10^{-6}$	[141]
SV40 LT (102-110)	pRb	NLFCSEEMD	ITC	$(0.44 \pm 0.06) \cdot 10^{-6}$	[141]
SV40 LT (7-117)	pRb	...NLFCSEEMPSSDDEAT	MT	$(1.95 \pm 0.37) \cdot 10^{-6}$	[163]
MCPyV LT (1-244)	pRb	...EDLFCDEESLSSPEPPSSSEPEEP PSSRSSPRQP	MT	$(5.13 \pm 3.81) \cdot 10^{-8}$	[163]
E1A-CR2 (106-139) C124S/E137C	pRb	ALGPVSMPNLVPEVI DLTSHEAGFPSSDDEDEEG	sF	$(1.1 \pm 0.3) \cdot 10^{-7}$	[84]
Motivos LxCxE celulares					
HDAC (409-428)	pRb	SSDKRIACEEEFSDSEEEGE	ITC	$(10.0 \pm 3.0) \cdot 10^{-6}$	[141]
HDAC1 (413-421)	pRb	RIACEEFS	ITC	$(20.0 \pm 4.0) \cdot 10^{-6}$	[141]
HDAC1 (413-421) F420L	pRb	RIACEEELS	ITC	Sin unión	[141]
HDAC1 (413-421) R413D	pRb	DIACEEFS	ITC	$(10.0 \pm 5.0) \cdot 10^{-6}$	[141]
HDAC1 (413-421) R413D/E417Y	pRb	DIACYEFS	ITC	$(3.27 \pm 0.07) \cdot 10^{-6}$	[141]
RIZ AR (297-341)	pRb	...EIRCDEKPEDLLEEPKTTSETLED CSECTPAM	ITC	$(6.40 \pm 0.70) \cdot 10^{-7}$	[150]
RIZ AR (297-341)	pRb	...EIRCDEKPEDLLEEPKTTSETLED CSECTPAM	EF	$1.10 \cdot 10^{-7}$	[150]
RIZ (309-319)	pRb	EIRCDEKPEDL	EF	$3.60 \cdot 10^{-7}$	[150]
LIN52 (12-34)	pRb	TDLEASLLSFEKLDRA SPDLWPE	ITC	Sin unión	[175]
LIN52 (12-34) pS28	pRb	TDLEASLLSFEKLDRA SPDLWPE	ITC	Sin unión	[175]
LIN52 (12-34)	p107	TDLEASLLSFEKLDRA SPDLWPE	ITC	$(5.9 \pm 0.9) \cdot 10^{-6}$	[175]
LIN52 (12-34) pS28	p107	TDLEASLLSFEKLDRA SPDLWPE	ITC	$(1.4 \pm 0.9) \cdot 10^{-6}$	[175]
LIN52 (12-34) pS28 L18A	p107	TDLEASALSFEKLDRA SPDLWPE	ITC	$(14.9 \pm 0.1) \cdot 10^{-6}$	[175]
LIN52 (12-34) pS28 E22A	p107	TDLEASLLSFAKLDRA SPDLWPE	ITC	$(5.5 \pm 0.3) \cdot 10^{-6}$	[175]
LIN52 (12-34) pS28 L24A	p107	TDLEASLLSFEKADRA SPDLWPE	ITC	$(7.4 \pm 0.9) \cdot 10^{-6}$	[175]
LIN52 (12-34) pS28	p130	TDLEASLLSFEKLDRA SPDLWPE	ITC	$(1.0 \pm 0.1) \cdot 10^{-6}$	[175]
Motivos E2F virales					
E1A-CR1-linker (37-121)	pRb	...HFEPPTLHELYDL...	ITC	$(0.9 \pm 0.3) \cdot 10^{-6}$	[84]
E1A-CR1 (27-105) G88C	pRb	...HFEPPTLHELYDL...	sF	$(3.5 \pm 0.5) \cdot 10^{-7}$	[180]
E1A-N-CR1(1-105) C6S/G88C	pRb	...HFEPPTLHELYDL...	sF	$(4.5 \pm 0.7) \cdot 10^{-7}$	[180]

Técnicas: FD: Flujo detenido. EF: Espectroscopia de fluorescencia. ITC: Calorimetría de titulación isotérmica. MT: Temoforesis a microescala. sF: FRET de molécula individual (“*single molecule*”).

Continuación Tabla 2.2.

Motivos E2F celulares					
E2F-2 (410-427)	pRb	<u>DDY</u> LWGLEAGEG <u>ISDLFD</u>	ITC	$(0.19 \pm 0.04) \cdot 10^{-6}$	[136]
E2F-2 (410-427) E420A	pRb	<u>DDY</u> LWGLEAGAG <u>ISDLFD</u>	ITC	$(0.32 \pm 0.04) \cdot 10^{-6}$	[136]
E2F-2 (410-427) G421S	pRb	<u>DDY</u> LWGLEAGES <u>ISDLFD</u>	ITC	$(10.2 \pm 1.4) \cdot 10^{-6}$	[136]
E2F-5 (323-346)	pRb	<u>DDY</u> NFNLDDEG <u>VCDLFD</u>	ITC	$(0.69 \pm 0.1) \cdot 10^{-6}$	[136]
E2F-1 (409-426)	pRb	<u>LDY</u> HFGLEEGEG <u>IRDLFD</u>	ITC	$(0.34 \pm 0.02) \cdot 10^{-6}$	[103]
E2F-1 (380-437)	pRb	... <u>LDY</u> HFGLEEGEG <u>IRDLFD</u> ...	ITC	$(0.16 \pm 0.01) \cdot 10^{-6}$	[103]
E2F-1 (409-426)	pRb	<u>LDY</u> HFGLEEGEG <u>IRDLFD</u>	ITC	$(0.07 \pm 0.01) \cdot 10^{-6}$	[98]
E2F-1 (409-426)	p107	<u>LDY</u> HFGLEEGEG <u>IRDLFD</u>	ITC	$(1.06 \pm 0.04) \cdot 10^{-6}$	[98]
E2F-2 (410-427)	pRb	<u>DDY</u> LWGLEAGEG <u>ISDLFD</u>	ITC	$(0.04 \pm 0.01) \cdot 10^{-6}$	[98]
E2F-2 (410-427) G415N	pRb	<u>DDY</u> LWNLEAGEG <u>ISDLFD</u>	ITC	$(0.38 \pm 0.02) \cdot 10^{-6}$	[98]
E2F-2 (410-427)	p107	<u>DDY</u> LWGLEAGEG <u>ISDLFD</u>	ITC	$(3.1 \pm 0.3) \cdot 10^{-6}$	[98]
E2F-2 (410-427) G415N	p107	<u>DDY</u> LWNLEAGEG <u>ISDLFD</u>	ITC	$(8.3 \pm 0.5) \cdot 10^{-6}$	[98]
E2F-3 (432-449)	pRb	<u>EDY</u> LLSLGEEEG <u>ISDLFD</u>	ITC	$(1.2 \pm 0.1) \cdot 10^{-6}$	[98]
E2F-3 (432-449)	p107	<u>EDY</u> LLSLGEEEG <u>ISDLFD</u>	ITC	$(7 \pm 1) \cdot 10^{-6}$	[98]
E2F-4 (390-407)	pRb	<u>HDY</u> IYNLDESE <u>GVCDLFD</u>	ITC	$(0.5 \pm 0.2) \cdot 10^{-6}$	[98]
E2F-4 (390-407) N395G	pRb	<u>HDY</u> IYGLDESE <u>GVCDLFD</u>	ITC	$(0.10 \pm 0.02) \cdot 10^{-6}$	[98]
E2F-4 (390-407)	p107	<u>HDY</u> IYNLDESE <u>GVCDLFD</u>	ITC	$(0.15 \pm 0.01) \cdot 10^{-6}$	[98]
E2F-4 (390-407) N395G	p107	<u>HDY</u> IYGLDESE <u>GVCDLFD</u>	ITC	$(0.15 \pm 0.01) \cdot 10^{-6}$	[98]
Motivo E2F y LxCxE en la misma proteína viral					
E1A-CR1-CR2 (27-139) S88C/C124S	pRb	...HFEPPTLHELYDL...ALGPVSMPNLV PEVIDLTSHEAGFPPSDEDEE	sF	$< 25 \cdot 10^{-9}$	[180]
E1A-N-CR1-CR2 (1-139) C6S/G88C/C124S	pRb	...HFEPPTLHELYDL...ALGPVSMPNLV PEVIDLTSHEAGFPPSDEDEE	sF	$< 25 \cdot 10^{-9}$	[180]

Técnicas: D: Flujo detenido. EF: Espectroscopia de fluorescencia. ITC: Calorimetría de titulación isotérmica. MT: Temoforesis a microescala. sF: FRET de molécula individual ("single molecule").

Para el motivo LxCxE, los primeros análisis cuantitativos de afinidad fueron realizados por Jones y colaboradores [181] sobre proteínas virales, reportando afinidad nanomolar media para HPV16 E7 (21-29) unido a pRb y una afinidad un tanto menor para la interacción del motivo en AdE1A con pRb. Estudios posteriores -realizados con técnicas más sensibles- reportaron afinidades mayores para el mismo fragmento de HPV16 E7 (21-29) ($K_D \approx 5$ nM) [100, 109, 135, 141] y para una proteína relacionada, el antígeno T grande, de Virus de simio 40 (SV40 LT) y de Poliomavirus de Células de Merkel (MCPyV LT), aunque con afinidades menores a la de HPV16 E7 [163]. En el caso de las instancias celulares de este motivo, la Histona Deacetilasa 1 (HDAC1) que presenta un motivo *IACEE*, mostró tener interacción mas débil y transiente con pRb que los blancos virales [141], y para el motivo *IRCDE* de la proteína RIZ se reportó una afinidad nanomolar alto a pRb [150].

El motivo LxCxE de HPV16 E7 (Figura 2.2 A) se encuentra cristalizado con el dominio AB de pRb (RbAB) (PDBID: 1GUX) (Figura 2.3 A). En el análisis de esta estructura se observa que los residuos L22 y C24 se unen fuertemente a un bolsillo hidrofóbico en el surco complementario a LxCxE, mientras que E26 establece puentes de hidrógeno con las cadenas laterales F721 y K722 en pRb. A su vez, puentes de hidrógeno entre la cadena principal del motivo LxCxE y los residuos K713 y N757 de pRb mantienen al péptido en posición fija (Figura 2.2 A). Por otro lado, ensayos de afinidad demostraron un descenso de la afinidad del motivo

en unas 5 a 10 veces tras mutar por Alaninas a L23 [175, 181] (Tabla 2.2). No obstante se ha reportado que cambios conservativos en el residuo hidrofóbico L22 por I -y también L28 por F- en HPV16 E7, no alteran su afinidad por pRb, por lo que otros residuos del mismo carácter podrían estar permitidos en estas posiciones [141]. Estos últimos datos sumados a los datos estructurales previos demuestran que los residuos centrales L, C, E, que constituyen el núcleo del motivo (Figura 2.2 A), son determinantes en la interacción con pRb.

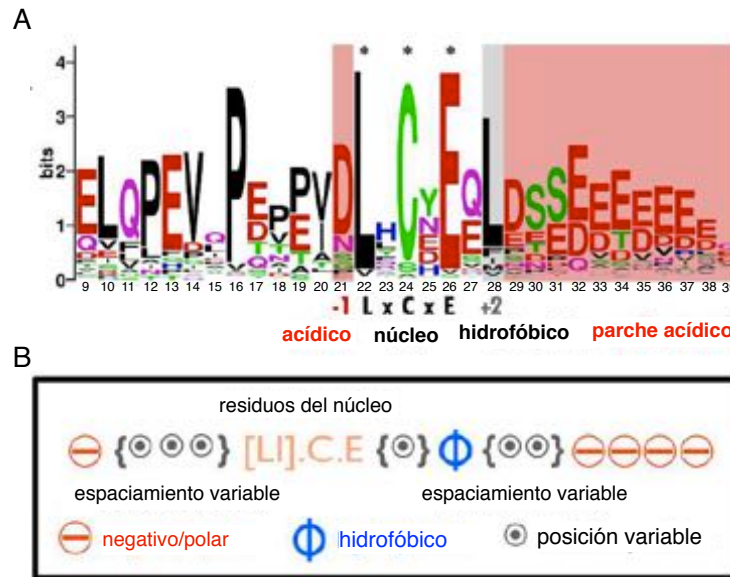


Figura 2.2. Residuos centrales y flanqueantes del motivo LxCxE A) Secuencia Logo de E7 de HPV mostrando un alto grado de conservación de los residuos del núcleo, así como las regiones flanqueantes que incluyen el residuo ácido en posición (-1) del núcleo del motivo, un cuarto residuo hidrofóbico en la posición (+2) del núcleo y el parche ácido y las Serinas fosforilables (verdes) hacia el C-terminal. B) Representación esquemática del núcleo del motivo LxCxE y sus regiones flanqueantes moduladoras. Mucho de los elementos flanqueantes pueden estar a distancias variables respecto del núcleo del motivo y el parche ácido puede aportar cargas negativas en un grado variable (Adaptado de [182])

También se vio que algunos contactos intramoleculares en el motivo LxCxE del péptido de HPV16 E7 serían importantes en la interacción con pRb, como los que establece Y25 con Y23. Si bien ninguno contacta directamente a pRb, su mutación produce una disminución de la afinidad del motivo con esta proteína globular. Esto sugiere que algunas conformaciones preformadas presentes en el péptido libre podrían modular dicha interacción [141].

Asimismo, parecería que un cuarto residuo “extra” de carácter hidrofóbico conservado en muchos motivos LxCxE -incluyendo a HPV16 E7, SV40 LT, HDAC, LIN52 y GATA-1- y ubicado en la posición +2 (Figura 2.2 A) también jugaría un rol importante en la interacción con pRb. Este residuo hidrofóbico se une a un amplio surco formado por los residuos V725, F739, I752 e I753 de pRb [135, 183] (Figura 2.3 A) y tiene un patrón de interacción similar en el complejo p107:LIN52 (LxSxE) [175](Figura 2.3 B). Se ha visto que la mutación de este residuo hidrofóbico (L24) en LIN52, es tan desestabilizante como las mutaciones de sus residuos

centrales L18, C20 y E22. Por otro lado, el amplio surco complementario al motivo en las proteínas “*pocket*”, permitiría variantes en la posición de este cuarto residuo hidrofóbico “extra”, como en el caso de AdE1A en donde está ubicado en la posición +3 [184] o en Ciclina D1, ubicado en la posición +4 y Ciclina D2, ubicado en la posición +3. No obstante, que tan lejos o cerca esté este cuarto residuo hidrofóbico “extra” respecto del núcleo del motivo, tendría efectos en la afinidad. Se ha visto que Ciclina D1 une con baja afinidad a pRb y es incapaz de unir a p130 [156], en tanto que Ciclina D2 une mas fuertemente a pRb [175]. Estas observaciones indican que las variaciones en posición de este cuarto residuo “extra”, podrían contribuir a explicar cambios en afinidades pero por ahora se requieren mayores estudios al respecto.

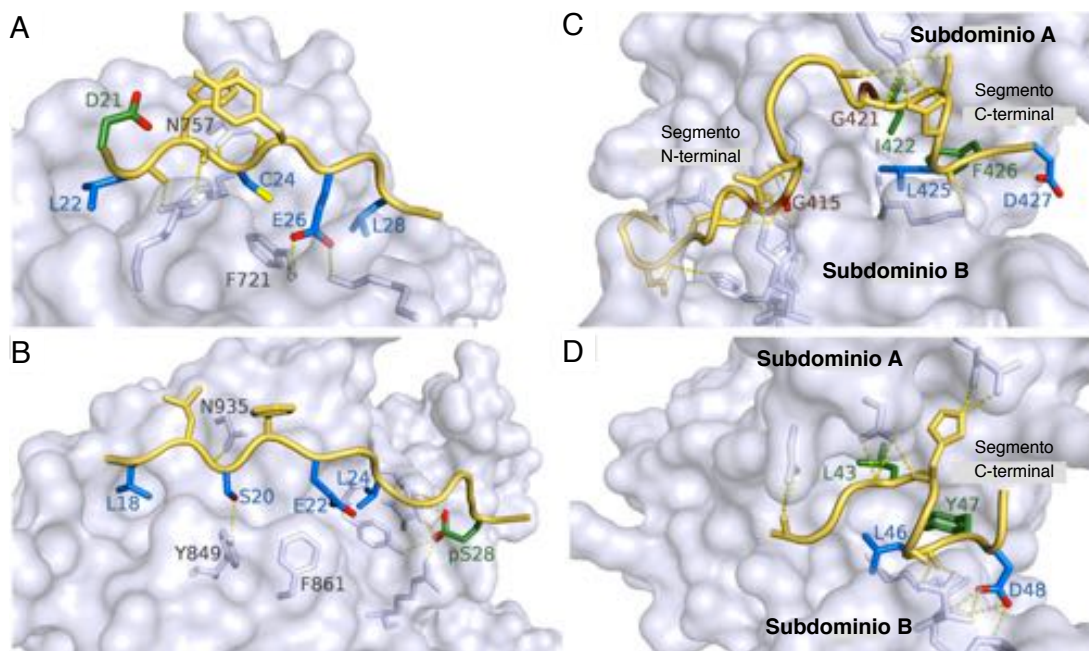


Figura 2.3 Estructuras representativas de las interacciones entre las proteínas “*pocket*” y los motivos LxCxE y E2F. A) Estructura del motivo LxCxE de HPV16 E7 unido al surco complementario en el subdominio B de pRb (PDB ID: 1GUX). Los residuos centrales L22, C24, E26 y L28 están azul. D21 en verde y las cadenas laterales de Y23 e Y25 en amarillo. B) Estructura del motivo LxSxE de LIN52 fosforilado en S28 unido al surco complementario en p107 (PDB ID: 4YOS). Los residuos centrales L18, S20, E22 y L24 están azul. pS28 (serina fosforilada) en verde y las cadenas laterales de L19 e Y21 en amarillo. C) Estructura del motivo E2F de E2F2 unido al surco entre los subdominios A y B de pRb (PDB ID: 1N4M). Los residuos centrales I422, L425, F426 y D427 están verde y azul. G415 y G421 del segmento N-terminal en rojo. G21 permite un pliegue que orienta adecuadamente al segmento N-terminal. En este segmento, contactos intermoleculares adicionales están coloreados en amarillo (cadenas laterales de residuos e interacciones puentes de hidrógeno). D) Estructura del motivo E2F de AdE1A del Adenovirus unido al surco entre los subdominios A y B de pRb (PDB ID: 2RG7). Los residuos centrales L43, L46, Y47 y D48 están verde y azul, en la hélice corta que forma el motivo. Contactos intermoleculares adicionales están coloreados en amarillo (cadenas laterales de residuos e interacciones puentes de hidrógeno). (Adaptado de [182]).

Residuos flanqueantes en el motivo LxCxE como moduladores de afinidad

Trabajos realizados por Singh *et al.* mostraron que la mutación D21R en HPV16 E7 produce un descenso de unas 8 veces en la afinidad con pRb, en tanto la mutación reversa R413D en HDAC 1 incrementa la afinidad por pRb unas 2 veces [141] (Tabla 2.2). Esto sugeriría que la presencia de un residuo ácido en posición (-1) (Figura 2.2 A) favorece la interacción con pRb, probablemente porque aporta una complementariedad de cargas a los residuos básicos K713 y K765 en el surco de pRb, posicionando correctamente al motivo (Figura 2.3 A).

A su vez, las proteínas HPV16 E7, HDAC 1 y SV40 LT poseen una región o parche cargado negativamente contiguo al motivo LxCxE (Figura 2.2) que según lo sugieren estudios cinéticos de interacción, establecería interacciones electrostáticas complementarias a la superficie positiva que rodea al surco LxCxE en pRb [109]. Los datos experimentales muestran que HPV16 E7, HDAC 1 y RIZ muestran mayor afinidad por pRb cuando estos parches acídicos están presentes (Tabla 2.2).

También se ha visto que la proteína celular HDAC1 y las proteínas virales HPV16 E7, SV40 LT, MPCyV LT y AdE1A, presentan dentro de estos parches acídicos contiguos al motivo LxCxE, una serie de residuos serinas que son fosforilables. La fosforilación *in vitro* de S31 y S32 en HPV16 E7 incrementa la afinidad por pRb unas 4 veces respecto del péptido no fosforilado [100, 109]. También se ha observado un incremento en la eficiencia de interacción con pRb cuando AdE1A está fosforilada en S132 [185]. En el caso de MPCyV LT se ha reportado que la mutación S220A disminuye la afinidad por pRb en experimentos de cultivo celular [186] y que la fosforilación del motivo no canónico LxSxE en LIN52 incrementa la afinidad por p107 [175].

En suma, todos estos datos evidencian la existencia de múltiples elementos flanqueantes a los centrales del motivo LxCxE: un residuo ácido en la posición (-1), un parche ácido hacia el C-terminal del motivo y serinas fosforilables dentro de este parche. Estos residuos flanqueantes si bien no son indispensables para la interacción, presentan una clara función moduladora de la afinidad del motivo por las proteínas “*pocket*”.

Determinantes de especificidad para interacciones LxCxE-proteínas “pocket”

Un estudio de Guiley y colaboradores, ha mostrado que el motivo no canónico LxSxE en LIN52 media la interacción con p107 y p130 como parte del complejo DREAM que regula diferenciación, proliferación celular y supresión tumoral. Este motivo difiere de LxCxE en que posee una serina en lugar de la cisteína, los cuales son estructuralmente similares, aunque el

hidroxilo (-OH) de Ser es mas polar que el grupo tiol (-SH) de Cis, lo que modifica el patrón de interacciones puente de hidrógeno que pueden establecer. Esto conduce a una disrupción de la interacción puente de hidrógeno entre el residuo E22 de LIN52 y F861 en p107 provocando una afinidad sub-óptima. No obstante, se ha visto que la fosforilación de S28 en LIN52 colaboraría incrementando su afinidad por p107 y p130, ya el grupo fosfato estabiliza puentes de hidrógeno con residuos adyacentes al surco LxCxE en la superficie de estas dos proteínas “*pocket*” [175] (Figura 2.3 B). Como estos residuos adyacentes al surco LxCxE difieren en p107 y p130 respecto de los de pRb, el motivo LxSxE de LIN52 fosforilado en S28 no es capaz de unirse a pRb.

Estos resultados sugieren que los determinantes de especificidad en la interacción del motivo LxCxE por pRb o p107/p130, están dados por la sustitución Cisteína por Serina y por los residuos adyacentes a los bolsillos hidrofóbicos del sitio complementario al motivo (surco LxCxE) en las proteínas “*pocket*”. La relación entre estos factores provee de un mecanismo por el cual la fosforilación puede regular la interacción entre el motivo LxSxE y este grupo de proteínas.

Especificidad de la interacción del motivo E2F y regiones flanqueantes

El *motivo E2F* es una hélice anfipática corta que se une a la interfaz entre los subdominios A y B de las proteínas “*pocket*” (sitio E2F) y que está presente en los miembros de la familia de factores de transcripción E2F (1 a 5) y en la proteína AdE1A de Adenovirus. Hasta el momento se han hecho 3 reportes de afinidad a pRb del motivo mínimo E2F más segmentos hacia el N- y/o C-terminal de la secuencia. Liu y Marmorstein, informaron una K_D micromolar para la secuencia AdE1A (37-121) [84], en tanto que Ferreon y colaboradores lo hicieron para secuencias AdE1A (1-105) y AdE1A (27-105) con algunas mutaciones puntuales en la secuencia pero alejados del motivo E2F reportando K_D del orden sub-micromolar. Los reportes para las interacciones a pRb de los miembros de la familia de factores de transcripción E2F muestran mayores afinidades, del orden nano a sub-micromolar [98, 103, 136](Tabla 2.2). En los factores de transcripción E2F, la superficie de interacción a pRb es bipartita e involucra un segmento N-terminal extendido y la hélice hacia el C-terminal del motivo E2F.

El motivo E2F está formado por tres residuos hidrofóbicos, en AdE1A por L43, L46 e Y47 y en los factores de transcripción E2F por I422, L425 y F426, los cuáles forman una hélice anfipática, en la que una de sus caras se une al surco conservado en pRb a través de interacciones de van der Waals con los residuos L476, F482 y K530 de pRb [136]. (Figura 2.3 C, D).

En AdE1A, los residuos L43 e Y47 son críticos para desplazar a E2F de su interacción con pRb tras producirse la infección [98]. D48, un residuo ácido conservado, establece puentes de hidrógenos con residuos en pRb, constituyéndose en un determinante adicional de la interacción. H44, que es específico de AdE1A en el motivo E2F, también establece puentes de hidrógeno con las cadenas carbonada y laterales de E464 y N472 de pRb (Figura 2.3 D) causando re-arreglos significativos en N472 [98]. El patrón de interacciones de AdE1A y los factores de transcripción E2F con pRb parece extenderse mas allá de los residuos centrales del motivo.

En los factores de transcripción E2F, el motivo E2F presenta un elemento adicional de interacción a pRb: un segmento extendido N-terminal “extra” al motivo mínimo en hélice, permitiendo establecer otros contactos hidrofóbicos y puentes de hidrógeno con el sitio complementario a E2F [103, 136]. En E2F2, el residuo G421 no interacciona con pRb, pero genera un pliegue en el péptido que orienta adecuadamente los segmentos N- y C-terminal del péptido en el surco de pRb (Figura 2.3 C). Se ha visto que la mutación G421S anula esta la flexibilidad dada por el residuo G, lo que provoca un descenso en la afinidad por pRb [136] (Tabla 2.2).

Las proteínas “*pocket*” también exhiben preferencias por los diferentes miembros de la familia de factores de transcripción E2F, así mientras pRb se une a E2F 1 y 2, p107 y p130 lo hacen exclusivamente con E2F 4 y 5. Estudios bioquímicos y estructurales muestran que esta selectividad estaría explicada por diferencias en los residuos que constituyen el sitio complementario a E2F en las proteínas “*pocket*” y que se unen a la hélice C-terminal del motivo E2F y además por los diferentes residuos que constituyen el segmento N-terminal del motivo entre los distintos miembros E2F, como G415 en E2F 2 y N395 en E2F 4 [98]. Otro factor determinante de especificidad estaría dado por la interacción del dominio C-terminal de pRb (CTD) con el dominio CM de E2F/DP. Esta interacción CM-CTD contribuye a la especificidad y están diferencialmente reguladas por patrones de fosforilación en el CTD [187]. Esto parece indicar que pRb evolucionó para unir diferencialmente y con gran afinidad a los distintos miembros de la familia E2F que son más potentes para activar el ciclo celular, mientras que p107 y p130 unen a E2F “inhibitorias”.

Por otro lado, el sitio E2F entre los dominios A y B en pRb también es blanco de interacción por parte del bucle interdominio del mismo pRb, como parte de un mecanismo de auto-inactivación mediada Ciclinas/CDK. Este bucle interdominio posee un serina (S608) [188] que tras fosforilarse, se estructura en una hélice y se une al sitio complementario a E2F, estableciendo puentes de hidrógeno que bloquean la interacción con el factor de transcripción

E2F [188]. Como resultado de ello, se produce al desplazamiento de este factor de transcripción y la activación del ciclo celular.

Este análisis de datos bioquímicos y estructurales muestra que en los motivos E2F y LxCxE -hallados en proteínas virales y del huésped- sus residuos centrales están bien definidos y son determinantes para la interacción con las proteínas “pocket”, mientras que los residuos dispuestos en las regiones flanqueantes –sean ellos sitios de fosforilación, residuos hidrofóbicos o regiones cargadas- son mas variables y pueden modular finamente la afinidad y especificidad de estas interacciones a las proteínas de la familia de pRb. En su conjunto, constituyen evidencia experimental de que siendo los SLiMs elementos clave en la interacción entre proteínas, la secuencia aminoacídica que hace a su contexto podría tener efecto crítico en sus propiedades, definiendo especificidades de interacción y funcionalidades.

OBJETIVO II.2.

Mecanismo de interacción entre AdE1A y RbAB

En esta sección, se buscó analizar y caracterizar biofísicamente los complejos [RbAB:AdE1A] (de la variante natural y mutantes), con el objetivo de describir su estequiometría de interacción y propiedades hidrodinámicas. También se buscó estudiar la posible estructuración de AdE1A tras producirse la unión con RbAB por difracción circular.

Para ello se trabajaron con las siguientes construcciones de AdE1A.WT y las variantes mutantes y de RbAB:

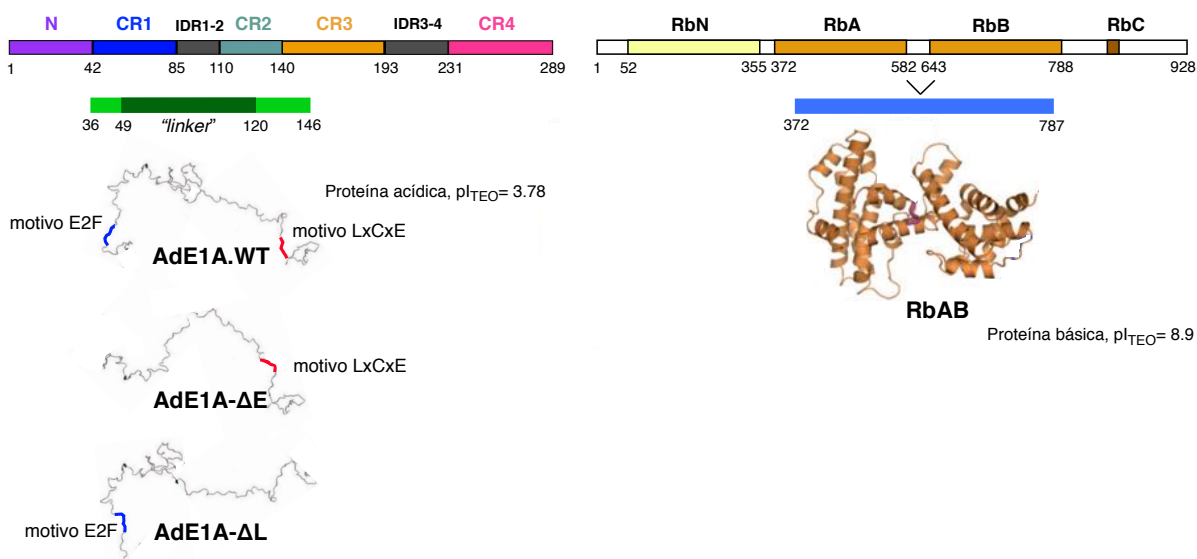


Figura 2.4. Fragmentos de las proteínas de trabajo utilizadas en esta sección. A la izquierda la secuencia de AdE1A y los límites de los fragmentos utilizados para la variante natural y las mutantes. A la derecha, la secuencia de RbAB y el fragmento de la proteína utilizado.

II.2.A. Estudio de las propiedades hidrodinámicas del complejo [RbAB:AdE1A] por SEC-SLS. Primer acercamiento al modelo de interacción.

Con la identificación de los dos sitios de contacto distintivos que posee AdE1A para interactuar con RbAB (motivos E2F y LxCxE) [108] y las evidencias estructurales de interacción entre el AdE1A-CR1 con pRb [84] y del complejo homólogo RbAB:HPV16 E7-CR2 [135] otros autores ya han propuesto que ambos sitios de interacción actúan sinérgicamente para desplazar a E2F de su unión con pRb. Teniendo en cuenta esto, formalmente dos modelos de complejos RbAB:AdE1A son igualmente válidos. En el primero de ellos, una molécula de AdE1A interactúa a través de sus motivos E2F y LxCxE con una misma molécula de RbAB (1:1), por lo cual la región “linker” se repliega de forma extendida sobre la proteína globular y

permite ambos contactos simultáneamente. En el segundo modelo, -sustentando por el hecho de que en los cristales del complejo RbAB con E1A-CR1 encontraron dos moléculas de RbAB por unidad asimétrica y porque en trabajos previos se ha sugerido que pRb puede formar oligómeros- en uno de los dos sitios de contacto de AdE1A interactúa con una molécula de RbAB, mientras que el segundo motivo lo hace con una segunda molécula de RbAB (2:2) (Figura 2.5) [84, 189].

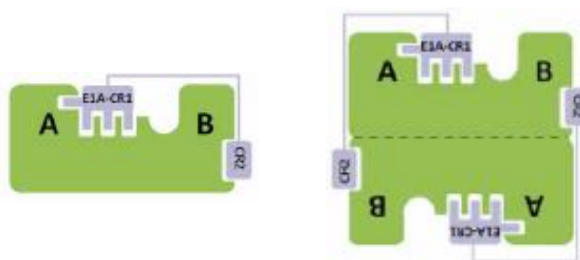


Figura 2.5. Modelos esquemáticos de la interacción entre RbAB y AdE1A.WT. A la izquierda, el modelo de interacción 1:1. A la derecha, el modelo formalmente posible 2:2. En ambos, pRb se representa en verde con sus sub-dominios indicados y AdE1A en gris, con sus sitios de contacto en CR1 y CR2. (Adaptado de [84]).

Una de las formas de distinguir entre estas dos posibilidades es mediante el estudio de sus propiedades hidrodinámicas por SEC-SLS. Aplicado a estos casos, el análisis de dispersión de luz láser estática acoplado a cromatografía de exclusión molecular, permite establecer el peso molecular del complejo independientemente de sus propiedades hidrodinámicas y a partir de sus propiedades de dispersión evaluar simultáneamente la estequiometría de la interacción, al permitir la determinación de la cantidad de componentes que lo forman. Este uso simultáneo de las dos técnicas es especialmente importante para el caso de IDPs, donde es difícil distinguir utilizando solo SEC entre un comportamiento hidrodinámico extendido y posible estados de oligomerización.

Para realizar esta caracterización, se analizaron por separado las proteínas individuales AdE1A y el complejo entre RbAB y AdE1A. Como esta serie de experimentos se realizó en simultáneo y a continuación a la caracterización de las propiedades hidrodinámicas y estados de oligomerización de AdE1A.WT, AdE1A-ΔE y AdE1A-ΔL libres en solución -presentadas en el Capítulo I- se retoman datos antes proporcionados.

En primer lugar se sembraron 100 μl de muestra AdE1A.WT a 270 μM, 100 μl de RbAB 35μM y posteriormente 100 μl de [RbAB:AdE1A.WT] formado por 60μM de cada componente, en una columna Superdex 75 analítica. La columna estaba acoplada a un sistema HPLC y a un refractómetro diferencial. Las señales de dispersión de luz a 90° (LS) y el índice de refracción (RI) a medida que las muestras eluían fueron debidamente recolectadas.

Como se puede observar en la Figura 2.4 y Tabla 2.3, este análisis reveló que AdE1A.WT eluye en un único pico, en estado monomérico y presenta un comportamiento hidrodinámico extendido (Figura 2.4 panel superior y ver Capítulo I). Por su parte, RbAB eluyó con un peso molecular aparente por SEC (PM_{apSEC}) de 45.9 kDa y presentó señal de índice de refracción correspondiente con una especie de peso molecular por SLS (PM_{SLS}) de 38.82 kDa. Considerando el error experimental, ambos valores son muy aproximados al peso molecular teórico (PM_{TEO}) de RbAB que es 42.1 kDa, indicando que esta proteína presenta un estado compacto ($PM_{apSEC}/PM_{TEO}= 1.09$) y monomérico ($PM_{SLS}/PM_{TEO}= 0.92$) en solución (Figura 2.6 panel medio y Tabla 2.3), lo que es concordante con datos previamente reportados para RbAB [133].

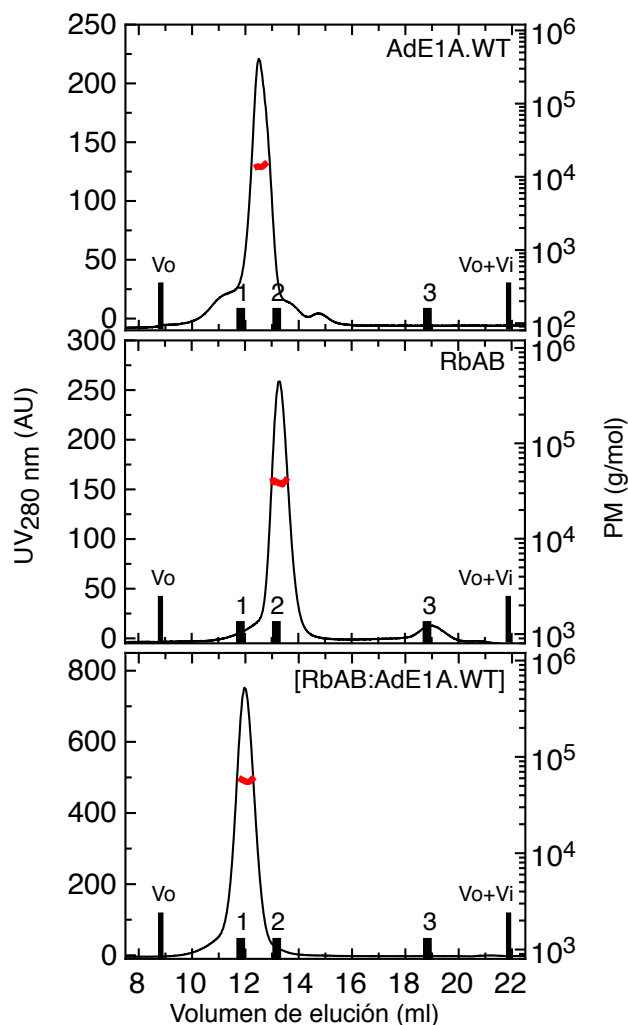


Figura 2.6. Análisis del complejo [RbAB:AdE1A.WT] por SEC-SLS. Perfiles de cromatogramas de exclusión molecular (Superdex 75) y señales de SLS para AdE1A.WT a 270 μ M (panel superior), RbAB a 35 μ M (panel medio) y el complejo estequiométrico [RbAB:AdE1A.WT] formado por 60 μ M de cada componente (panel inferior). Todas las medidas fueron hechas en tampón fosfato de sodio 20 mM, 200 mM NaCl, 1 mM DTT pH 7.0. Las barras negras corresponden a los volúmenes de exclusión de los marcadores moleculares utilizados BSA 66kDa (1), MBP 45 kDa (2) y Lisozima 14.3 kDa (3). Vo y Vo + Vi están indicados. Línea negra: perfil SEC. Línea roja: señal de SLS indicando el valor de PM (g/mol).

Tabla 2.3 Análisis de la estructura cuaternaria y propiedades hidrodinámicas de las proteínas de AdE1A libres y en complejo con pRb.

	Ve (ml)	PM _{TEO} (kDa)	PM _{SLS} (kDa)	PM _{apSEC} (kDa)	Estado de oligomerización PM _{SLS} /PM _{TEO}	Comportamiento hidrodinámico PM _{apSEC} /PM _{TEO}	R _H (nm) ¹
AdE1A.WT	12.51	12.5	13.93	54.4/ 59.68±0.03 ^a	1.11 (monómero)	4.35/ 4.77 ^a (extendido)	3.07/ 3.42±0.01 ^a
AdE1A-ΔE	12.66	12.1	12.96	52.7	1.07 (monómero)	4.36 (extendido)	3.03
AdE1A-ΔL	12.57	12.2	11.58	53.7	0.95 (monómero)	4.40 (extendido)	3.05
RbAB	13.30	42.1	38.82	45.9	0.92 (monómero)	1.09 (compacto)	2.91
[RbAB:AdE1A.WT]	11.96	54.6	56.3	61.2	1.03 (complejo 1:1)	1.12	3.25
[RbAB:AdE1A-ΔE]	12.51	54.2	53.9	65.3	0.99 (complejo 1:1)	1.20	3.32
[RbAB:AdE1A-ΔL]	11.96	54.3	56.2	69.1	1.03 (complejo 1:1)	1.27	3.40

^a Resultado de promediar tres medidas independientes de DLS

¹ el radio hidrodinámico experimental se calculó de la relación $\log R_H = -0.204 + 0.357 \log PM_{apSEC}$

Para el complejo [RbAB:AdE1A.WT], los resultados mostraron que el mismo eluyó en un único pico con un PM_{apSEC} 61.2 kDa y presentó señal de índice de refracción homogénea correspondiente a un PM_{SLS} 56.3 kDa. El PM_{SLS} hallado experimentalmente fue muy aproximado al PM_{TEO} esperado para el complejo formado por un monómero de AdE1A y un monómero de RbAB (54.6 kDa), lo que es indicativo que el complejo [RbAB:AdE1A.WT] sigue una estequiometría 1:1 (PM_{SLS}/PM_{TEO}= 1.03). Por otro lado, la razón PM_{apSEC}/PM_{TEO}= 1.12 evidencia la formación de un complejo que es semi-compacto (Figura 2.6 panel inferior y Tabla 2.3) con un radio hidrodinámico calculado según la formulación de Uversky [120] de R_H 3.25 nm. Como al estado libre RbAB es una proteína globular (y su R_H 2.93 nm) y AdE1A.WT está considerablemente extendida (y su R_H 3.12 nm), los resultados sobre el complejo sugieren que tras la interacción la proteína viral de naturaleza desordenada se repliega de forma extendida sobre pRb lo que permitiría el contacto simultáneo de ambos sitios de interacción.

Estos resultados están en acuerdo con el reportado por Clements y colaboradores quienes co-expresaron el complejo pRb_{his}:E1A-CR1-3, formado por pRb entre los aminoácidos 376 y 792 (dominio central AB) + marca de histidina (PM_{TEO} 49.9 kDa) y AdE1A entre los aminoácidos 36 y 189 poseyendo ambos sitios de interacción, E2F y LxCxE (PM_{TEO} 17.2 kDa). Para este complejo aislado de manera soluble, sus experimentos combinando las técnicas de cromatografía de exclusión molecular y sedimentación al equilibrio (SE) evidenciaron mayoritariamente una estequiometría 1:1. No obstante es de notar que sus datos experimentales mostraron un PM_{apSEC} 80.2 kDa para este complejo cuyo PM_{TEO} es 67.1 kDa,

mostrando una mayor discordancia entre el valor experimental y el teórico, lo que les sugirió cierto grado de oligomerización del complejo, probablemente porque ellos utilizaron una construcción de AdE1A con el dominio CR3 que podría presentar tendencias a oligomerizar como se ha visto para el CR3 de HPV-E7 [114].

Los experimentos de SEC-SLS también fueron realizados sobre las variantes mutantes AdE1A- Δ E y AdE1A- Δ L libres y en complejo con RbAB. En sus formas libres ambas mutantes presentaron un estado monomérico y un comportamiento hidrodinámico extendido en solución (Ver Capítulo I y Tabla 2.3).

El complejo [RbAB:AdE1A- Δ E] eluyó en un único pico con un PM_{apSEC} 69.1 kDa y presentó señal de índice de refracción homogénea correspondiente a un PM_{SLS} 56.2 kDa (Figura 2.7, Tabla 2.3). La razón $PM_{SLS}/PM_{TEO} = 1.03$ indicó que este complejo sigue una estequiometría 1:1 y $PM_{apSEC}/PM_{TEO} = 1.27$ reveló un grado de compactación del complejo menor que la de [RbAB:AdE1A.WT] y un R_H 3.32 nm, levemente mayor. En tanto el complejo [RbAB:AdE1A- Δ L] eluyó en un único pico con un PM_{apSEC} 65.3 kDa y presentó señal de índice de refracción homogénea correspondiente a un PM_{SLS} 53.9 kDa (Figura 2.7, Tabla 2.3). Este complejo también sigue una estequiometría 1:1 ($PM_{SLS}/PM_{TEO} = 0.99$) y presenta un grado de compactación ($PM_{apSEC}/PM_{TEO} = 1.20$) menor a la presentada por [RbAB:AdE1A.WT] y un R_H 3.40 nm, mayor pero no muy distante del de [RbAB:AdE1A.WT] .

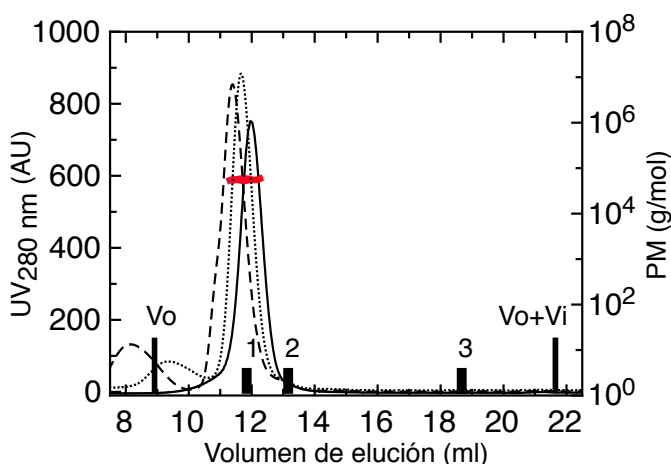


Figura 2.7. Análisis del complejo [RbAB:AdE1A.WT] y RbAB:AdE1A mutantes por SEC-SLS. Perfiles de cromatogramas de exclusión molecular (Superdex 75) y señales de SLS para los complejos [RbAB:AdE1A.WT] (línea continua), [RbAB:AdE1A- Δ E] (línea punteada) y [RbAB:AdE1A- Δ L] (línea de rayas) formados con 60 μ M de cada componente. Todas las medidas fueron hechas en tampón fosfato de sodio 20 mM, 200 mM NaCl, 1 mM DTT pH 7.0. Las barras negras corresponden a los volúmenes de exclusión de los marcadores moleculares utilizados BSA 66kDa (1), MBP 45 kDa (2) y Lisozima 14.3 kDa (3). Vo y Vo + Vi están indicados.

Para estas variantes mutantes sus complejos con RbAB eluyeron con un peso molecular aparente por SEC superior al de [RbAB:AdE1A.WT] y presentaron una estequiometría “n”=1, algo que era esperable, ya que en estos casos un solo extremo de AdE1A está unido a RbAB mientras que el otro extremo de la cadena está “libre” generando a un estado menos compacto.

En resumen, este análisis reveló que todos los complejos se forman con un monómero de AdE1A y un monómero de RbAB (estequiometría 1:1). La compactación observada por SEC-SLS para el complejo [RbAB:AdE1A.WT] es compatible con una estructura que asume que ambos extremos de AdE1A están unidos a RbAB, por lo que para la proteína viral decrece su radio hidrodinámico comparado con su estado libre. Para las mutantes se observa que, dado el hecho de que unen a RbAB sólo a través de un único sitio de contacto, el otro extremo permanece libre dotando a estos complejos de un radio hidrodinámico ligeramente superior al del complejo [RbAB:AdE1A.WT].

II.2.B. Caracterización espectroscópica por Dicroísmo Circular del complejo RbAB:AdE1A. Evaluando procesos de plegamiento-interacción en la formación del complejo.

Debido a su naturaleza altamente dinámica, una característica de las IDPs es que pueden experimentar transiciones de desorden-a-orden en diferentes condiciones. Así, muchas proteínas desordenadas al unirse a sus diferentes blancos de interacción fisiológicos pueden plegarse en diferentes estructuras de manera total o parcial -sobre sus elementos de reconocimiento- en procesos de plegamiento-acoplado-a-interacción [51]. En el Capítulo I se presentó la caracterización de AdE1A lograda por Dicroísmo Circular (CD) y otras técnicas biofísicas que evidenciaron su naturaleza desordenada en el estado nativo libre en solución y su capacidad para poblar estructuras canónicas (α -hélice) y no canónicas (PPII) en distintas condiciones (solvente TFE y temperatura). Además, existe evidencia de que AdE1A experimenta una transición de desorden-a-hélice entre los residuos 56 a 78 de la región CR1 tras interactuar con los dominios TAZ-2 [76] y NCBP de los co-activadores transcripcionales CBP/p300 [82], estando estos residuos comprendidos en la región “*linker*” del fragmento de AdE1A utilizado en el presente trabajo. Podía ser esperable entonces, que tras la interacción con RbAB, además del motivo E2F que ya se vio que estabiliza α -hélice en complejo con RbAB (Figura 2.3 D), regiones más extensas que comprendan el “*linker*” AdE1A experimenten un proceso de plegamiento-acoplado-a-interacción.

Para evaluar esta posibilidad se realizó un estudio por Dicroísmo Circular del complejo. La espectroscopía de CD es una técnica útil que permite cuantificar el contenido global de estructura secundaria de un péptido o una proteína y revelar en forma consistente a una reacción de plegamiento-acoplado-a-interacción, si los componentes que interaccionan forman un complejo con distinta cantidad de estructura secundaria relativa a la que contiene y aporta cada componente aislado [190], como bien se ejemplifica en la Figura 2.8. En el mismo se puede observar como el espectro de CD del complejo presenta mayor cantidad de señal de α -hélice, que el espectro suma de sus componentes libres, consecuente con la estructuración de la IDP.

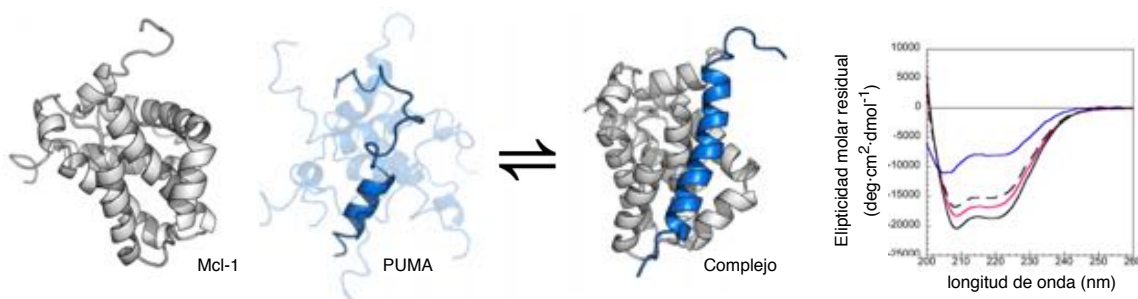


Figura 2.8. Ejemplo de caracterización de un complejo por Espectroscopía de CD. Sistema modelo de plegamiento-acoplado-a-interacción de la región BH3 de PUMA que forma una α -hélice continua al unirse a la proteína globular Mcl-1. Representación de Mcl-1 en gris (PDBID:1WSX), ensambles de estructuras construido con Chimera (UCSF) del péptido PUMA libre en azul, y el complejo formado con ambas proteínas unidas (PDBID: 2ROC) en gris y azul. El análisis por Dicroísmo Circular muestra como PUMA se une a Mcl-1 con un incremento en la helicidad del sistema. El complejo 1:1 (línea negra continua) evidencia mayor señal de α -hélice que el espectro del mismo sistema en el que los componentes no interactúan (línea negra de rayas) y que resulta de la sumatoria de los espectros de CD de PUMA libre (línea azul) y de Mcl-1 libre (línea roja). Todos estos espectros están en elipticidades molares. (Adaptado de [190])

Para evaluar este fenómeno en la formación del complejo [RbAB:AdE1A.WT], en primer lugar se llevó a cabo una medida del espectro de CD de las muestras de AdE1A.WT y de RbAB por separados a concentración 5 μ M en tampón fosfato de sodio pH 7.0 a 20 °C. La elipticidad cruda fue registrada y graficada para cada una de las dos especies trabajadas (Figura 2.9 B), y a partir de esta fue calculada la elipticidad molar residual media ($[\theta]_{MRW}$) para cada proteína individual. Sus espectros de CD típicos están representados en la Figura 2.7 A.

El espectro de CD de AdE1A.WT, que ya había sido individualmente ensayado, mostró su perfil de IDP con mínimo a 200 nm y escasa señal a 220 nm. Por otro lado, el espectro de CD de RbAB, extensamente caracterizado por Chemes y colaboradores [133], mostró un espectro típico de estructura α -hélice con dos mínimos de similar intensidad a 208 y 220 nm y un máximo ubicado a 190 nm consistente con su estructura cristalográfica [135, 191] que evidencia un alto contenido de este tipo de estructura secundaria canónica, ya que se trata

doble motivo de ciclinas, donde cada uno de ellos presenta un arreglo característico de 5 α -hélices.

Luego se midió el espectro de CD del complejo [RbAB:AdE1A.WT] resultado de la mezcla de cada componente a 5 μM en condiciones de medida: tampón fosfato de sodio pH 7.0 a 20 °C. El espectro del complejo en valores de elipticidad cruda fue adquirido y representado en la Figura 2.9 C (línea negra) y fue comparado con el espectro resultado de la suma aritmética de los espectros crudos de cada uno de los componentes del complejo obtenidos por separado (línea azul a rayas), que es el espectro predicho para un sistema en el que los componentes *no* interactúan. Como se puede observar, a diferencia del ejemplo mostrado en la Figura 2.8, en este caso AdE1A.WT se une a RbAB dando lugar a un complejo cuyo espectro de CD se solapa perfectamente con el espectro suma de los componentes, lo que demuestra que contienen idéntica cantidad de estructura secundaria a la predicha por la sumatoria de los espectros por separado.

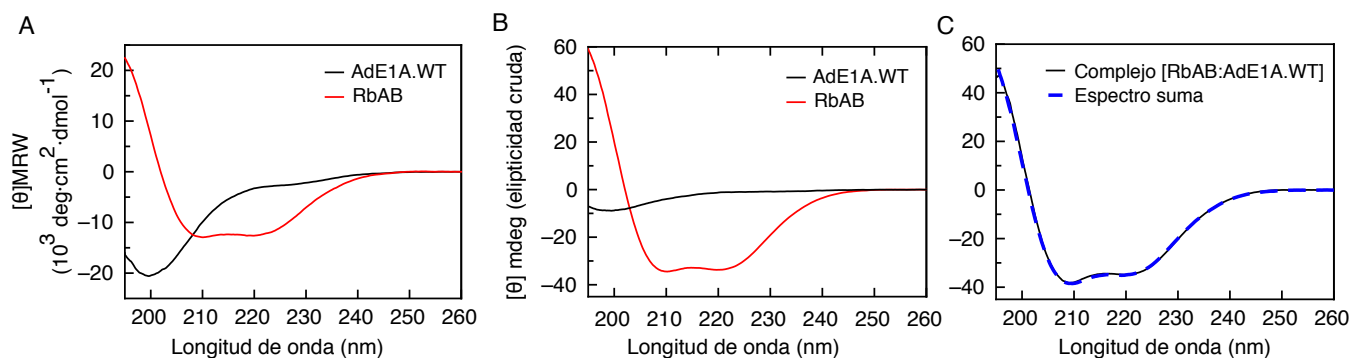


Figura 2.9. Caracterización del complejo [RbAB:AdE1A.WT] por Espectroscopía de CD. A) Espectros de CD individuales en elipticidad molar característicos de AdE1A.WT (línea negra) y RbAB (línea roja). B) Para este experimento: Espectros de CD individuales en elipticidad cruda de AdE1A.WT (línea negra) y RbAB (línea roja). C) Espectro de CD del complejo [RbAB:AdE1A.WT] a una razón molar 1:1 (5 μM de cada componente) (línea negra) comparado con el espectro de CD resultado de la suma aritmética de los espectros en elipticidad cruda de AdE1A.WT y RbAB (línea azul a rayas).

Si bien los datos estructurales muestran que el motivo E2F de AdE1A se encuentra formando una pequeña α -hélice de alrededor de 6 residuos (L43 a D48) al interactuar con el surco interdominio A-B de RbAB [84] (Figura 2.3 D), lo cierto es que su contribución a la señal total de CD sería muy baja en comparación con la aportada por el doble motivo de ciclina del dominio central de RbAB (más de 150 residuos en α -hélice). Además, sería un tanto difícil diferenciar esta contribución de aquella que se estima contiene inicialmente AdE1A.WT en el estado nativo libre -el cual fue de un 6.96% de residuos en α -hélice y que correspondería a alrededor de 7/8 residuos en dicha conformación sobre un total de 114- (Ver Capítulo I). Más allá de estos detalles este experimento demuestra que la formación del complejo con RbAB no

conduce a un plegamiento-acoplado-a-interacción significativo de AdE1A detectable por esta metodología, mas allá de la evidente estabilización de estructura α -hélice a nivel local en el motivo mínimo E2F [84].

En conclusión, estos experimentos demuestran que la proteína viral se une a RbAB para generar el complejo [RbAB:AdE1A.WT] en el cual mantendría una conformación altamente desordenada.

II.2.C. Análisis termodinámico de la interacción RbAB-AdE1A.

Las descripciones realizadas hasta este punto en este trabajo sobre la interacción entre AdE1A y RbAB implicaron el manejo de técnicas que permitieron la descripción cualitativa del sistema. Una caracterización más acabada requiere de la cuantificación precisa de los parámetros termodinámicos de esta interacción y la correlación de estos datos con las estructuras interactuantes involucradas [192]. Por otra parte, como se vio al inicio de este capítulo, este tipo de reportes para la interacción RbAB-AdE1A son parciales y escasos o se informan afinidades de variantes mutantes de la proteína viral [84, 193], además de que el trabajo en experimentos biofísicos con estas proteínas en forma controlada es difícil. Por lo tanto no hay un modelo cuantitativo de la interacción entre estas proteínas, que permita entender cómo se produce a nivel molecular una interacción de alta afinidad y de cuenta del desplazamiento efectivo de E2F celular por parte de AdE1A.

Para llegar a describir el modelo cuantitativo de esta interacción se fueron descartando las hipótesis planteadas al principio (Figura 2.1), con un estudio detallado de las contribuciones energéticas de las diferentes regiones funcionales AdE1A en la interacción con RbAB-AdE1A. Para ello se utilizaron los péptidos (obtenidos por síntesis química, ver Materiales y Métodos) y proteínas (obtenidos por expresión recombinante –Capítulo I-) presentados en la Figura 2.10.

Estos reactivos fueron utilizados en experimentos con dos metodologías diferentes: Espectroscopía de Fluorescencia (EF) para lo cual fueron marcados con la sonda fluorescente FITC a través de ensayos directos y de competencia y por Calorimetría de Titulación Isotérmica (ITC, por sus siglas en inglés).

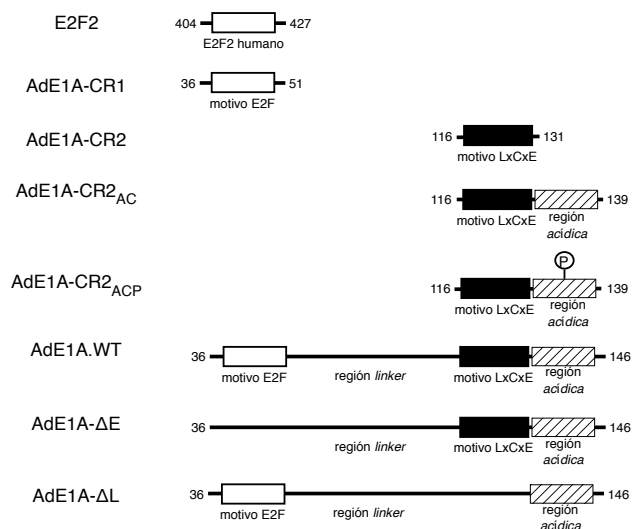


Figura 2.10: Esquemas de péptidos y proteínas utilizados en este trabajo para evaluar el aporte de las diferentes regiones de AdE1A en la interacción RbAB- AdE1A. Para todos los reactivos se señalan los límites de los fragmentos y las regiones en ellos comprendidos.

II.2.C.1. Determinación de estequiometrías y constantes de disociación por Espectroscopía de Fluorescencia

Titulaciones directas

Los experimentos de espectroscopía de fluorescencia (EF) se han convertido en procedimientos de rutina para estudiar interacciones proteína-ligando. Esta metodología sólo requiere que se produzca un cambio en el rendimiento cuántico que sea consecuente a la interacción, independientemente que se esté observando la fluorescencia del ligando, fluorescencia intrínseca de la proteínas o la fluorescencia de una sonda unida a un ligando que es sensible a la interacción [194]. Un parámetro asociado a la interacción entre dos moléculas donde una de ellas está marcada en forma fluorescente, es la variación de la anisotropía de fluorescencia (r) con el tamaño de la partícula. En este sentido por ejemplo, el valor que asume r de una sonda unida a un ligando al estado libre se modifica al producirse la unión con la proteína, porque se forma un complejo de mayor tamaño, comúnmente asociado a un mayor tiempo rotacional del fluoróforo, dando por resultado un aumento en el valor de anisotropía, r . Así, estos cambios observables son luego analizados por algún método que permite determinar constante de equilibrio de asociación (K_A) –o de disociación ($K_D = K_A^{-1}$)- y estequiometría (n) de reacción.

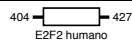
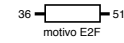
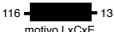

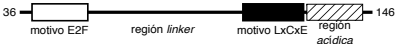


Los ensayos cuantitativos de interacción péptido-ligando basados en EF ya se encontraban puestos a punto en el laboratorio y mostraron gran reproducibilidad y precisión de

resultados para otros sistemas. Para continuar con este grado satisfactorio de resultados se trabajó con condiciones conocidas que aseguran la estabilidad de las muestras: tampón fosfato de sodio 20 mM, 200 mM NaCl, 2 mM DTT y 0.1% Tween-20, pH 7.0 y a 20 °C. Mientras el tampón fosfato de sodio 20 mM y 200 mM NaCl constituyó el medio probado en el que se caracterizaron propiedades y estabilidad de AdE1A y RbAB [133], el Tween-20 fue usado para prevenir la agregación de los complejos en las reacciones -permitiendo titulaciones sin pérdida de muestra- y el DTT fue agregado para asegurar condiciones reductoras -fundamental en el caso de los péptidos y proteínas conteniendo la cisteína 124 (del motivo LxCxE) altamente reactiva que media la dimerización de los mismos en condiciones oxidantes-.

Para estos experimentos se utilizaron los péptidos y proteínas derivatizadas con el fluoróforo FITC (Figura 2.10). En Materiales y Métodos se comenta en detalle estos ensayos y los ajustes realizados. Brevemente, para todas las interacciones, una cubeta conteniendo una cantidad fija de sonda fluorescente (péptido o proteína fluoresceínada) fue titulada mediante el agregado de una solución concentrada de RbAB. Para cada punto de la curva se esperó 2 minutos para que la reacción alcance el equilibrio y luego la señal emitida fue adquirida. Esta señal, r , fue aumentando con la formación de los complejos en comparación con las especies libres, lo que permitió seguir los cambios en la curva de asociación.

Las titulaciones fueron realizadas en dos rangos de concentraciones. A altas concentraciones proteicas (al menos 10 veces superior a la K_D estimada) se determinaron las estequiometrías de reacción. A concentraciones sub-estequiométricas (en el rango de las K_D) las medidas permitieron una estimación de la constante de disociación al equilibrio (K_D) para todos los complejos estudiados.

Tabla 2.4. Estequiometrías y Constantes de disociación de las interacciones entre péptidos y proteínas AdE1A y E2F2 humana con RbAB

	Pep o Prot	n	K_D (nM)	ΔG (cal/mol)
 E2F2 humano	E2F2	1	1.03 ± 0.03	-12057 ± 16
 motivo E2F	AdE1ACR1	1	119.1 ± 2.1	-9286 ± 10
 motivo LxCxE	AdE1ACR2	1	104.6 ± 2.9	-9362 ± 16
 motivo LxCxE región acidica	AdE1ACR2 _{Ac}	1	73.4 ± 2.2	-9577 ± 18
 motivo E2F región linker motivo LxCxE región acidica	AdE1A-WT	1	0.024 ± 0.004	-14244 ± 97
 región linker motivo LxCxE región acidica	AdE1A- ΔE	1	47.1 ± 1.7	-9828 ± 21
 motivo E2F región linker región acidica	AdE1A- ΔL	1	206.7 ± 3.1	-8966 ± 9

ΔG se calculó como $\Delta G = RT \cdot \ln(K_D)$, con $RT = 582$ cal/mol
n= estequiometría de la interacción.

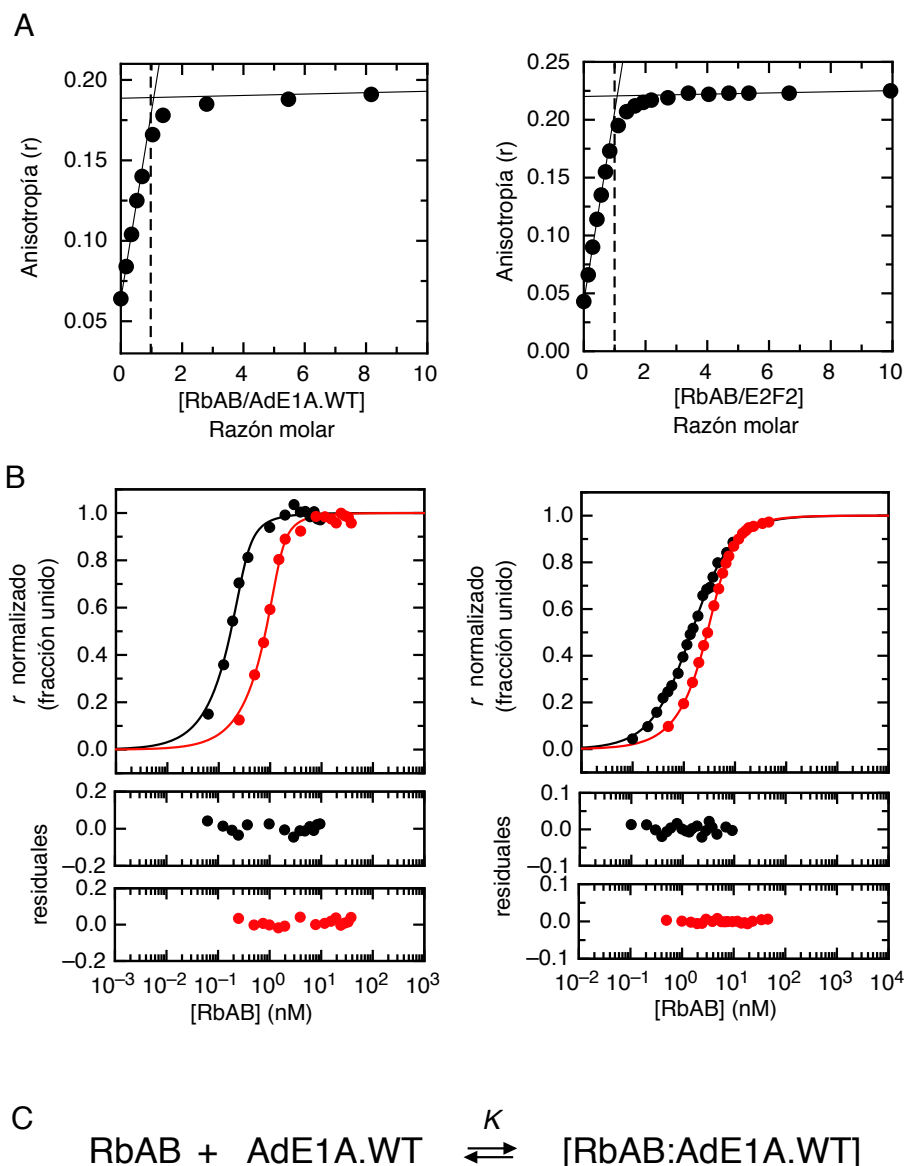


Figura 2.11: Medidas espectroscópicas para la determinación de estequiometrías y constantes de disociación y modelo de asociación bimolecular de 2 estados. A) Curva de asociación a concentraciones saturantes para determinar la estequiometría de unión realizada en tampón 20 mM fosfato de sodio, 200 nM NaCl, 1mM DTT, 0.1% Tween 20, pH 7.0 a 20 °C para las asociaciones RbAB- AdE1A.WT (trabajando con sonda AdE1A.WT-FITC a 2 nM, izquierda) y RbAB-E2F2 (trabajando con sonda E2F2-FITC a 30 nM, derecha). La estequiometría fue determinada por la extrapolación de los ajuste lineales (líneas rectas) de las señales iniciales y finales de anisotropía. B) Curvas de titulación al equilibrio ajustado a un modelo de interacción 1:1. Intensidades de anisotropía de fluorescencia normalizada a dos concentraciones sub-estequiométricas distintas (círculos negros y rojos) ajustadas a curvas de unión 1:1 (líneas negras y rojas) para obtener K_D y los residuales del ajuste, para las asociaciones RbAB- AdE1A.WT (trabajando con sonda AdE1A.WT-FITC a 0.5 nM y 2 nM, izquierda) y RbAB-E2F2 (trabajando con sonda E2F2-FITC a 1nM y 5 nM, derecha). C) modelo de asociación bimolecular de dos estados (especies iniciales libres y complejo final unido) para el caso de la interacción de AdE1A.WT con RbAB.

La Tabla 2.4 muestra los valores de estequiometrías, K_D y ΔG asociados que fueron hallados para las distintas interacciones testeadas. En la Figura 2.11 A, se muestran dos ejemplos representativos de curvas con las que se determinaron las estequiometrías. Como se puede observar, el valor de la señal de anisotropía creció linealmente hasta una razón molar 1:1 (agente titulado : agente titulante) donde alcanzó un valor constante que indica la saturación

de los sitios de unión. Esto revela que hay un único sitio de unión para cada sonda por monómero de RbAB, por lo que estas interacciones siguen una estequiometría 1:1. Todas las mediciones restantes saturaron a la misma razón molar 1 : 1, indicando estequiometrías igual a 1 (Tabla 2.4). Por otro lado, dado que todas las reacciones presentaron estequiometrías $n= 1$, se utilizó para el ajuste de las curvas de unión (Figura 2.11 B) un modelo de asociación bimolecular de dos estados (especies iniciales libres y complejo final unido) (Figura 2.11 C).

El resto de las curvas de asociación para determinar estequiometrías y constantes de disociación se encuentran en el Anexo (Figura A2.1 a Figura A2.8). En la Tabla A2.1 del Anexo se muestran todas las mediciones realizadas para las interacciones ensayadas y las distintas concentraciones de sonda a las que se realizaron las curvas de titulación junto con los valores encontrados para estequiometrías, K_D de ajuste de datos crudos y normalizados y de ajuste global, además de los valores iniciales y finales de anisotropía (Más detalles en Tabla A2.1 y en Materiales y Métodos).

Interacción de motivos mínimos celulares y virales a RbAB:

En el análisis de los resultados, podemos ver que el péptido E2F2 humano, que comprende al motivo homónimo de interacción a RbAB, se unió a esta última con una $K_D= 1\text{nM}$ mientras que el péptido viral AdE1ACR1, que también contiene el motivo E2F, se unió a RbAB con una $K_D= 119\text{nM}$. Estas diferencias muestran que pese a que se trata del mismo motivo, el péptido celular es 100 veces más afín por RbAB que el viral. Si se observan nuevamente las estructuras cristalinas de ambos complejos [RbAB:E2F2] y [RbAB:E1A-CR1] (Figuras 2.3 C y D) esta diferencia en afinidad a favor del motivo celular se debería a la presencia del segmento extendido N-terminal anexo al motivo mínimo en hélice que presenta E2F2 humano, el cual media adicionales contactos hidrofóbicos y puentes de hidrógeno con RbAB. En el caso del péptido viral AdE1ACR2 que contiene el motivo LxCxE, este interactuó con RbAB con una $K_D= 105\text{nM}$, similar afinidad a la presentada por el motivo E2F viral.

Análisis de la contribución de la región ácida a la interacción del motivo LxCxE con RbAB:

Por su parte AdE1ACR2_{AC}, que posee el parche ácido cargado negativamente contiguo al motivo LxCxE, se unió a RbAB con una mayor afinidad, presentando $K_D= 73\text{nM}$. Si bien esta diferencia es pequeña (cambio en la afinidad en menos de la mitad o medio-“fold”), entendida en el contexto de otros datos cuantitativos -discutidos en la primer sección de este Capítulo y que revelan el aporte que realizan parches cargados negativamente inmediatos al motivo LxCxE en otras proteínas (Tabla 2.2)-, estaría también cumpliendo un rol modulador

de afinidad mostrando un incremento en la misma dado probablemente por la complementariedad de cargas a la superficie positiva que rodea al surco de LxCxE en RbAB.

Análisis de la contribución de la región “linker” a la interacción con RbAB:

Si se comparan las K_D de los fragmentos que contienen los motivos mínimos de interacción con las mutantes de motivos, es posible evaluar la contribución de la región “linker”. Por un lado al comparar a interacción de AdE1A- Δ E con RbAB frente a la de AdE1ACR2_{AC} y por otro lado la interacción de AdE1A- Δ L frente a la de AdE1ACR1, se puede observar que hay poca diferencia entre sus K_D 's y ΔG 's (Tabla 2.4), lo que indica que en las condiciones ensayadas, la región “linker” aportaría muy poco significativamente a la energía de unión con RbAB.

Interacción de AdE1A.WT conteniendo ambos motivos de unión a RbAB:

Sorprendentemente, para AdE1A.WT, que posee ambos motivos unidos por la región “linker”, su interacción con RbAB rindió una $K_D = 24$ pM, dando lugar a un complejo [RbAB:AdE1A.WT] altamente estable energéticamente ($\Delta G = -14.2$ kcal/mol). Esta afinidad es unas 40 veces superior a la presentada por E2F2 humano a RbAB y alrededor de 4000 veces más alta que la que presentan los motivos virales testeados individualmente. Esto revela que hay un efecto sinérgico en la interacción simultánea de ambos motivos cuando están unidos por el “linker” en AdE1A a RbAB, generando tal incremento en la afinidad global de AdE1A.WT por RbAB.

***Concluyendo**, mientras los motivos virales individuales presentan afinidades de interacción a RbAB similares entre sí, estas son significativamente menores a la que presenta E2F2 humano por RbAB. Por otro lado, la presencia de la región “linker” en la proteína viral no aporta per se significativamente a la interacción con RbAB desde el punto de vista termodinámico, como lo demuestran la similitud en los parámetros K_D y ΔG de las interacciones de los péptidos virales con motivos mínimos en comparación con las proteínas con uno de los motivos mutados. No obstante, cuando el “linker” se encuentra uniendo ambos motivos en AdE1A.WT, estos podrían unir en forma simultánea a RbAB, dando como resultado una muy alta afinidad de interacción y generando un complejo energéticamente muy estable.*

II.2.C.2. Determinación de las contribuciones energéticas asociadas a la interacción de RbAB-AdE1A por Calorimetría de Titulación Isotérmica

El análisis cuantitativo de los sistemas estudiados no se acaba con la determinación de sus constantes de disociación. La descripción de los factores que gobiernan los eventos de asociación molecular requiere de la determinación de parámetros termodinámicos que incluyen los cambios en la energía libre de Gibbs (ΔG), la entalpía (ΔH) y la entropía (ΔS) como así también el cambio en la capacidad calorífica (ΔCp). El método por excelencia para determinar estos parámetros lo constituye la Calorimetría de Titulación Isotérmica (ITC). Esta metodología permite la medición directa de los calores involucrados (absorbidos o liberados) en las interacciones a una temperatura constante, y bajo condiciones adecuadas de análisis de los datos permite determinar en forma directa cambios de entalpía (ΔH), como también constantes de asociación (K_A , siendo $K_A^{-1}=K_D$) y la estequiometría de reacción (n) y deducir consecuentemente el cambio de energía libre (ΔG) y el cambio de entropía (ΔS) asociados [192]. Estos datos termodinámicos son importantes para elucidación de mecanismos y su relación con estructuras asociadas en las interacciones proteína-ligando.

Para todas las interacciones ensayadas se utilizaron concentraciones de reactivo en celda de 5, 15 ó 30 μM y en la jeringa, el titulante 10 veces más concentrado. Para la mayoría de las determinaciones, estas concentraciones aseguraron que el parámetro adimensional C ($C = [\text{Reactivo}_{\text{celda}}]/K_D$) -que idealmente debiera estar entre 10 y 1000 [195]- estuviese entre 20 y 800. Esto es fundamental, ya que la confianza en el valor hallado de K_A (K_D^{-1}) -que surge del ajuste de la curva- está determinada por la forma de la misma, la cual depende a su vez del valor C [192, 195]. Si la interacción es muy débil se deben usar altas concentraciones de reactivo para que C sea > 10 ; si la interacción es muy afín, para que C sea < 1000 , la concentración debe ser baja, pero esto trae aparejado consigo una poca cantidad de calor de asociación intercambiado en la titulación. Esto último fue el caso de la interacción RbAB-AdE1A.WT la cual es altamente afín - $K_D = 24 \text{ pM}$ determinado por EF-, por lo que se utilizó una concentración de compromiso (con $C > 1000$), que si bien no permitió el ajuste de la curva para hallar K_D por ITC - muy por debajo de su límite de sensibilidad- la cantidad de calor intercambiado permitió la determinación confiable del ΔH asociado.

Las determinaciones se realizaron a una temperatura de trabajo de 20°C . Como en el caso de las interacciones RbAB- AdE1ACR1 y RbAB- AdE1A- ΔL la cantidad de calor intercambiado en la titulación fue muy bajo para lograr un ajuste confiable de estas isotermas, se ajustaron temperaturas y concentraciones, realizando titulaciones en una serie de

temperaturas y utilizando un equipo nano-ITC (Ver Materiales y Métodos) que reduce hasta 10 veces la cantidad de proteínas necesaria por titulación. Por este motivo, para la interacción RbAB- AdE1ACR1 se informa los datos obtenidos a 10°C.

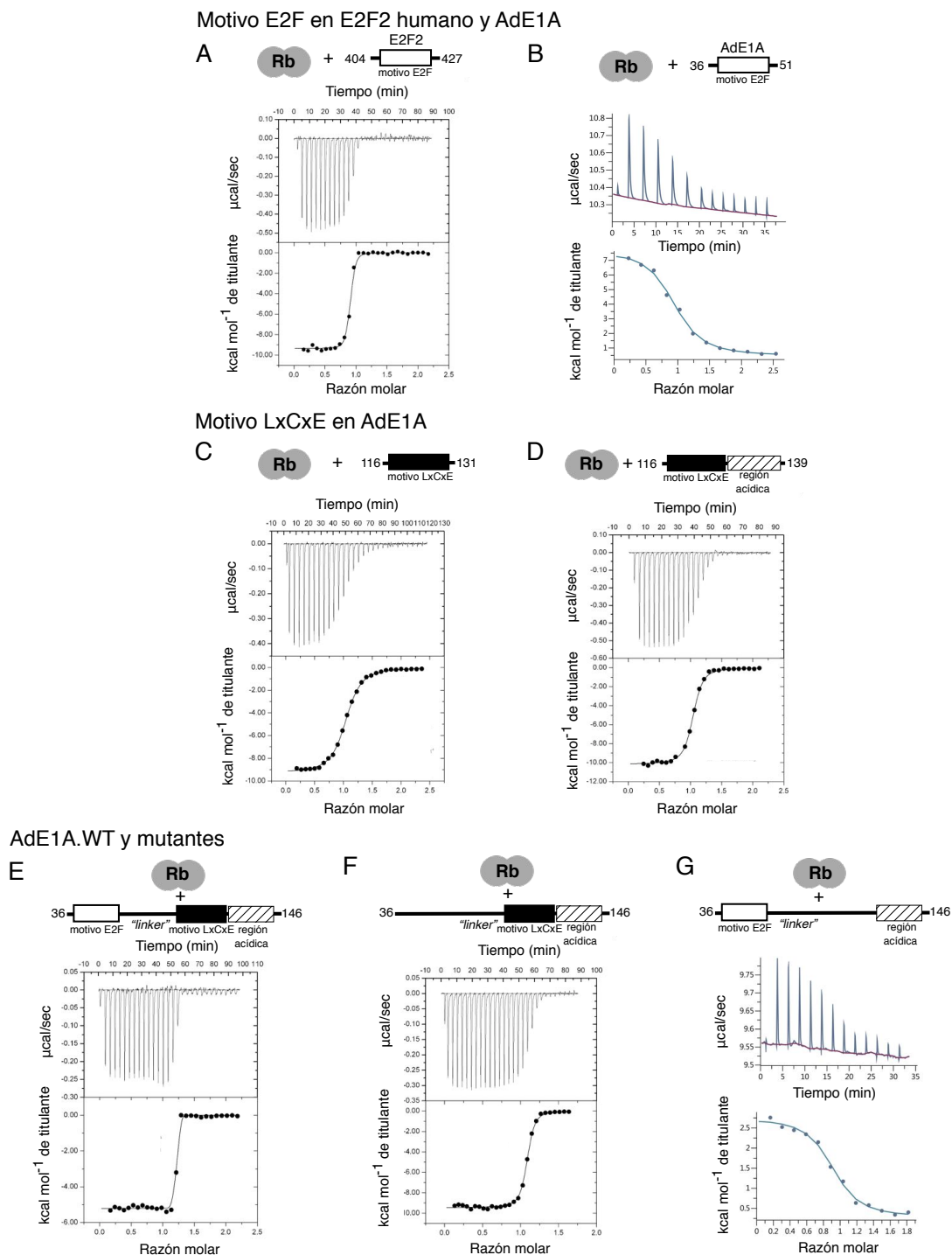
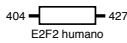
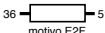

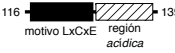
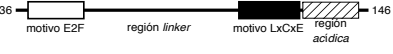
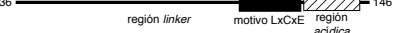
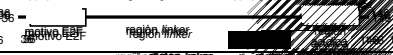


Figura 2.12: Curvas representativas de ITC de las interacciones entre RbAB y los diferentes péptidos y proteínas testeados. Las mediciones fueron hechas en fosfato de sodio 20mM, 200mM NaCl, 5 mM 2-mercapto etanol a pH 7.0. La celda fue cargada con RbAB aislado y la jeringa con los diferentes péptidos o proteínas titulantes. Curvas de asociación para A) RbAB (5 μ M) y E2F2humano (50 μ M); B) RbAB (30 μ M) y AdE1ACR1(300 μ M); C) RbAB (15 μ M) y AdE1ACR2 (150 μ M); D) RbAB (15 μ M) y AdE1ACR2_{AC} (150 μ M); E) RbAB (15 μ M) y AdE1A.WT (150 μ M); F) RbAB (15 μ M) y AdE1A- Δ E (150 μ M); G) RbAB (30 μ M) y AdE1A- Δ L(300 μ M). Las mediciones fueron hechas a 20 °C excepto por la titulación RbAB- AdE1ACR1 medida a 10 °C.

La Tabla 2.5 resume los valores medios para los parámetros termodinámicos ajustados y calculados a partir de las isotermas, algunas de ellas representadas en la Figura 2.12. En el Anexo, se presentan todas las curvas ITC realizadas para las interacciones directas con RbAB desde Figura A2.9 a Figura A2.16. Como se puede observar en la Tabla 2.5, todas las asociaciones siguieron una estequiometría $n=1$ (con valores entre 0.81 y 1.19), es decir que por cada mol de monómero de RbAB existe un sitio único de interacción para cada titulante utilizado. Todas las asociaciones mostraron poseer un carácter exotérmico a la temperatura ensayada, excepto por las interacciones RbAB- AdE1ACR1 y RbAB- AdE1A- Δ L, que mostraron ser endotérmicas.

Tabla 2.5. Parámetros termodinámicos de las interacciones entre péptidos y proteínas AdE1A y E2F2 humana con RbAB hallados por ITC

	Pep o Prot	n	ITC				EF	
			ΔH (cal/mol)	$-T\Delta S$ (cal/mol)	ΔG (cal/mol)	K_D (nM)	K_D (nM)	ΔG (cal/mol)
	E2F2	0.94 ± 0.01	-8646 ± 174	-2170 ± 302	-10817 ± 247	8.7 ± 3.6	1.03 ± 0.03	-12057 ± 16
	AdE1ACR1	0.86 ± 0.02	7205 ± 338	-15005 ± 366	-7800 ± 139	1535 ± 392	119.1 ± 2.1	-9286 ± 10
	AdE1ACR2	1.03 ± 0.01	-9522 ± 79	573 ± 92	-8949 ± 47	213 ± 17	104.6 ± 2.9	-9362 ± 16
	AdE1ACR2_{AC}	1.02 ± 0.01	-10127 ± 122	511 ± 161	-9615 ± 104	68.5 ± 12.3	73.4 ± 2.2	-9577 ± 18
	AdE1A-WT	1.19 ± 0.01	-5284 ± 40	-8960 ± 105§	<i>nd</i>	<i>nd</i>	0.024 ± 0.004	-14244 ± 97
	AdE1A-ΔE	0.92 ± 0.01	-10784 ± 95	452 ± 140	-10332 ± 103	19.9 ± 3.8	47.1 ± 1.7	-9828 ± 21
	AdE1A-ΔL	0.92 ± 0.05	2447 ± 291	-10553 ± 477	-8106 ± 378	904 ± 578	206.7 ± 3.1	-8966 ± 9

Todas las mediciones fueron hechas por triplicado excepto por las titulaciones RbAB - AdE1ACR1 y RbAB - AdE1A.WT las cuales fueron hechas por duplicado y RbAB - AdE1A- Δ E por cuadruplicado.

Las mediciones fueron hechas a 20 °C excepto por la titulación RbAB- AdE1ACR1. Para esta interacción se reporta la medida a 10 °C.

§ Calculado usando ΔG obtenido de los datos de EF.

nd: no determinado

ΔH y K_D son valores experimentales y fueron obtenidos con errores luego de los ajustes. ΔG (de ITC y EF) fue calculado como $\Delta G = RT \ln K_D$ y $-T\Delta S = \Delta G - \Delta H$. Los errores de ΔG y $-T\Delta S$ fueron obtenidos por propagación.

Los valores C para cada experimento de ITC reportado en la tabla fueron: RbAB-E2F2: 575; RbAB-AdE1ACR1: 24, RbAB-AdE1ACR2: 71; RbAB-AdE1ACR2_{AC}: 219; RbAB-AdE1A- Δ E: 754 y RbAB-AdE1A- Δ L: 26.

Contribuciones energéticas en la interacción de motivos mínimos a RbAB

En el caso del péptido E2F2 humano, éste se unió a RbAB con una $K_D = 8.7$ nM (Figura 2.12 A). Esta K_D también está en el orden nanomolar bajo, como la hallada por EF (1nM), mostrando un buen acuerdo considerando que se tratan de determinaciones hechas por técnicas distintas (y que difieren generalmente entre sí en unas 2 a 10 veces los valores de K_D). La formación del complejo [RbAB:E2F2] energéticamente estable, está entálpicamente dirigido (ΔH negativo) y también favorecido entrópicamente ($-T\Delta S$ negativo). Por su parte, el péptido

viral AdE1ACR1 se unió a RbAB con una $K_D = 1.5 \mu\text{M}$ que difiere del hallado por EF (119nM) no sólo al uso de metodologías distintas, sino a que la reportada para ITC fue determinada a 10°C (283.15 K), siendo este un parámetro dependiente de temperatura. Sorprendentemente, la reacción entre AdE1ACR1 y RbAB se encuentra entrópicamente dirigida y entálpicamente desfavorecida (Tabla 2.5). y a la temperatura ensayada es endotérmica, en contraposición a lo caracterizado para la reacción entre RbAB y E2F2 humano. Analizando todos estos datos termodinámicos conjuntamente con las estructuras cristalográficas de sus complejos (Figura 2.3 C y D), la transición que experimenta el motivo mínimo E2F de desordenado-a- α -hélice -que genera complementariedad óptima para contactos- acoplado a interacción con el sitio complementario a E2F en RbAB, estaría dirigido entrópicamente (si bien el plegamiento tiene un costo entrópico desfavorable, la desolvatación de superficies que acompaña esta estructuración es favorable y preponderante en la interacción –se discutirá más adelante-). La contribución entálpica en tanto, es más importante en el caso de la interacción RbAB- E2F2, en donde además del plegamiento-acoplado-a-interacción descrito, se producen contactos hidrofóbicos y puentes de hidrógeno estabilizantes adicionales con RbAB, dados por la presencia del segmento extendido N-terminal extra al motivo E2F2 humano. Estos contactos “extras” contribuirían preponderantemente a la interacción (componente entálpico negativo), generando además un complejo más estable que [RbAB:E1A-CR1] ($\Delta\Delta G = 3 \text{ kcal/mol}$).

El péptido viral AdE1ACR2 interactuó con RbAB con una $K_D = 213 \text{ nM}$ (Figura 2.12 C), en excelente acuerdo con lo hallado por EF. Esta interacción está levemente desfavorecida entrópicamente y dirigida entálpicamente, consecuencia de las interacciones complementarias formadas entre el péptido y el surco LxCxE en RbAB.

La contribución entálpica de la región acídica a la interacción del motivo LxCxE con RbAB

El péptido AdE1ACR2_{AC} (motivo LxCxE + parche acídico contiguo) se unió a RbAB con K_D en excelente acuerdo a lo hallado por EF, en una reacción también dirigida entálpicamente y levemente desfavorecida entrópicamente (Tabla 2.5). El efecto estabilizante de la región acídica respecto del motivo mínimo LxCxE - AdE1ACR2- ($\Delta\Delta G = 0.7 \text{ kcal/mol}$), es puramente debido a un componente entálpico, lo que se correlaciona con el carácter electrostático descrito para la interacción de esta región con RbAB. Esto revela el rol modulador del parche acídico en la afinidad del motivo LxCxE con RbAB.

Contribuciones de la región “linker” a la interacción con RbAB

Para AdE1A- Δ E la interacción con RbAB (Figura 2.12 F) reportó valores de K_D y ΔG aproximados a los obtenidos por EF y en comparación con el cambio de energía libre de la interacción RbAB- AdE1ACR2_{AC}, hay mínimas diferencias. Por otro lado si se comparan las interacciones entre AdE1A- Δ L y AdE1ACR1 con RbAB llevadas a cabo a las mismas temperaturas (un análisis más extenso de estas dos interacciones se presenta después) se observan también sólo mínimas diferencias entre K_D 's y el ΔG 's. En conjunto, esto los resultados evidencian el aporte poco significativo que la región “*linker*” genera por sí misma, haciendo suponer que no existe un tercer sitio o motivo de interacción independiente en dicha región que contribuya globalmente a la energía de unión con RbAB.

Interacción de AdE1A.WT conteniendo ambos motivos de unión a RbAB

Para la interacción entre AdE1A.WT y RbAB (Figura 2.12 E), no fue posible hallar una K_D por ITC, por estar muy debajo de su límite de sensibilidad –del orden picomolar hallado por EF- por lo que sólo se determinó un valor de ΔH (-5.3 kcal/mol) asociado a la reacción, en tanto que $-T\Delta S$ fue calculado usando el dato de ΔG obtenido de EF (Tabla 2.5). Estos datos muestran que a diferencia de lo que ocurre con las interacciones individuales de los motivos E2F y LxCxE en AdE1A con RbAB, dirigidos exclusivamente por un factor entrópico o entálpico, AdE1A.WT interacciona con RbAB en una reacción favorecida tanto entálpica como entrópicamente (levemente dominante). Además como se puede observar, los valores que asumen sus parámetros termodinámicos no resultan de la suma aritmética de contribuciones complementarias, es decir que ni su ΔH , $-T\Delta S$ ni ΔG resultan de la sumatoria de los valores de ΔH , $-T\Delta S$ y ΔG de la interacción RbAB- AdE1ACR1 más los de la de RbAB- AdE1A- Δ E, ni las de la sumatoria entre los parámetros termodinámicos de RbAB- AdE1ACR2_{AC} más los de RbAB- AdE1A- Δ L. Todo ello es indicativo de que en la interacción entre AdE1A.WT y RbAB se produce una compensación entálpico-entrópica en la que los motivos actúan sinérgicamente estando unidos a través la región “*linker*”, dando lugar al complejo [RbAB:AdE1A.WT] altamente estable.

Resumiendo, estos resultados de ITC muestran por un lado que los factores termodinámicos que dirigen las reacciones de interacción entre los motivos mínimos E2F y LxCxE en AdE1A con RbAB implican procesos de plegamiento-acoplado-a-interacción generando complementariedad de superficies, así como una serie clave de interacciones hidrofóbicas y puentes de hidrógeno con residuos complementarios en RbAB, dando lugar a

complejos estables. La presencia de regiones extras flanqueantes a un motivo mínimo de interacción son energéticamente cruciales, así por ejemplo, en E2F2 humano el segmento N-terminal adicional al motivo cambia el perfil energético de la reacción con RbAB con respecto a lo que ocurre con AdE1ACR1, tornándose mayoritariamente en un proceso entálpico. Por otro lado, las contribuciones energéticas de los motivos E2F y LxCxE en AdE1A no son aditivas cuando estando enlazados por la región "linker", reaccionan simultáneamente con RbAB, dando lugar a un proceso donde hay una compensación entálpico-entrópica que origina un complejo ([RbAB:AdE1A.WT]) que es incluso energéticamente más estable que los formados por motivos individuales y que el complejo celular [RbAB:E2F2humano].

En este trabajo se logró medir todos los parámetros termodinámicos de la interacción entre AdE1A y RbAB -analizando en una sola serie cada uno de sus fragmentos- en condiciones de trabajo adecuadas y controladas, lo que permite realizar análisis cuantitativos consistentes. Los valores hallados tanto por EF como por ITC muestran un gran acuerdo generando confiabilidad en los resultados aportados. Además, los resultados aquí presentados son también consistentes con los reportados en literatura para la interacción de fragmentos equivalentes (tanto de E2F2 como de AdE1A) con RbAB (ver Tabla 2.2), lo que implica mutua validación y confianza en dichos datos hallados de manera independiente por distintos laboratorios.

II.2.C.3. Análisis de la transición desorden-a-orden del motivo E2F viral y contribución del "linker" por titulaciones calorimétricas dependientes de temperatura. Interacciones entre AdE1ACR1 y AdE1A- Δ L con RbAB

Como fue mencionado con anterioridad, la mayoría de las titulaciones de ITC para las interacciones ensayadas se realizaron a 20°C, sin embargo para las interacciones de RbAB con AdE1ACR1 y AdE1A- Δ L los primeros ensayos realizados a esta temperatura habían mostrado que la cantidad de calor intercambiado con el sistema durante la titulación era un tanto insuficiente para hallar parámetros termodinámicos con suficiente grado de confianza. No obstante estos primeros análisis ya fueron lo suficientemente informativos, para permitir describir estas interacciones como procesos endotérmicos a la temperatura ensayada, correspondiendo justamente a interacciones en las que se encontraba involucrado el motivo E2F viral, el cual experimenta una transición de desordenado-a-ordenado acoplado a la interacción con RbAB.

Por otro lado, se ha visto de modo notable que los eventos de desorden-a-orden acoplado a interacciones que experimentan muchas IDPs al unirse a otra proteína, están acompañados de cambios significativos de entalpía dependientes de la temperatura, incluso dentro de un rango experimentalmente accesible, haciendo que los parámetros termodinámicos determinados a una única temperatura sean menos informativos de lo que típicamente se asume [196, 197].

Con estas consideraciones arriba expuestas e intenciones de generar datos confiables e interpretables para estas interacciones, se hicieron en principio ensayos de ITC -utilizando un nano-calorímetro- a una temperatura de 10°C y luego fueron finalmente completados ensayos en dos series de temperaturas: 10, 15, 20 y 30°C para la interacción RbAB-AdE1ACR1 y a 10, 20 y 30°C para la interacción RbAB-AdE1A-ΔL. Con los datos obtenidos en estas titulaciones se realizó un análisis global que permite encontrar contribuciones energéticas y parámetros estructurales de las interacciones, como las áreas y número de residuos involucrados. Además la comparación entre las interacciones entre RbAB-AdE1ACR1 y RbAB-AdE1A-ΔL, permitió dilucidar el rol del “*linker*” en la interacción.

Interpretación de los parámetros termodinámicos para el sitio CR1 (motivo E2F)

Los datos obtenidos en estas series de titulaciones dependientes de temperatura, se resumen en la Tabla 2.6.

Tabla 2.6. Titulaciones calorimétricas dependiente de Temperatura para las interacciones entre AdE1ACR1 y AdE1A-ΔL con RbAB

Titulante	°C	#	n	ΔH (cal/mol)	$-T\Delta S$ (cal/mol)	ΔG (cal/mol)	K_D (nM)
AdE1ACR1	10	2	0.86 ± 0.02	7205 ± 338	-15005 ± 365	-7800 ± 139	1535 ± 391
AdE1ACR1	15	1	0.86 ± 0.01	4130 ± 122	-12071 ± 157	-7941 ± 99	1200 ± 205
AdE1ACR1	20	1	0.73 ± 0.06	1100 ± 144	-9102 ± 469	-8003 ± 446	1080 ± 828
AdE1ACR1	30	1	0.85 ± 0.02	-5380 ± 246	-2686 ± 297	-8066 ± 168	969 ± 279
AdE1A-ΔL	10	2	0.91 ± 0.03	8675 ± 599	-16462 ± 623	-7787 ± 172	1590 ± 487
AdE1A-ΔL	20	3	0.92 ± 0.05	2447 ± 291	-10553 ± 477	-8106 ± 378	904 ± 578
AdE1A-ΔL	30	1	0.94 ± 0.03	-3450 ± 225	-4645 ± 305	-8095 ± 206	922 ± 326

número de determinaciones

ΔH y K_D son valores experimentales y fueron obtenidos con errores luego de los ajustes. ΔG fue calculado como $\Delta G = RT \ln K_D$ y $-T\Delta S$ como $-T\Delta S = \Delta G - \Delta H$. Los errores de ΔG y $-T\Delta S$ fueron obtenidos por propagación.

En primer lugar, ambas interacciones RbAB- AdE1ACR1 y RbAB- AdE1A-ΔL presentaron estequiometrías n=1, independientemente de la temperatura ensayada. Como se puede observar, tanto una como otra interacción presentaron el mismo comportamiento: a las

menores temperaturas ensayadas las reacciones se caracterizaron por estar entrópicamente dirigidas (valores $-T\Delta S$ negativos) y entálpicamente desfavorecidas (valores ΔH positivos), pero con el incremento de la temperatura se produjo una inversión de las contribuciones energéticas y pasaron a estar entrópicamente desfavorecidas y entálpicamente dirigidas, experimentando compensación entálpico-entrópica y por consecuencia una transición de interacción endotérmica a exotérmica. Estas variaciones se simplificaron de manera visual en la Figura 2.13, donde se puede apreciar los perfiles energéticos de ambas interacciones a las temperaturas ensayadas y la compensación de las contribuciones energéticas con el aumento de la temperatura (Figura 2.13 A y D). A su vez, ΔG para ambos casos, permaneció relativamente constante (Figura 2.13 B y E). Esto último sería más importante que el fenómeno de compensación entálpico-entrópica en sí mismo, porque implica que el proceso de plegamiento-acoplado-a-interacción del motivo E2F viral a RbAB sería robusto a fluctuaciones en la temperatura, en un rango de temperaturas cercanas a la biológicamente toleradas [196].

Análisis global de los datos calorimétricos. Parámetros estructurales para la interacción entre el motivo E2F viral y RbAB

A continuación, se realizó el análisis global de los datos calorimétricos dependientes de temperatura, el cual es importante porque minimiza la incertidumbre de los datos termodinámicos encontrados cerca de la temperatura de transición (temperatura en la que el calor intercambiado con el sistema es muy bajo) y porque permite el cálculo de la capacidad calórica a presión constante (ΔC_p) haciendo posible una interpretación más robusta de los parámetros termodinámicos que están involucrados en la interacción [196].

Los cambios en ΔH en función de la temperatura, como los aquí observados, hacen experimentalmente accesible el cálculo de ΔC_p . Para las dos interacciones estudiadas este parámetro fue hallado como la pendiente que resulta del análisis de regresión lineal del gráfico ΔH vs temperatura (Figura 2.13 C y F). Para la interacción RbAB- AdE1ACR1 ΔC_p fue de -629 cal/molK, mientras que un valor muy similar ($\Delta C_p = -606$ cal/molK) fue hallado para la interacción RbAB- AdE1A- Δ L. Estos valores altos de ΔC_p hallados implican que estas interacciones son sitio-específicas, y su signo negativo, que hay remoción de grandes cantidades de agua – desolvatación o efecto hidrofóbico- de las superficies interactuantes durante la formación de los complejos [198] [RbAB:AdE1ACR1] y [RbAB:AdE1A- Δ L].

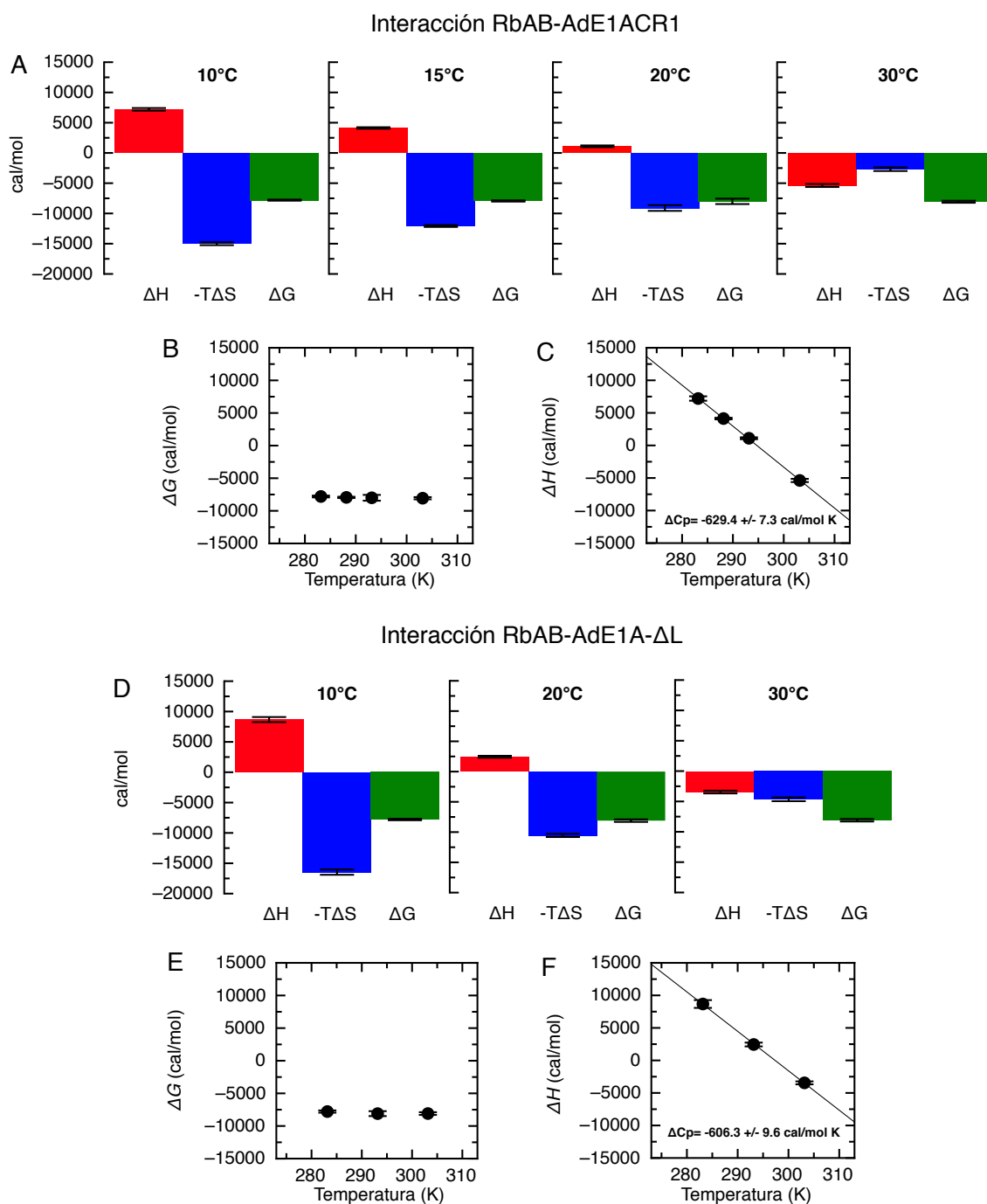


Figura 2.13: Análisis de los parámetros termodinámicos de las titulaciones calorimétricas dependiente de temperatura para las interacciones RbAB-AdE1ACR1 y RbAB-AdE1A-ΔL. A, D) Variaciones de ΔH (barra roja), $-T\Delta S$ (barra azul) y ΔG (barra verde) con las temperaturas crecientes. Como se puede observar mientras ΔH y $-T\Delta S$ evidencian una compensación entálpico-entrópica, ΔG permanece constante. B, E) Gráfico ΔH vs *Temperatura*. De la pendiente del ajuste de regresión lineal fue obtenido ΔC_p . C, F) Gráfico ΔG vs *Temperatura*. Como se puede observar, hay muy poca variación de ΔG con temperaturas crecientes.

ΔC_p constituye el parámetro básico para interpretar los datos termodinámicos en términos estructurales ya que se asocia a cambios en la cantidad de área superficial expuesta al solvente (ΔASA_T) tras producirse la interacción [199, 200]. Esta asociación está dada por relaciones empíricas propuestas que permiten conocer las contribuciones de cambios en superficies polares (ΔASA_p) y no polares (ΔASA_{np}) involucradas en la interfaz de interacción. ΔASA_{np} , ΔASA_p y su sumatoria (ΔASA_T), fueron calculadas como se encuentra descrito en Materiales y Métodos y los valores presentados en la Tabla 2.7.

Tabla 2.7. Parámetros estructurales derivados del análisis de las titulaciones calorimétricas dependiente de Temperatura para las interacciones entre AdE1ACR1 y AdE1A- ΔL con RbAB

Titulante	ΔC_p (cal/mol K)	ΔASA_{np} (Å^2)	ΔASA_p (Å^2)	ΔASA_T (Å^2)
AdE1ACR1	-629.4 ± 7.3	-1623 ± 19	-389 ± 11	-2013 ± 247
AdE1A- ΔL	-606.3 ± 9.6	-1534 ± 26	-323 ± 15	-1857 ± 418

El ΔASA_T calculado para la interacción RbAB- AdE1ACR1 fue de -2013 Å^2 y son superficies no polares las que están principalmente involucradas en la interacción ($\approx 80\%$ de la superficie total). Esto coincide además con el carácter hidrofóbico de la interacciones de la hélice E2F con el sitio complementario a E2F en RbAB. Los datos de ΔASA_T calculados están en muy buen acuerdo con el valor $\Delta ASA_T = 1623 \text{ Å}^2$ encontrado a partir de datos cristalográficos del complejo [RbAB:E1A-CR1] (PDBID 2R7G) en el trabajo de Liu y Marmorstein [84]. Las pequeñas discrepancias que podrían deberse al uso de metodologías distintas para calcular ΔASA_T y/o al hecho de que los complejos pueden estar menos compactos en solución (abarcando mas superficie) en comparación con el grado de compactación del complejo cristalizado [198].

Contribuciones entrópicas en la interacción entre el motivo E2F viral y RbAB

Los altos valores negativos de ΔC_p y extensas superficies no polares involucradas en estas interacciones, sugieren entonces que la entropía de solvatación -efecto hidrofóbico- es el principal componente entrópico dirige el proceso de plegamiento-acoplado-a-interacción del motivo E2F viral con RbAB en el rango de temperaturas evaluado. Es decir, que el aumento de entropía del agua en todo el sistema es el factor clave en este proceso de interacción. Ahora bien, además del efecto hidrofóbico (ΔS_{solv}), otras contribuciones entrópicas están involucradas en el proceso: la entropía rotacional-traslacional (ΔS_T) y la entropía configuracional (ΔS_{config}). La sumatoria de estos 3 componentes da la entropía total asociada a la interacción (ΔS_{int}).

Estas contribuciones fueron calculadas y presentadas en la Tabla 2.8. Como se puede observar, ΔS_{rt} se asume con un valor fijo de -8 cal/K mol [192] para ambas interacciones estudiadas, mientras que los valores de ΔS_{config} y ΔS_{solv} mostraron poca variación para el rango de temperaturas ensayado para ambas interacciones. En ambos casos, ΔS_{rt} constituye una contribución desfavorable a la interacción entre el motivo E2F viral y RbAB porque implica una reducción de los grados de libertad traslacional y rotacional tras formarse el complejo. De igual modo, ΔS_{config} también es desfavorable a la interacción, porque la transición del motivo E2F de desordenado-a- α -hélice acoplado a interacción con RbAB reduce su entropía conformacional. No obstante, la principal contribución entrópica es el efecto de desolvatación de las superficies interactuantes la cual dirige el signo de ΔS_{int} ($\Delta S > 0$ y $-T\Delta S < 0$ en estas interacciones).

Tabla 2.8. Contribuciones entrópicas para las interacciones entre AdE1ACR1 y AdE1A- Δ L con RbAB

Titulante	°C	ΔS (cal/mol K)	ΔS_{rt}^* (cal/mol K)	ΔS_{solv} (cal/mol K)	ΔS_{config} (cal/mol K)	X_{res}
AdE1ACR1	10	52.99 ± 1.29	- 8	193.39 ± 2.24	-132.4 ± 2.6	31
	15	41.89 ± 0.54	- 8	182.38 ± 2.12	-132.5 ± 2.2	31
	20	31.05 ± 1.59	- 8	171.55 ± 1.99	-132.5 ± 2.6	31
	30	8.86 ± 0.98	- 8	150.44 ± 1.74	-132.6 ± 2.0	31
AdE1A- Δ L	10	58.14 ± 2.21	- 8	186.29 ± 2.24	-120.2 ± 3.1	28
	20	35.99 ± 1.63	- 8	165.25 ± 1.99	-121.3 ± 2.6	28
	30	15.32 ± 1.01	- 8	144.92 ± 1.74	-121.6 ± 2.0	28

* por consideraciones teóricas y empíricas ΔS_{rt} asume valores entre -4 y -10 cal/K mol, tomándose como referencia el valor de -8 cal/K mol [195].

Por otro lado, se estimó el número de residuos involucrados en las interacciones para el complejo [RbAB:AdE1ACR1] para cada temperatura, encontrando unos 31 residuos involucrados en cada complejo respectivamente (Tabla 2.8). Si bien esta es un cálculo un tanto grosero, resulta sorprendente la concordancia con la estimación de un total de 32 residuos interactuantes (23 de RbAB y 9 de AdE1ACR1) encontrados a partir del análisis de la estructura cristalográfica del complejo [RbAB:E1A-CR1] (PDBID: 2R7G) (Figura 2.3 D) y localizados en menos de 4Å uno respecto de otro residuo en la otra proteína interactuante (4Å se consideró como el límite superior en distancia para que se puedan dar enlaces no-covalentes).

En resumen, la interacción específica del motivo E2F en AdE1A con RbAB está fuertemente guiada por un efecto entrópico hidrofóbico que resulta de la desolvatación de las

superficies interactuantes -fundamentalmente no polares- cuando se produce el plegamiento-e-interacción entre ambos y que supera la penalidad entrópica de plegar el motivo. Además, esta interacción presenta un fenómeno de compensación entálpico-entrópica en un rango de temperaturas cercanas a la biológicamente tolerada y es robusta a estas fluctuaciones térmicas porque su energía libre asociada permanece constante. Esto implica que independientemente de cambios térmicos (dentro de un rango que no conlleve a la desnaturalización de RbAB) esta asociación va a proceder, con una afinidad moderada regulada fundamentalmente por la penalidad entrópica asociada al plegamiento del motivo E2F.

II.2.C.4. Análisis de la contribución de la región “linker” según parámetros estructurales

Si la región “linker” contribuye a la interacción con RbAB se esperaría que la especie AdE1A- Δ L presente un ΔASA_T mayor y más residuos estén en contacto con RbAB, comparativamente con el fragmento AdE1ACR1 en la interacción analizada. No obstante, como ya se mencionó, los parámetros termodinámicos muestran similitud en los valores asumidos temperatura a temperatura para sus interacciones con RbAB. Esto se refleja en una similitud de los parámetros estructurales derivados: para la interacción del motivo E2F (AdE1A-CR1) con RbAB el ΔASA_T fue de -2013 \AA^2 mientras que para la interacción de RbAB con el fragmento que además contiene el “linker” (AdE1A- Δ L) el ΔASA_T fue -1857 \AA^2 (Tabla 2.7). Por otro lado el número de residuos involucrados en la interacciones fue similar (28 vs 31) (Tabla 2.8). Estos resultados revelan que el “linker” no aporta interacciones “extras” en esta región que formen contactos y establezcan la unión con RbAB.

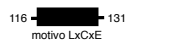
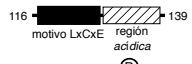

II.2.C.5. Efecto de la fosforilación de la serina 132 en la interacción entre AdE1A y RbAB

Varios estudios sobre AdE1A demostraron que esta proteína viral posee múltiples sitios que pueden ser fosforilados tanto *in vitro* como *in vivo*, conduciendo a cambios en sus propiedades bioquímicas y efectos celulares [185, 201, 202]. Mal y colaboradores demostraron que cuando AdE1A se encuentra fosforilada, sería más eficiente para asociarse a Rb, así como para irrumpir con el complejo celular [RbAB:E2F2humano] [202]. Por otro lado, Whalen y colaboradores reportaron en experimentos independientes, que la fosforilación en la serina 132 (S132) por CKII, podría ser responsable de este incremento en la eficiencia de asociación con RbAB, con un consecuente aumento en la capacidad transformante de AdE1A [185], ya que como fue planteado al inicio de este capítulo dicha fosforilación en la región contigua al motivo LxCxE modularía su fuerza de interacción con RbAB, como ocurre con otros motivos LxCxE descriptos.

Con el fin de estudiar el efecto de los residuos flanqueantes en la interacción de un motivo mínimo con su compañero de interacción, se cuantificó el aporte que produce la fosforilación en S132 sobre la interacción de AdE1A con RbAB. Para ello se realizaron los mismos experimentos de interacción con RbAB por EF e ITC utilizando un péptido sintético AdE1ACR2_{ACP}, en el que S132 está fosforilada.

En la Tabla 2.9 se presentan los resultados de estos experimentos y la Figura 2.14 muestra la isoterma de la interacción entre RbAB y AdE1ACR2_{ACP} (Figura 2.14 C), en comparación con las obtenidas para la interacción con RbAB de los otros dos péptidos CR2 de AdE1A. Todas las curvas de ITC obtenidas para el análisis de la interacción RbAB-AdE1ACR2_{ACP}, se encuentran en la Figura A2.13 del Anexo. Tanto por EF como por ITC, la estequiometría de reacción para la interacción en estudio fue n=1, como era esperable.

Tabla 2.9. Efecto de la fosforilación en S132 en la interacción entre AdE1A y RbAB

Pep o Prot	n	ITC				EF		ΔG (cal/mol)
		ΔH (cal/mol)	$-T\Delta S$ (cal/mol)	ΔG (cal/mol)	K_D (nM)	K_D (nM)		
 AdE1ACR2	1.03 ± 0.01	-9522 ± 79	573 ± 92	-8949 ± 47	213 ± 17	104.6 ± 2.9	-9362 ± 16	
 AdE1ACR2 _{AC}	1.02 ± 0.01	-10127 ± 122	511 ± 161	-9615 ± 104	68.5 ± 12.3	73.4 ± 2.2	-9577 ± 18	
 AdE1ACR2 _{ACP}	1.19 ± 0.01	-7900 ± 107	-2021 ± 168	-9921 ± 129	40.6 ± 9.1	19.6 ± 0.5	-10339 ± 16	

Todas las mediciones fueron hechas por triplicado y a 20 °C.

ΔH y K_D son valores experimentales y fueron obtenidos con errores luego de los ajustes. ΔG (de ITC y EF) fue calculado como $RT \cdot \ln K_D$ y $-T\Delta S$ como $\Delta G - \Delta H$. Los errores de ΔG y $-T\Delta S$ fueron obtenidos por propagación.

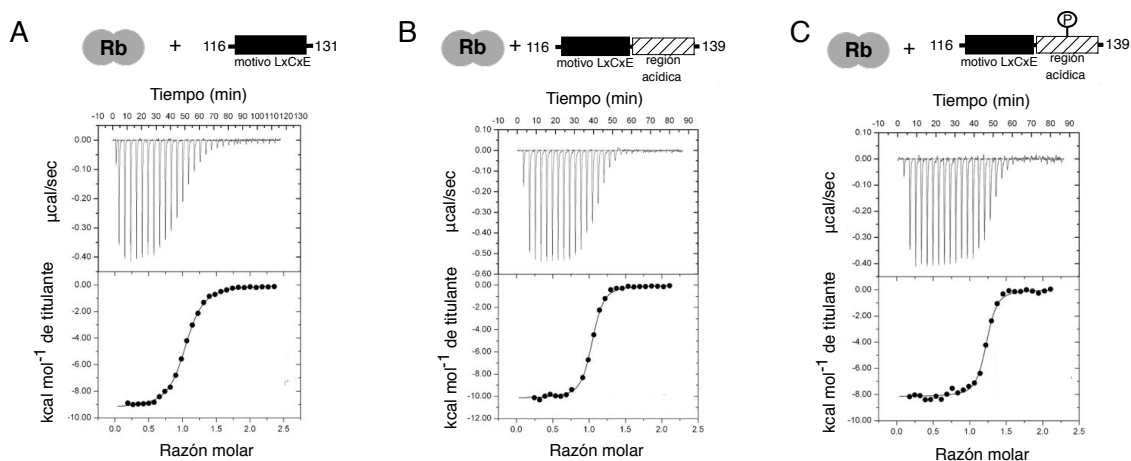


Figura 2.14: Curvas representativas de ITC para el análisis del efecto de la fosforilación en S132 en la interacción del motivo LxCxE viral con RbAB. Las mediciones fueron hechas en fosfato de sodio 20mM, 200mM NaCl, 5 mM 2-mercapto etanol a pH 7.0. La celda fue cargada con RbAB aislado y la jeringa con los diferentes péptidos o proteínas titulantes. Curvas de asociación para A) RbAB (15μM) y AdE1ACR2 (150μM); B) RbAB (15μM) y AdE1ACR2_{AC} (150μM); C) RbAB (15μM) y AdE1ACR2_{ACP}(150μM). . Las curvas A y B son las mismas presentadas con anterioridad, aquí nuevamente representadas a los fines comparativos.

Por EF y por ITC, la interacción entre RbAB y AdE1ACR2_{ACP} presentó K_D y ΔG similares. Como se puede ver en Tabla 2.9, la fosforilación aumentó la afinidad por RbAB unas 5 veces (revelado por ambas técnicas -EF e ITC-), respecto del motivo mínimo LxCxE – AdE1ACR2-. La diferencia de energía libre de unión entre ambos péptidos con RbAB, $\Delta\Delta G = -1$ kcal/mol, fue significativa en estos dos experimentos independientes, indicando que la fosforilación conduce a la formación de un complejo más estable energéticamente.

El análisis de las contribuciones energéticas en la interacción entre el péptido fosforilado con RbAB, muestra que esta reacción está favorecida y dirigida entálpicamente, pero sorprendentemente y a diferencia de lo que ocurre con los otros 2 péptidos CR2 de AdE1A, también hay una contribución entrópica favorable ($-T\Delta S$ negativo). Es decir que la fosforilación en S132 conduciría a un cambio tal, que la reacción entre el péptido y RbAB pasa de estar entrópicamente desfavorecida a favorecida. Una posible explicación a este fenómeno, sería que esta fosforilación afecta la conformación local de la región acídica conduciendo a una estabilización de estructura tipo PPII sobre el motivo LxCxE en AdE1A al interactuar con RbAB. Esto podría ser así, pues se ha visto que la fosforilación en S31 y S32, próximos al sitio LxCxE en HPV16-E7 estabiliza contenido de estructura tipo PPII y consecuentemente presenta afinidad incrementada a RbAB [109, 203]. En un paralelismo con el caso de la interacción del motivo E2F con RbAB, el cual al experimentar un plegamiento a α -hélice acoplado a interacción en un proceso fundamentalmente entrópico, probablemente aquí también estén implicados procesos de desolvatación de superficies que conduzcan a la estabilización de PPII -que es una hélice levógira extendida- optimizando la consecuente interacción con la superficie complementaria en RbAB. No obstante, se diferencia en que la penalidad entrópica asociada no es tan significativa (sigue siendo un proceso fundamentalmente entálpico), probablemente porque la conformación tipo PPII ya esté estabilizada previa a la interacción con RbAB, como fue demostrado para los péptidos fosforilados HPV16-E7-PP, que al estado libre, presentaron una propensión a estructura PPII incrementada.

***En conclusión,** la fosforilación de S132 en el CR2 de AdE1A produce un incremento en la afinidad de interacción con RbAB, generando un complejo que es energéticamente más estable que su contraparte no fosforilada. Esto podría deberse a que con dicha fosforilación se produciría la estabilización de estructura secundaria no canónica PPII (o pre-estructuración) en la región acídica conducente a una optimización en la interacción entre AdE1A y RbAB, reduciendo la penalidad entrópica que está asociada a un proceso plegamiento-acoplado-a-interacción.*

Conclusiones Parciales II.2.C.:

A este punto, tanto los resultados de EF como los de ITC, evidencian que la región “linker” en la proteína viral no aporta per se energéticamente de modo significativo a dicha interacción. Esto quiere decir que en su secuencia no habría otro sitio de interacción (o varios) que pueda/n mediar contactos con RbAB que sean estabilizantes del complejo. Con esto, la primer hipótesis (Modelo A, Figura 2.1 A) que se había planteado acerca del mecanismo de interacción alta afinidad entre la proteína viral y RbAB, queda descartada. No obstante, se ha demostrado que cuando el “linker” está conectando ambos motivos en AdE1A, ya no se observa una interacción a RbAB con afinidad moderada -como cuando cada motivo se une de manera individual-. En este caso los motivos se unen simultáneamente a RbAB con una afinidad global sorprendentemente alta (picomolar) generando un complejo energéticamente muy estable. Esto llevó a preguntarse: ¿puede ser que la primer interacción de cualquiera de los dos motivos de AdE1A con su sitio complementario en RbAB, esté favoreciendo ampliamente la segunda interacción con esta proteína? Con esto se planteó la hipótesis del mecanismo de interacción mediado por modulación alostérica clásica (Modelo B Figura 2.1 B)

II.2.D. Análisis de posible efecto cooperativo alostérico “en trans” entre los motivos de interacción de AdE1A con RbAB.

Además de los datos termodinámicos obtenidos en los experimentos vistos anteriormente, evidencias previas habían propuesto un posible efecto cooperativo entre los motivos E2F y LxCxE en AdE1A al interactuar con RbAB [108]. En este proceso cooperativo de tipo alostérico clásico la interacción primaria con RbAB de uno de los dos motivos induciría una alteración conformacional en la proteína globular RbAB, promoviendo un aumento en la afinidad del segundo motivo de interacción por su sitio complementario en esta proteína.

Para evaluar si existe un comportamiento alostérico en *trans* entre el motivo E2F (AdE1ACR1) y el motivo LxCxE (AdE1ACR2) en la interacción de AdE1A con RbAB, se realizó una serie de experimentos de titulación de ITC, bajo la condición en la que la región “linker” no se encontrase uniendo ambos motivos entre sí. Por un lado, para evaluar el efecto alostérico de la interacción del motivo LxCxE sobre el segundo donde interactúa el motivo E2F viral, un complejo pre-formado [RbAB:AdE1ACR2] (30 μ M) fue titulado con cantidades crecientes de AdE1A-CR1 y con AdE1A- Δ L (motivo E2F + región “linker”). Se titula también con AdE1A- Δ L, para evaluar si el posible efecto alostérico puede influir en la interacción del “linker” a RbAB. Estas titulaciones fueron ensayadas a 10°C atendiendo al hecho de que 20°C, el calor intercambiado de reacción era bajo y había efectos de agregación mayores, haciendo más

difícil su análisis (Figura 2.15, Serie A). Por otro lado, para evaluar el efecto alostérico de la interacción del motivo E2F sobre el sitio donde interactúa el motivo LxCxE en una segunda serie de reacciones realizadas a 20°C, un complejo pre-saturado [RbAB:AdE1ACR1] y otro [RbAB:AdE1A-ΔL] (ambos a 30 μM), fueron titulados con cantidades crecientes de AdE1A-CR2. Se pre-satura RbAB con AdE1A-ΔL, para evaluar si es necesaria la presencia del “linker” para poder inducir algún efecto alostérico. (Figura 2.15, Serie B). Si existiese efecto alostérico alguno, se debería ver un cambio significativo en la afinidad (y la K_D) en la interacción del motivo complementario con RbAB, cuando RbAB está pre-saturada con el primer motivo de interacción de AdE1A.

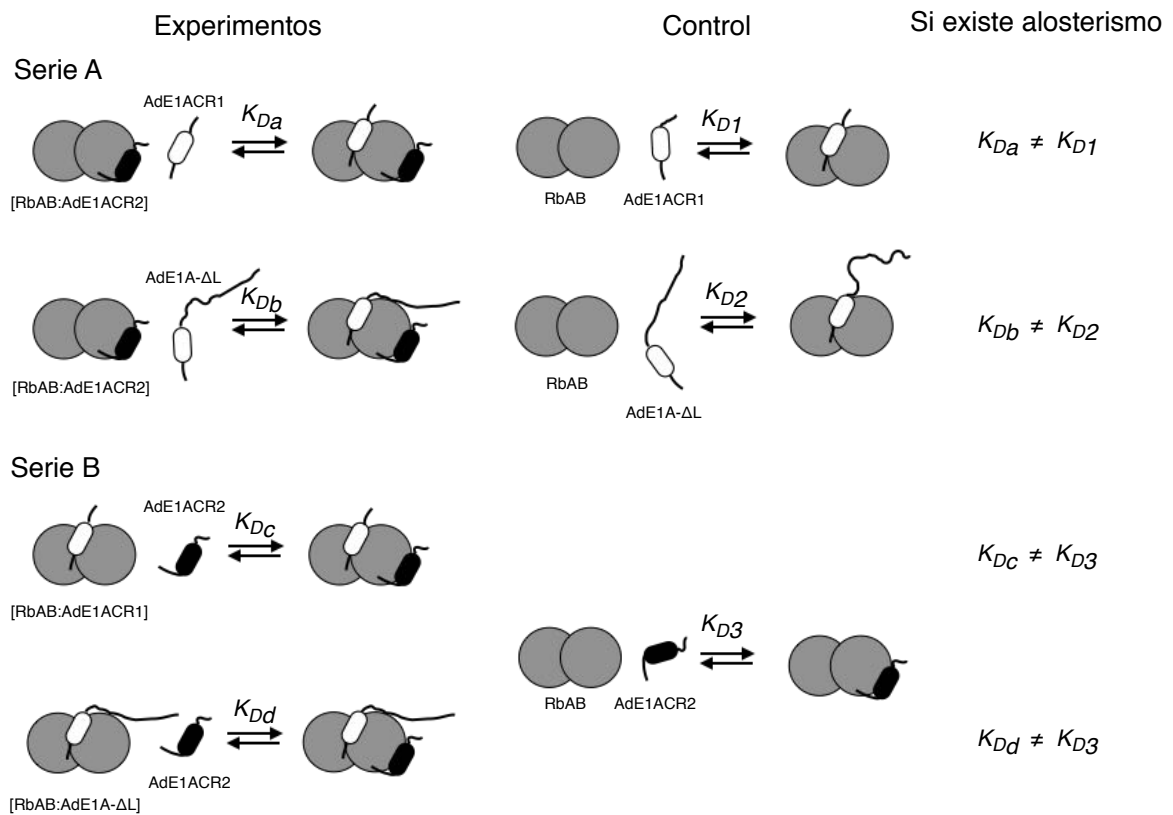


Figura 2.15: Esquema del experimento para probar posibles efectos alostéricos “en trans” entre los motivos de interacción de AdE1A con RbAB. La Serie A permite evaluar el efecto alostérico del motivo LxCxE (AdE1ACR2) sobre el segundo sitio donde interactúa el motivo E2F (AdE1ACR1 y AdE1A-ΔL). Se titula también con AdE1A-ΔL, para evaluar si el posible efecto alostérico puede influir en la interacción del “linker” a RbAB. La Serie B permite evaluar el efecto alostérico del motivo E2F (AdE1ACR1) sobre el segundo sitio donde interactúa el motivo LxCxE (AdE1ACR2). Se pre-satura RbAB con AdE1A-ΔL, para evaluar si es necesaria la presencia del “linker” para poder inducir algún efecto alostérico. Los experimentos sin pre-saturar a RbAB sirven de control. Existirían efectos alostéricos si las constantes de disociación son significativamente distintas de las K_D de los controles.

Tabla 2.10. Evaluación del efecto alostérico en “trans” entre los motivos de interacción de AdE1A con RbAB

Celda	Titulante	n	ΔH (cal/mol)	$-T\Delta S$ (cal/mol)	ΔG (cal/mol)	K_D (nM)
RbAB	AdE1ACR1	0.86 ± 0.02	7205 ± 338	-15005 ± 366	-7800 ± 139	1535 ± 392
RbAB + AdE1ACR2	AdE1ACR1	0.89 ± 0.01	6875 ± 214	-14779 ± 238	-7904 ± 72	1327 ± 159
RbAB	AdE1A- ΔL	0.91 ± 0.03	8675 ± 599	-16462 ± 623	-7787 ± 173	1590 ± 487
RbAB + AdE1ACR2	AdE1A- ΔL	0.86 ± 0.02	6795 ± 320	-14847 ± 353	-8051 ± 149	994 ± 252
RbAB	AdE1ACR2	1.03 ± 0.01	-9522 ± 79	573 ± 92	-8949 ± 47	213 ± 17
RbAB + AdE1ACR1	AdE1ACR2	1.19 ± 0.01	-8301 ± 89	-565 ± 110	-8867 ± 37	245 ± 27
RbAB + AdE1A- ΔL	AdE1ACR2	1.13 ± 0.01	-8873 ± 64	78 ± 76	-8795 ± 41	277 ± 20

Todas las mediciones fueron hechas por triplicado excepto por las titulaciones RbAB- AdE1ACR1, [RbAB:AdE1ACR2] - AdE1ACR1 y [RbAB:AdE1ACR2] - AdE1A- ΔL las cuales fueron hechas por duplicado.

Las mediciones fueron hechas a 10 °C para aquellas titulaciones en las que se usaron como agente titulante a AdE1ACR1 y AdE1A- ΔL . Para aquellas en las que se usó AdE1ACR2 como titulante, fueron hechas a 20 °C.

ΔH y K_D son valores experimentales y fueron obtenidos con errores luego de los ajustes. ΔG fue calculado como $\Delta G = RT \ln K_D$ y $-T\Delta S$ como $-T\Delta S = \Delta G - \Delta H$. Los errores de ΔG y $-T\Delta S$ fueron obtenidos por propagación.

La Tabla 2.10 muestra los resultados de esta serie de experimentos y en la Figura 2.16 se encuentran las isothermas representativas de los mismos. El resto de las isothermas se encuentran entre la Figuras A2.17 y Figura A2.20 del Anexo. Todas las interacciones ensayadas siguieron una estequiometría n=1, validando que no existen cambios en las estequiometrías de unión.

El péptido viral AdE1ACR1 se unió al complejo [RbAB:AdE1ACR2] en una reacción endotérmica a la temperatura ensayada (Figura 2.16 B). Como se puede observar, todos los parámetros termodinámicos son muy concordantes con los de la interacción entre AdE1ACR1 y RbAB sola. Esto indica que la unión del motivo mínimo LxCxE a RbAB no está produciendo cambio alguno que modifique la afinidad del segundo motivo (E2F en AdE1ACR1) por RbAB.

Similares resultados se obtuvieron tras titulación del complejo [RbAB:AdE1ACR2] con AdE1A- ΔL (Figura 2.16 D) la cual muestra el mismo perfil energético que la reacción entre AdE1A- ΔL y RbAB aislada (Figura 2.16 C, Tabla 2.10). La concordancia entre los parámetros termodinámicos de estas dos reacciones, también es indicativo que la unión del motivo LxCxE a RbAB no conduce a cambios que se traduzcan en mayor afinidad del motivo E2F por RbAB. Además, la región “*linker*” sería un anexo pasivo en la interacción con RbAB bajo estas condiciones ensayadas, ya que el perfil energético de las reacciones con AdE1A- ΔL no varían significativamente de aquellas en las que se utilizó AdE1ACR1 como titulante.

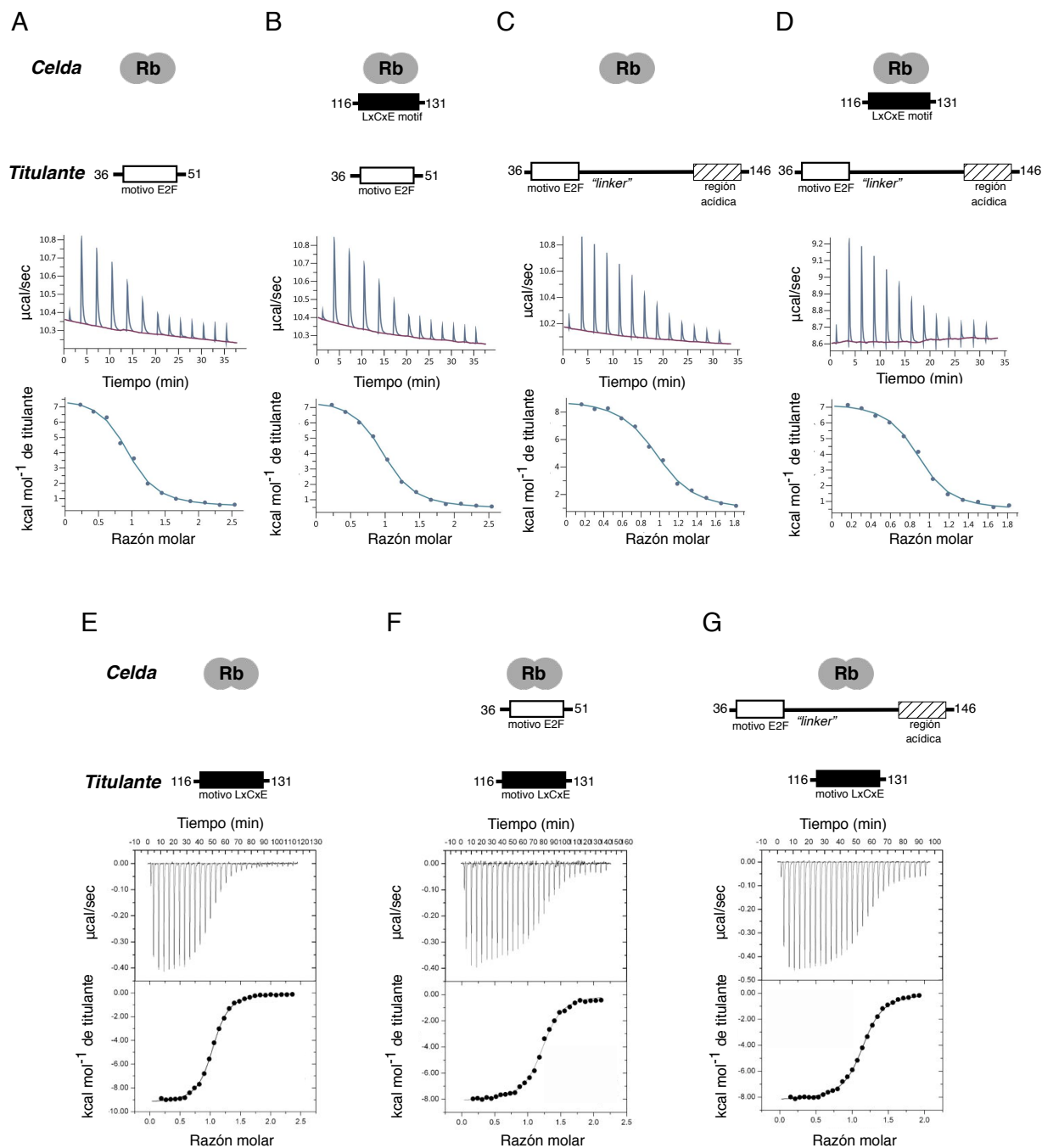


Figura 2.16: Curvas representativas de ITC para el análisis del efecto alostérico en *trans* entre los motivos de interacción de AdE1A con RbAB. Las mediciones fueron hechas en fosfato de sodio 20mM, 200mM NaCl, 5 mM 2-mercapto etanol a pH 7.0. La celda fue cargada con RbAB aislado o en complejo pre-ensamblado y la jeringa con los diferentes péptidos o proteínas complementarias. Curvas de asociación para **Serie A:** A) RbAB (30µM) y AdE1ACR1(300µM) a 10°C; B) [RbAB:AdE1ACR2] (30µM) y AdE1ACR1(300µM) a 10°C; C) RbAB (30µM) y AdE1A-ΔL(300µM) a 10°C; D) [RbAB:AdE1ACR2] (30µM) y AdE1A-ΔL (300µM) a 10°C; **Serie B:** E) RbAB (15µM) y AdE1ACR2 (150µM) a 20°C; F) [RbAB:AdE1ACR1] (15µM) y AdE1ACR2 (150µM) a 20°C; G) [RbAB:AdE1A-ΔL] (15µM) y AdE1ACR2 (150µM) a 20°C. Las curvas A y E son las mismas presentadas con anterioridad, aquí nuevamente representadas a los fines comparativos del experimento del efecto cooperativo alostérico. La curva C, es similar a la presentada en la sección anterior, pero esta fue determinada a 10°C.

En los ensayos complementarios, el péptido viral AdE1ACR2 se unió al complejo [RbAB:AdE1ACR1] con un perfil energético similar a las reacciones entre AdE1ACR2 y RbAB aislada (Figura 2.16 E) y entre AdE1ACR2 y [RbAB:AdE1A- Δ L] (Figura 2.16 G). En sumatoria, estos resultados indican que la unión primaria del motivo mínimo E2F (ni aun en presencia del “linker” en el complejo pre-saturado) a RbAB no produce cambio alguno que se asocie a una modificación en la afinidad del segundo motivo (LxCxE en AdE1A) por la proteína RbAB.

Conclusiones Parciales II.2.D.

La ausencia de cambios en la afinidad y energías asociadas a la interacción de un motivo viral de AdE1A con RbAB, estando unido a esta misma proteína el segundo motivo complementario de AdE1A (presente en otra molécula o en trans), indica que no hay efecto cooperativo alostérico en trans entre dichos sitios de interacción. De este modo, el Modelo B de interacción de mecanismo de interacción entre AdE1A y RbAB quedó también descartada. La tercer hipótesis (Modelo C, Figura 2.1 C), en la que el “linker” estaría jugando un rol crucial al aumentar la concentración efectiva del segundo motivo de interacción en las regiones cercanas a su sitio complementario en RbAB una vez unido el primer motivo, parece ser la más plausible y tras considerar los aportes de RMN, fue tratada en el último capítulo.

Hasta aquí se hicieron descripciones cuantitativas acerca del mecanismo de interacción de AdE1A con RbAB. No obstante a esta interacción bipartita, AdE1A debe desplazar al factor celular E2F del complejo endógeno que forma con RbAB, para que de esta forma se produzca el reingreso al ciclo celular. Una vez evaluadas las contribuciones de cada segmento de AdE1A en la unión a RbAB surge la pregunta de *¿qué componentes en AdE1A son estrictamente necesarios para generar el desplazamiento efectivo de E2F celular de su complejo con RbAB? ¿cumple la región “linker” alguna función en este proceso?*

II.2.E. Evaluación de la capacidad efectiva de AdE1A y fragmentos para desplazar eficientemente E2F2 de su complejo con RbAB. Ensayos de Competición por EF

Las evidencias estructurales muestran que el motivo E2F en AdE1A establece contactos con el sitio complementario a E2F en RbAB, compitiendo con E2F2 celular por la misma superficie, y pudiendo prevenir su interacción con RbAB [77, 84] (en Introducción, Figura I.11 C). Sin embargo las afinidades aquí reportadas, sugieren que el motivo E2F viral por sí sólo no

sería capaz de desplazar a E2F2 de su interacción con RbAB. En cambio AdE1A.WT, con una afinidad por RbAB 40 veces superior a la de E2F2, se esperaría que posea una capacidad efectiva para competir y desplazar a E2F2 del complejo [RbAB: E2F2].

Para evaluar estas hipótesis, se llevó a cabo una serie de experimentos de competición basados en EF. Estos experimentos se realizaron en las mismas condiciones que las titulaciones directas: tampón fosfato de sodio 20 mM, 200 mM NaCl, 2 mM DTT y 0.1% Tween-20, pH 7.0 y a 20 °C. Para ello se tituló un complejo pre-ensamblado [RbAB:E2F2-FITC] (en una relación molar 1:1) en concentración 5 μ M, con cantidades crecientes de: E2F2, AdE1ACR1, AdE1ACR2, AdE1.WT y AdE1A- Δ L, estos últimos sin derivatizar con FITC. La disrupción del complejo [RbAB:E2F2-FITC] y desplazamiento de E2F2-FITC fue seguida con el decrecimiento de la señal anisotrópica hasta que r alcanzó su valor correspondiente a la especie E2F2-FITC libre (Figura 2.17 A). Luego, los valores fueron normalizados y los IC_{50} fueron estimados directamente de las curvas.

La Figura 2.17 B, presenta las curvas obtenidas en el experimento de competición para cada especie competidora ensayada. Para todos los casos, los valores de IC_{50} estimados del experimento de competición, fueron alrededor de 20 a 60 veces más altos que los valores de K_D obtenidos de manera directa, lo que se debió a que la concentración complejo [RbAB: E2F2] trabajada (5 μ M) es completamente saturante (siendo la $K_D = 1$ nM), por lo que se encontraba muy estabilizado como para ser desplazado fácilmente.

Para AdE1ACR2 (motivo LxCxE) no hubo cambio en la señal anisotrópica con el agregado de concentraciones crecientes ya que no se une al mismo sitio de interacción que E2F2 en RbAB, por lo que incapaz de desplazar E2F2-FITC y constituyó un control negativo del experimento. E2F2 humano fue capaz de desplazar a su forma derivatizada presentando un IC_{50} de 30 nM, es decir 30 veces por encima de su K_D (1 nM).

AdE1A- Δ L y AdE1ACR1, que poseen el motivo E2F viral, mostraron un comportamiento similar entre sí y necesitaron de concentraciones 50 a 70 veces más altas que E2F2 para desplazar el 50% de E2F2-FITC de su interacción con RbAB ($IC_{50} = 1500$ y 2000 nM respectivamente). Esto es concordante con sus menores afinidades por RbAB, y se evidencian por las relaciones entre sus valores de K_D , las cuáles muestran ser alrededor de 100-200 veces mayor que la K_D de la interacción RbAB- E2F2 (Tabla 2.11). Con esto se demuestra por un lado, su poca eficiencia para desplazar E2F2 de su complejo con RbAB y por el otro, que el “*linker*” (que está presente en AdE1A- Δ L) no contribuye *per se* a dicho desplazamiento. Por su parte AdE1A.WT presentó una $IC_{50} = 1.5$ nM, y revela una habilidad de desplazamiento potenciada en unas 1000 veces respecto al motivo individual E2F viral (AdE1ACR1) (Tabla

2.11) para desplazar E2F2 del complejo con RbAB, lo que está en acuerdo con su medida directa de K_D de orden picomolar (Tabla 2.5).

A nivel técnico, los valores obtenidos en medidas con péptidos derivatizados con FITC (titulaciones directas por EF) y sin derivatizar (competición) y los cambios relativos entre K_D e IC_{50} (Tabla 2.11), son indicativos de que la presencia de la molécula de FITC no modifica sustancialmente sus propiedades de interacción con RbAB.

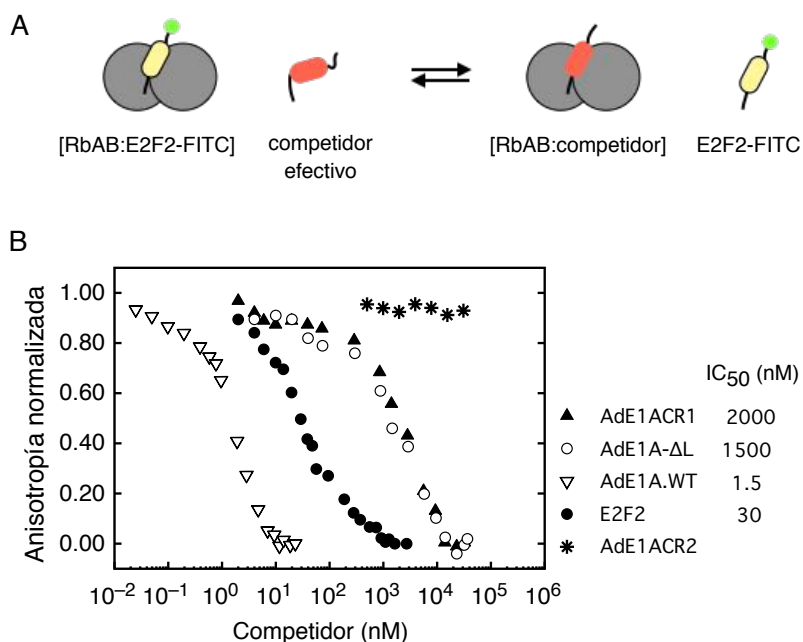


Figura 2.17: Experimento de competición de E2F2 unido a RbAB por diferentes péptidos y proteínas. A) Esquema del experimento de competición, donde E2F2-FITC en el complejo pre-formado con RbAB es competido por otro péptido o proteína. Si este competidor es efectivo, logra su desplazamiento. B) Curvas de competición con valores normalizadas de anisotropía para cada uno de los competidores. La eficiencia de competición de cada especie es distinta, lo que se evidencia por un rango dinámico alto de concentraciones (10^{-2} a 10^5 nM). La concentración del complejo [RbAB:E2F-FITC] trabajada fue de 5 μ M. Los valores IC_{50} estimados de las curvas para cada especie están a la derecha de la figura.

Tabla 2.11 Relaciones entre valores de K_D e IC_{50} para los péptidos y proteínas ensayados

	K_D / K_D	IC_{50} / IC_{50}
AdE1ACR1 / E2F2	120	70
AdE1A-ΔL / E2F2	200	50
AdE1A.WT / AdE1ACR1	0.00025	0.001
AdE1A.WT / E2F2	0.025	0.05

En la primera columna se presentan las razones de los valores de K_D de interacción con RbAB, hallados por EF en los experimentos de titulación directa, de las proteínas y péptidos ensayados. En la segunda columna, los valores de las razones entre las IC_{50} de este experimento de competición para las proteínas y péptidos ensayados. Como se puede observar hay un muy acuerdo entre los valores de estas razones, lo que valida las conclusiones del experimento.

En conclusión, estos resultados demuestran que el motivo E2F viral en AdE1A no es capaz de desplazar eficientemente a E2F2 humano de su interacción con RbAB. Tampoco la

presencia de la región "linker" contribuiría al motivo E2F a producir dicho desplazamiento. Para lograrlo, se requiere de la presencia e interacción simultánea a RbAB, del segundo motivo de interacción - LxCxE- estando conectados a través de la región "linker". En conclusión, AdE1A.WT es una proteína viral eficiente a la hora de desplazar a E2F2 de su complejo celular con RbAB, para lo cual requiere la integridad y trabajo conjunto de sus componentes: el motivo E2F, el motivo LxCxE y la región "linker".

CONCLUSIONES II

En este capítulo de la tesis se presentaron múltiples análisis cualitativo y cuantitativo/termodinámico tendientes a determinar un modelo mecanístico de la interacción entre AdE1A con RbAB y del desplazamiento efectivo de E2F2 de su complejo celular endógeno con RbAB.

Primero, la revisión bibliográfica de las interacciones mediadas por los motivos lineales E2F y LxCxE a miembros de la familia de proteínas “*pocket*”, mostraron que los residuos centrales conservados de estos SLiMs, son esenciales para la interacción dado que establecen contactos claves con residuos en la superficie de la proteína complementaria. No obstante a ello, residuos en el entorno inmediato pueden modular finamente la afinidad y especificidad de estas interacciones. Así, la secuencia aminoacídica que contextualiza un dado SLiMs, tendría un efecto significativo en sus preferencias de interacción y funcionalidad [193].

A continuación, la caracterización biofísica del complejo formado por AdE1A y RbAB, demostró estequiometría 1:1, compatible con una estructura en la que AdE1A se compacta sobre RbAB y se une a ella por sus extremos. La formación de este complejo no traería aparejado un cambio significativo en el contenido de estructura secundaria de AdE1A -evaluado por CD-, indicando que la proteína viral preserva su naturaleza desordenada tras la interacción.

En tercer lugar, con el análisis termodinámico en solución se observó que, en el caso de E2F2 celular, la región adicional N-terminal al motivo, lo dota de una afinidad a RbAB 100 veces superior respecto al motivo viral homólogo y su perfil energético revela una mayor contribución entálpica. Los motivos virales E2F y LxCxE en AdE1A se unen a RbAB, de modo independiente con afinidades de interacción moderadas y similares entre sí. En caso del motivo E2F, un proceso plegamiento-acoplado-a-interacción generaría complementariedad de superficies y se encuentra entrópicamente dirigida por un efecto de desolvatación que supera el costo entrópico del plegado. Para el motivo LxCxE, la interacción está entálpicamente dirigida a consecuencia de las interacciones hidrofóbicas y puentes de hidrógeno establecidos en la unión con RbAB. La presencia del parche ácido contiguo al motivo LxCxE viral, contribuye energéticamente de forma mínima pero suficiente para modular su afinidad de interacción con RbAB, como también ocurre para otras proteínas como HPV16 E7, HDAC 1 y RIZ al interactuar con RbAB [100, 109, 141, 150]. La fosforilación en S132 de AdE1A, tendría el mismo efecto, aportando a la eficiencia con la que el motivo LxCxE une a RbAB. De modo sorprendente AdE1A.WT, que posee ambos motivos de interacción unidos por la región “*linker*”, se une a RbAB con una afinidad muy alta (K_D del orden picomolar) 4000 veces superior a sus motivos

individuales y 40 veces superior a la interacción de E2F2 celular por RbAB, generando un complejo energéticamente muy estable.

El análisis termodinámico demostró que la presencia de la región “*linker*” en la proteína viral no estaría aportando *per se* significativamente a la interacción con RbAB, es decir, cuando no se encuentra uniendo los motivos E2F y LxCxE en AdE1A. No obstante su presencia es clave, dado el gran incremento en la afinidad por RbAB de la proteína AdE1A.WT, en comparación con los motivos individuales. Estos resultados rechazan el hipotético Modelo A, en el que interacción de alta afinidad entre AdE1A y RbAB estaría mediada por una estabilización producida por contactos directos del “*linker*”. Por otro lado, como el *alosterismo* clásico no logró explicar esta interacción de alta afinidad, el hipotético Modelo B, queda también descartado. Estos resultados dan lugar a ponderar la tercer hipótesis, el Modelo C, en el cual el “*linker*” en AdE1A jugaría un rol clave produciendo un aumento efectivo de la concentración del segundo motivo de interacción a RbAB sobre la región cercana a su sitio de interacción favoreciendo su unión, una vez que el primer motivo se encuentra unido (Figura 2.19).

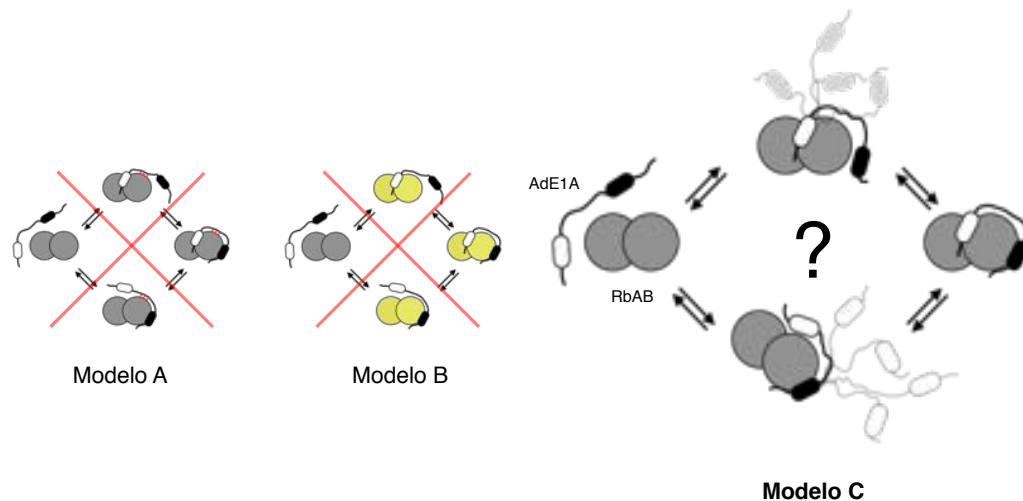


Figura 2.19. Modelos de interacción de alta afinidad entre AdE1A y RbAB. Los Modelos A y B fueron descartados, y el modelo C será evaluado más adelante.

Por último este trabajo reveló que el motivo E2F viral individualmente (y aún con la región “*linker*” presente) es incapaz de desplazar eficientemente al factor celular E2F2 de su complejo endógeno con RbAB, por un efecto de competencia directa. Sin embargo, cuando la región “*linker*” está uniendo ambos motivos (en AdE1A.WT), el efecto cooperativo permitiría a AdE1A desplazar fácilmente a E2F2, para dar lugar a un complejo ([RbAB:AdE1A.WT]) que es incluso energéticamente más estable que el complejo endógeno.

CAPITULO III

**Estudio estructural de alta resolución de AdE1A libre y en complejo con RbAB
por Resonancia Magnética Nuclear**

La caracterización estructural de AdE1A en alta resolución es de gran importancia pues proporciona información clave que es inaccesible por métodos biofísicos convencionales. Aunque algunos grupos de investigación han hecho avances al respecto, *la descripción estructural de AdE1A se encuentra aun lejos de estar completa*. A los fines de caracterizar la interacción de esta proteína viral con RbAB, faltan conocer detalles que permitan responder *¿cómo se caracterizan y comportan en forma aislada las diferentes regiones de AdE1A que están involucradas en esta interacción? Y ¿en qué forma dichas propiedades influyen en la interacción con RbAB?*

Por otro lado, también se hace necesaria una caracterización a nivel estructural de alta resolución del complejo entre RbAB y AdE1A, a fin de encontrar un modelo mecanístico que complemente la descripción cuantitativo/termodinámica hasta ahora lograda en este trabajo. Las pocas descripciones a nivel atómico de esta interacción hechas hasta el momento sólo han logrado retratar de manera parcial esta interacción, por lo que aquí *se buscará estudiarla desde un enfoque que considere todas las regiones en la proteína viral que estarían involucradas en la formación del complejo con RbAB*. Esta perspectiva de trabajo permitirá responder: *¿Qué regiones en AdE1A interactúan con RbAB y cuáles residuos median los contactos? ¿Qué residuos en los SLiMs de AdE1A contactan con RbAB? El “linker”, ¿media por caso algún contacto con RbAB? ¿Cuál es su estado conformacional tras formarse el complejo?*

Estos estudios se harán por Resonancia Magnética Nuclear (RMN), ya que esta metodología permite estudiar IDPs en solución, teniendo acceso a sus propiedades dinámicas y a la alta flexibilidad conformacional que las caracteriza. Con esto se tratará de comprender a nivel estructural los mecanismos que llevan a AdE1A a interactuar e inactivar a RbAB y que en consecuencia le permiten secuestrar y tomar control del ciclo celular.

OBJETIVOS ESPECIFICOS:

- I. Hacer un análisis estructural de alta resolución de la proteína viral AdE1A y las mutantes de motivos, complementando así sus estudios hechos por métodos biofísicos. Particularmente:
 - a. Asignar resonancias de los residuos del esqueleto carbonado de AdE1A.
 - b. Detectar elementos de estructura secundarios a nivel de residuos que pudieran estar pre-formados. Y analizar propensiones conformacionales de AdE1A en estado nativo libre en solución, sobre todo en las regiones que median contactos con RbAB.
 - c. Estudiar su dinámica conformacional y flexibilidad en el estado libre en solución

- d. Analizar si las mutantes introducen cambio alguno en la conformación nativa y en la dinámica de AdE1A.
- II. Analizar a nivel de detalles estructurales la interacción entre AdE1A y RbAB. Específicamente:
- Mapear los residuos en AdE1A involucrados en la interacción con RbAB, determinando si otros aparte de los ubicados en los motivos de interacción, se encuentran involucrados en dicha interacción.
 - Evaluar el rol de los residuos flanqueantes a los motivos E2F y LxCxE en la interacción con RbAB.
 - Analizar el estado conformacional de la región “*linker*” de AdE1A tras producirse la unión a RbAB.
 - Determinar los *regímenes de intercambio* de los residuos de AdE1A involucrados en la interacción con RbAB y así caracterizar las propiedades dinámicas de la interacción en estudio.

Los experimentos de RMN presentados en esta sección se realizaron en colaboración con el laboratorio del Dr. Gary W. Daughdrill, del Centro para el Descubrimiento de Drogas e Innovación (CDDI) de *University of South Florida* (USF) en la ciudad de Tampa, Florida, EEUU, como parte de un proyecto conjunto y de una beca Fulbright –Ministerio de Educación de la Nación, de la que fui beneficiario en el año 2017. El laboratorio del Dr. Daughdrill cuenta con dos equipos resonadores que fueron utilizados para realizar toda la serie de experimentos planteados en este capítulo. A continuación, se resumen los experimentos realizados en cada uno de ellos.

Tabla 3.1 Experimentos de RMN realizados en esta tesis.

		AdE1A.WT		AdE1A-ΔE		AdE1A-ΔL	
Estado de la proteína		¹⁵ N	¹⁵ N- ¹³ C	¹⁵ N	¹⁵ N- ¹³ C	¹⁵ N	¹⁵ N- ¹³ C
Varian 800 MHz	Libre	HSQC TROSY	HNCACB HNCO CBCACONH	HSQC TROSY	HNCACB HNCO CBCACONH	HSQC TROSY	HNCACB HNCO CBCACONH
	Unido a RbAB (1:1)	HSQC TROSY Titulación		TROSY		TROSY	
Varian 600 MHz	Libre	Relajación R1, R2, R1ρ, HN-NOE		Relajación R1, R2, R1ρ, HN-NOE		Relajación R1, R2, R1ρ, HN-NOE	

OBJETIVO III.1

Caracterización estructural de AdE1A libre

III.1.A. Evidencias de desorden intrínseco en AdE1A. Asignación de resonancias de AdE1A.WT, AdE1A- Δ E y AdE1A- Δ L.

Los experimentos de correlación heteronuclear bidimensionales constituyen uno de los pilares fundamentales en la investigación de proteínas por espectroscopía RMN, porque son fáciles de obtener y la buena dispersión de picos proveen un primer acercamiento acerca de las propiedades de estas biomoléculas. Particularmente el experimento bidimensional ^1H ^{15}N de Coherencia Cuántica Simple Heteronuclear ($^1\text{H}^{15}\text{N}$ -HSQC, por sus siglas en inglés) – el cual requiere solo muestras uniformemente marcadas con ^{15}N - es muy útil, porque muestra todas las correlaciones entre los núcleos ^1H y ^{15}N amida unidos covalentemente, proveyendo un pico por cada amida del esqueleto carbonado (excepto prolinas) e incluso algunos picos de cadenas laterales de grupos amida, indol y en ciertos casos guanidinio. Por lo tanto, es considerado una “huella digital” de la proteína [204]. Como las resonancias de estas amidas están afectadas por el entorno químico en el que se encuentran, permiten una rápida evaluación de si la proteína en estudio se encuentra plegada (proteína globular) o es una IDP. En el caso de proteínas globulares, el espectro HSQC presenta una serie de resonancias bien dispersas, particularmente en la dimensión ^1H , abarcando entre 3.5 a 4 ppm, ya que al estar estructurada, los entornos químicos de cada amida es particular, como lo muestra el ejemplo de la Ubiquitina humana en la Figura 3.1 A. En contraposición, las IDPs presentan un espectro pobremente disperso en la dimensión ^1H de 1 a 1.5 ppm, y un número importante de resonancias solapadas en la zona central del mismo, en consecuencia a su falta de estructura y alta flexibilidad que “asemejan” entre sí a los entornos químicos de las amidas, como se puede observar para la cola C-terminal de FCP1 -una proteína que regula la síntesis y procesamiento de mARNs- en la Figura 3.1 B [205, 206].

Cuando es necesaria información sitio-específica para profundizar el estudio de las proteínas, las resonancias observadas en un espectro como el HSQC deben ser asignadas. En este aspecto, la caracterización de IDPs puede ser desafiante porque la cadena polipeptídica es inherentemente flexible y rápidamente se interconvierte en múltiples conformaciones, complicando la asignación de resonancias sitio-específicas. No obstante, esto es posible en muchos casos, como en el de las proteínas globulares, llevarlo a cabo con experimentos de triple resonancia con muestras doblemente marcadas con isótopos ^{13}C y ^{15}N [207, 208]. Estos experimentos tridimensionales establecen conectividades secuenciales de resonancias de ^{13}C y

^{15}N bien resueltas en muestras uniformemente marcadas con ambos isótopos, y proveen de un método más general y robusto para obtener asignaciones de resonancia sin ambigüedades [209].

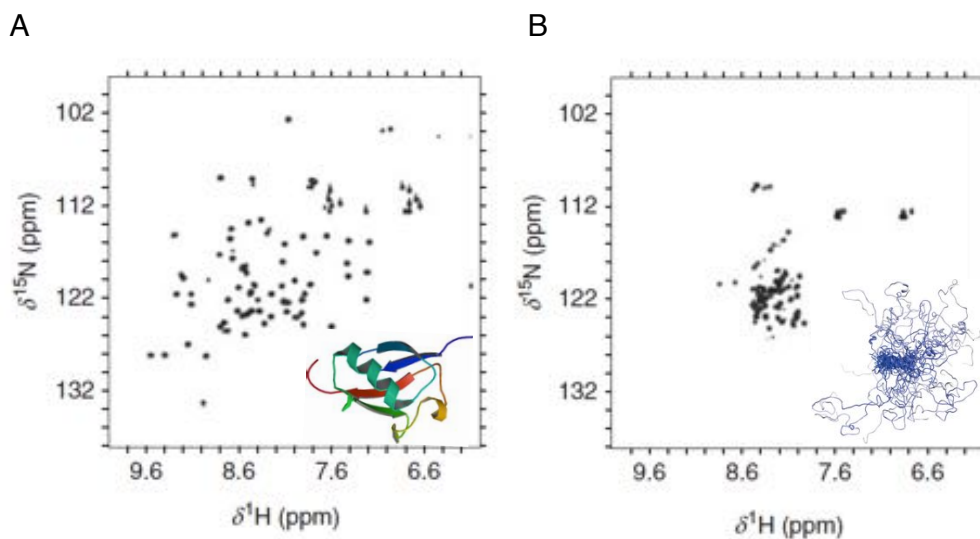
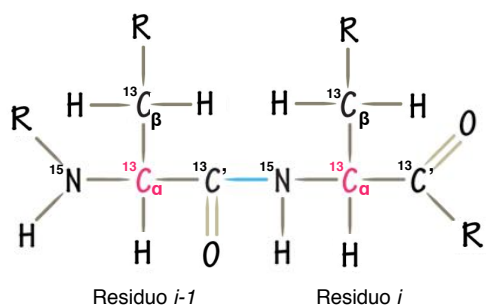


Figura 3.1 Espectros $^1\text{H}/^{15}\text{N}$ -HSQC de una proteína globular y una proteína desordenada. A) Espectro de la ubiquitina humana, mostrando su buena dispersión de picos. Abajo a la derecha, la representación de cintas obtenido de su estructura resuelta. B) Espectro de la cola C-terminal de FCP1, que muestra poca dispersión y solapamiento de picos. Abajo a la derecha, representación del ensamble conformacional de la cola y también el extremo N-terminal (que forma una hélice –en el medio) (Adaptado de [206]).

Con el objetivo de avanzar en la caracterización en alta resolución de las propiedades estructurales-conformacionales de AdE1A en estado libre en solución a nivel su análisis por RMN comenzó con la adquisición de espectros HSQC y experimentos tridimensionales para su asignación.

El trabajo en AdE1A comenzó con las asignaciones de las resonancias de sus núcleos ^1H amida, que están en abundancia natural y de los núcleos ^{13}C y ^{15}N amida para AdE1A.WT. Para esto, se expresó la proteína doblemente marcada con dichos isótopos (Ver Capítulo I y Materiales y Métodos). Para lograr la asignación, se realizaron adquisiciones de espectros trabajando con los siguientes experimentos tridimensionales -en el Varian 800MHz-:

- $^{15}\text{N}/^{13}\text{C}$ -HN(CA)CB, que permite la asignación de los $^{13}\text{C}_\alpha$ y $^{13}\text{C}_\beta$ de los residuos o aminoácidos “i” e “i-1”.
- $^{15}\text{N}/^{13}\text{C}$ -HNCO, que permite asignar $^{13}\text{C}'$ del residuo “i-1”.
- $^{15}\text{N}/^{13}\text{C}$ -CBCA(CO)NH, con el que se logra también asignar $^{13}\text{C}_\alpha$ y $^{13}\text{C}_\beta$ del aminoácido “i-1”.



Los espectros resultantes de los primeros tres experimentos fueron asignados de manera secuencial como se encuentran listados. En la Figura 3.2, se muestra un ejemplo de la asignación secuencial, que puede describirse como sigue: en principio, trabajando sobre el espectro $^{15}\text{N}^{13}\text{C}\text{-HN}(\text{CA})\text{CB}$, representado en las dimensiones ^1H (en ppm, eje X) vs ^{13}C (en ppm, eje Y), se asignaron los $^{13}\text{C}_\alpha$ (en azul) y $^{13}\text{C}_\beta$ (en rojo) del aminoácido “i”, a partir de valores tabulados de desplazamientos químicos (δ) de referencia promedios \pm desviaciones estándar (en ppm), los cuáles son característicos para cada núcleo según el tipo de aminoácido y asumen valores cercanos a la media en IDPs [210] (Figura 3.2 A). Naturalmente el trabajo de asignación comenzó por una glicina (G77 de la secuencia de AdE1A.WT), ya que éstos residuos son más fácilmente identificables al poseer únicamente el núcleo $^{13}\text{C}_\alpha$. Los $^{13}\text{C}_\alpha$ y $^{13}\text{C}_\beta$ del aminoácido “i-1” se identificaron como resonancias de menor intensidad compartiendo con el aminoácido “i” el mismo valor en la dimensión ^1H . Los δ de $^{13}\text{C}_\alpha$ y $^{13}\text{C}_\beta$ son invariables, por lo que el aminoácido “i-1” identificado en el paso anterior (E76), pudo ser posicionado como aminoácido “i”, al moverse en las dimensiones ^1H y ^{15}N (tercera dimensión) a otras zonas del espectro. Esto último es importante pues los δ de los núcleos ^1H y ^{15}N se reconocen como tales sólo cuando el aminoácido es “i” y no “i-1”, permitiendo así las asignaciones de dichos núcleos. A su vez, esto permitió identificar núcleos $^{13}\text{C}_\alpha$ y $^{13}\text{C}_\beta$ de un nuevo “i-1”. A continuación, los valores de δ de los núcleos previamente asignados (para la G77) fueron utilizados para posicionarse en el espectro $^{15}\text{N}^{13}\text{C}\text{-HNCO}$ -representado también en las dimensiones ^1H vs ^{13}C -, permitiendo la asignación del carbono carbonílico ($^{13}\text{C}'$) del aminoácido “i-1” (E76), los cuáles también asumen valores de δ característicos según el tipo de aminoácido (Figura 3.2 B). Idéntico procedimiento se realizó sobre el espectro $^{15}\text{N}^{13}\text{C}\text{-CBCA}(\text{CO})\text{NH}$ que permitió asignar nuevamente $^{13}\text{C}_\alpha$ y $^{13}\text{C}_\beta$ del aminoácido “i-1” (E76) (Figura 3.2 C).

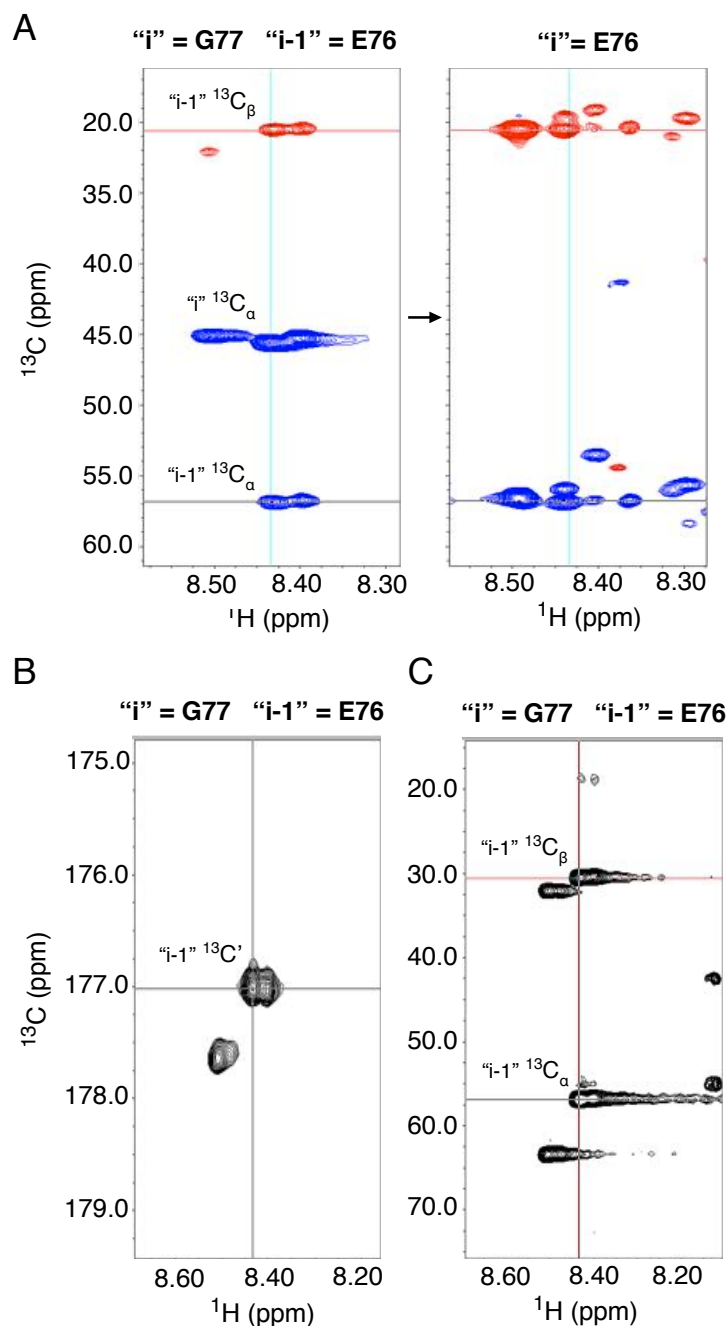


Figura 3.2 Espectros de los experimentos tridimensionales para llevar a cabo la asignación secuencial de residuos en la proteína AdE1A.WT. Los mismos fueron ubicados en el orden en el que fueron analizados. A) Espectro HNCACB, con el $^{13}\text{C}_\alpha$ identificado para el aminoácido "i" G77, y los $^{13}\text{C}_\alpha$ y $^{13}\text{C}_\beta$ del aminoácido "i-1" E76. B) Espectro HNCO, con el $^{13}\text{C}'$ identificado del aminoácido "i-1" E76. C) Espectro CBCACONH con los $^{13}\text{C}_\alpha$ y $^{13}\text{C}_\beta$ del aminoácido "i-1" E76 identificados.

La conectividad entre estos tres experimentos que fueron realizados sobre AdE1A.WT y analizados secuencialmente, redujeron las incongruencias al momento de realizar las asignaciones, al proveer información que es complementaria y solapante permitiendo asignar para esta proteína de 114 residuos, el 90% de los núcleos $^{13}\text{C}_\alpha$, el 86% de los $^{13}\text{C}_\beta$, el 75% de los $^{13}\text{C}'$, el 75% de los ^1H amida y el 75% de los ^{15}N amida. Los valores experimentales de los δ

para cada uno de estos núcleos en la secuencia de AdE1A.WT están tabulados en la Tabla A3.1 de Anexos.

Por otro lado, se realizó el experimento bidimensional $^1\text{H}^{15}\text{N}$ -HSQC de AdE1A.WT. El espectro fue adquirido en muestras uniformemente marcadas con ^{15}N a concentraciones de proteína de 300, 420 y 700 μM . El espectro $^1\text{H}^{15}\text{N}$ -HSQC (Figura 3.3), caracterizado por una baja dispersión de resonancias en la dimensión ^1H , presentó alta calidad, en gran medida porque fue tomado en alta sensibilidad (800 MHz) y porque la muestra con la que se lo adquirió estaba altamente concentrada (700 μM), dando una alta relación señal/ruido.

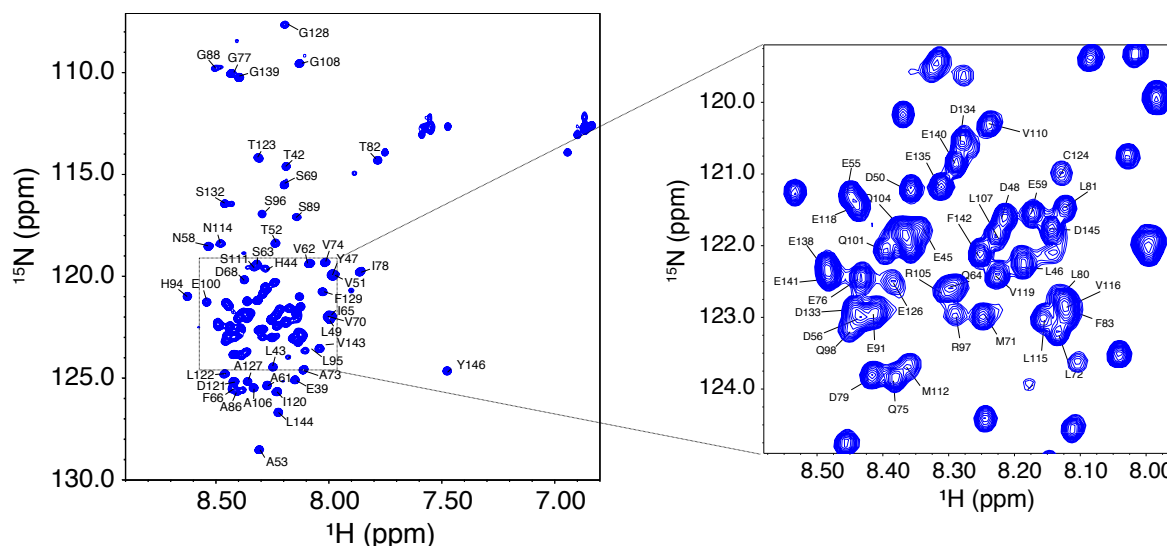


Figura 3.3. Espectro HSQC de AdE1A.WT. El espectro HSQC fue obtenido para una muestra uniformemente marcada con ^{15}N a 700 μM en un equipo Varian 800 MHz a 25°C. Sus correspondientes asignaciones se encuentran señaladas. A la derecha, detalle de la región central del espectro con mayor densidad de picos y sus asignaciones.

Las picos en el espectro $^1\text{H}^{15}\text{N}$ HSQC colectado para AdE1A.WT fueron asignados utilizando los valores de δ de los núcleos ^1H y ^{15}N amida asignados previamente en los experimentos 3D. La Figura 3.3, también muestra el resultado de la asignación y un detalle de la misma de los picos de la región central de dicho espectro. En total, se logró asignar 85 de 114 aminoácidos contenidos en la secuencia de trabajo de AdE1A.WT. Los restantes residuos en la secuencia no fueron asignados en el espectro $^1\text{H}^{15}\text{N}$ HSQC por las siguientes razones:

- a. Prolinas. Todos los aminoácidos poseen un ^1H unido al ^{15}N amida y como consecuencia presentan una correlación ^1H - ^{15}N en el espectro $^1\text{H}^{15}\text{N}$ HSQC que proyecta sus correspondientes picos. La excepción la constituyen las prolinas, que por ser cíclica, su ^{15}N amida no está unida a un ^1H , y por lo tanto no presenta picos en el espectro $^1\text{H}^{15}\text{N}$ HSQC. Además, en la serie de experimentos tridimensionales

estándar realizados, las prolinas no pueden ser determinadas, ya que no aparecen como aminoácidos “i” por la misma razón de no poseer un ^1H unido al ^{15}N amida, no pudiendo asignarse sus núcleos ^{15}N correspondientes. Las Prolinas no asignadas son 18 en la secuencia de AdE1A y se listan a continuación: P40, P41, P54, P57, P67, P84, P85, P87, P90, P92, P93, P99, P102, P109, P113, P117, P130 y P131.

- b. Otros residuos. Como es común para los residuos en el extremo N-terminal de una secuencia, la resonancia o picos de los primeros aminoácidos no son observables debido a intercambios en el solvente [211], en este caso no pudieron asignarse los residuos G(-2), S(-1) y M(0), que no corresponden a la secuencia natural de AdE1A y S36, H37, F38 de la secuencia de AdE1A. Finalmente, algunos residuos no pudieron ser asignados porque en los experimentos tridimensionales, sus $^{13}\text{C}_\alpha$ y $^{13}\text{C}_\beta$ cuando fueron analizados como aminoácidos “i”, no pudieron ser claramente discernidos de otros con δ muy similares, por lo que los δ de sus núcleos ^1H y ^{15}N no fueron asignados. Éstas fueron 6 residuos: E60, E103, H125, D136 y E137.

Idénticos procedimientos fueron aplicados para lograr la asignación de los núcleos $^{13}\text{C}_\alpha$, $^{13}\text{C}_\beta$, $^{13}\text{C}'$, ^1H amida y ^{15}N amida, de las dos especies mutantes AdE1A- ΔE y AdE1A- ΔL . Los δ experimentales de estos núcleos están tabulados en las Tabla A3.2 y A3.3 de Anexos. La Tabla 3.2, resume el porcentaje de núcleos asignados en ambas mutantes.

Al igual que para AdE1A.WT, los valores de δ de los núcleos ^1H y ^{15}N , fueron utilizados para asignar las resonancias o picos en el espectro $^1\text{H}^{15}\text{N}$ HSQC colectado para cada mutante. Para AdE1A- ΔE se logró la asignación de 81 de 113 aminoácidos contenidos en su secuencia (Figura 3.4 A). Los restantes residuos no se asignaron en el espectro $^1\text{H}^{15}\text{N}$ HSQC por las mismas razones por la que varios residuos en AdEA1.WT no fueron asignados. Ellos fueron:

- a. Prolinas: P40, P41, P54, P57, P67, P84, P85, P87, P90, P92, P93, P99, P102, P109, P113, P117, P130 y P131.
- b. Otros residuos: G(-2), S(-1) y M(0), que no corresponden a la secuencia natural de AdE1A y S36, H37 y F38 de la secuencia, ubicados en el extremo N-terminal. Los residuos: A44, A46, A47 de la porción mutada y los residuos E60, E103, H125, D136 y E137, cuyos $^{13}\text{C}_\alpha$ y $^{13}\text{C}_\beta$ fueron indistinguibles de otros con δ muy similares en los experimentos 3D.

En el caso de AdE1A- ΔL se logró la asignación de 81 de 113 aminoácidos contenidos en su secuencia (Figura 3.4 B). Los residuos no asignados fueron:

- Prolinas: P40, P41, P54, P57, P67, P84, P85, P87, P90, P92, P93, P99, P102, P109, P113, P117, P130 y P131.
- Otros residuos. G(-2), S(-1) y M(0), que no corresponden a la secuencia natural de AdE1A y S36, H37 y F38 de la secuencia de AdE1A, ubicados en el extremo N-terminal. Otros residuos cuyos $^{13}\text{C}_\alpha$ y $^{13}\text{C}_\beta$ estaban solapados con otros: A123, A124 y A126 de la porción mutada y los residuos E60, E103, E118, D136 y E137.

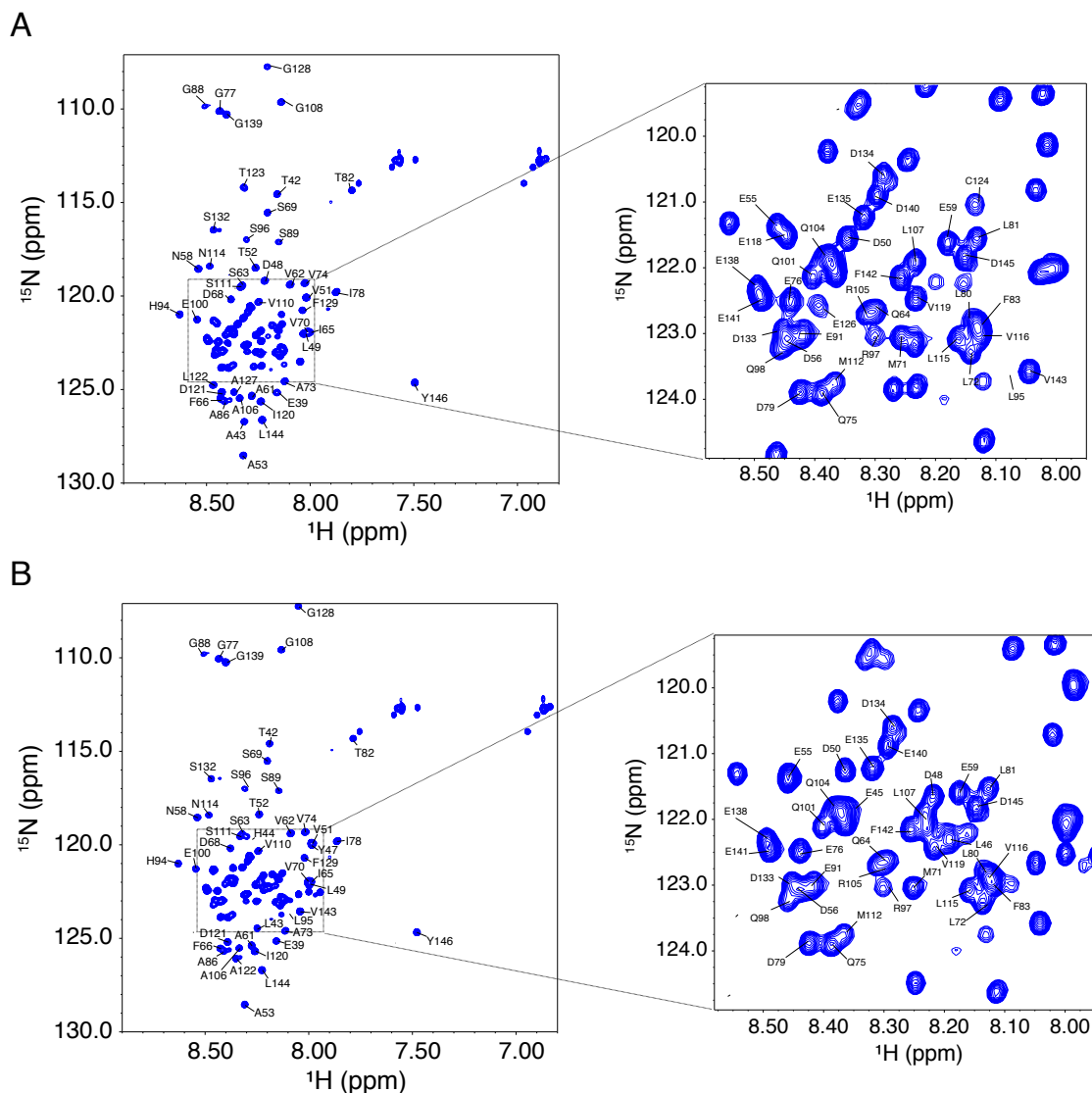


Figura 3.4. Espectro HSQC de variantes mutantes de AdE1A. A) Espectro HSQC de AdE1A- Δ E obtenido para una muestra uniformemente marcada con ^{15}N a $975\ \mu\text{M}$ y sus correspondientes asignaciones. A la derecha, detalle de la región central del espectro con mayor densidad de picos y sus asignaciones. B) Espectro HSQC de AdE1A- Δ L obtenido para una muestra uniformemente marcada con ^{15}N a $850\ \mu\text{M}$ y sus correspondientes asignaciones. A la derecha, detalle de la región central del espectro con mayor densidad de picos y sus asignaciones. Ambos adquiridos a 25°C en un equipo Varian 800 MHz.

Tabla 3.2. Porcentaje de núcleos asignados en las variantes mutantes de AdE1A

	% de núcleos asignados en los experimentos 3D					Residuos asignados en el HSQC	Total de residuos
	¹³ C _α	¹³ C _β	¹³ C'	¹⁵ N	¹ H		
AdE1A.WT	90	88	75	75	75	85	114
AdE1A-ΔE	88	85	73	73	73	81	113
AdE1A-ΔL	88	85	73	73	73	81	113

En resumen, tanto para la variante natural de AdE1A, como para las mutantes, los espectros HSQC obtenidos presentaron una dispersión de las resonancias en un rango no mayor a 1.5 ppm en la dimensión ¹H, y muchos picos se encontraron solapados en la región central del espectro. Estas características en el espectro HSQC son típicas de proteínas intrínsecamente desordenadas, y constituyen evidencia estructural de alta resolución que están en acuerdo con las caracterizaciones biofísicas previas logradas para estas proteínas. Asimismo, y a pesar de la acumulación de señales en la zona central del espectro, la asignación de las resonancias alcanzadas para las secuencias natural y mutantes de AdE1A fueron en todos los casos superiores al 70%, fundamental para lograr interpretaciones posteriores realizadas para los experimentos de interacción con RbAB.

III.1.B. Identificación de elementos de estructura secundaria en AdE1A

Para muchos casos de IDPs la presencia de estructuras secundarias estables está condicionada a la formación de un complejo tras producirse la interacción, ya que al estado libre de la proteína estas estructuras son transitorias y se interconvierten rápidamente en un número elevado de conformaciones. No obstante, la caracterización por RMN, debido a su alta resolución permite describir algunos elementos de estructura secundaria que pueden encontrarse pre-estructurados en las formas libres de las IDPs. Como ejemplo, la región del dominio N-TAD de p53 que interactúa con Mdm2 (discutido en Introducción) se predice desordenada, pero con experimentos de RMN: datos de desplazamientos químicos ($\Delta\delta$), experimentos HNOE y distribución de probabilidad de poblaciones, se ha logrado identificar una pequeña α -hélice pre-estructurada en dicha región [23, 48](Figura 3.5 A).

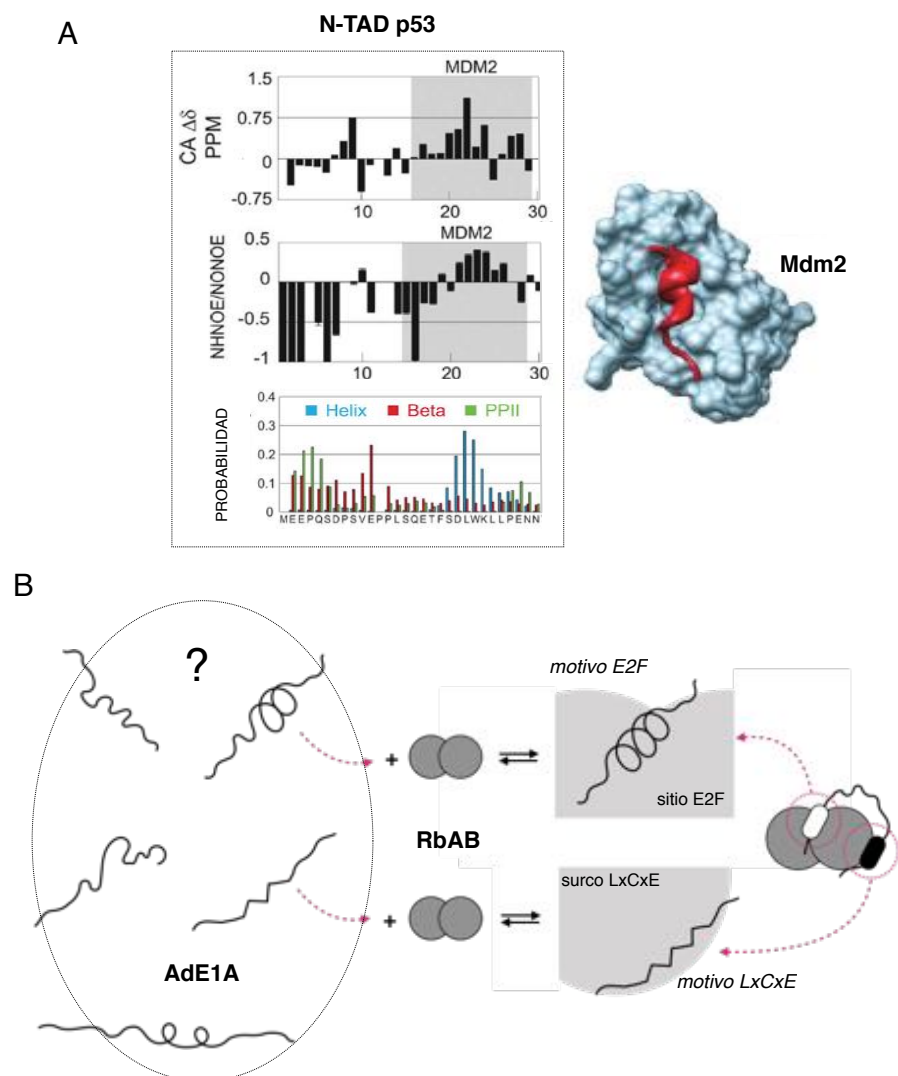


Figura 3.5. Presencia de elementos de estructura secundaria en IDPs evaluados por RMN. A) Análisis por RMN de elementos de estructura secundaria pre-formados en el N-TAD de p53 en el estado libre en solución, los cuáles revelan que los residuos de la región que se une a Mdm2 (residuos 20-25) se encuentra pre-estructurados en α -hélice. Se muestran valores por residuos de $^{13}\text{C}_\alpha\Delta\delta$, HNOE y tendencia estructural ($\approx 30\%$ α -hélice) (Adaptado de [48]). A la derecha la estructura cristalina de la región N-TAD de p53 unida a Mdm2 (PDB ID: 1YCR) (Adaptado de [23]). B) Posibles escenarios para distintos fragmentos de AdE1A que podrían poblar estructuras secundarias canónicas y no canónicas al estado libre, como por ejemplo los motivos LxCxE y E2F, los cuáles se presentan estructurados en su forma unida.

De la misma manera que N-TAD p53, determinadas regiones en AdE1A podrían presentar estructuras secundarias o estar parcialmente estructuradas en el estado libre, particularmente sobre los residuos de los motivos E2F y LxCxE -de los que existen evidencias de orden en su estado unido- como así también en otras partes de la extensión del “*linker*”. Esto es útil determinar para caracterizar el contenido mismo de estructura en AdE1A, pero también para entender mecanismos de interacción sitio-específicos (Modelo de *selección conformacional* vs modelo de *ajuste inducido*, Figura 3.5 B), y echar luz a si los costos

entrópicos asociados a plegamiento -como en el caso del motivo E2F- es menor por estar preformada una hélice y por tanto si está más o menos favorecida energéticamente la interacción con RbAB. Experimentos como el CD tienen baja resolución para proveer de información útil en estos casos, pero el RMN es especialmente ventajoso en ello.

Con el objetivo de identificar la presencia de estructuras secundarias locales a nivel de residuos en la proteína AdE1A en estado libre en solución, se procedieron a realizar los siguientes análisis por RMN:

III.1.B.1. Análisis de los desplazamientos químicos secundarios de los residuos en AdE1A en estado libre

Los cambios o desviaciones en los desplazamientos químicos de los núcleos -obtenidos experimentalmente- con respecto a valores de referencia estándar tabulados para aminoácidos en conformación “*random coil*”, definen formalmente a los desplazamientos químicos secundarios ($\Delta\delta$). Estos $\Delta\delta$ contienen información apreciable sobre la estructura y dinámica de las proteínas [212] y si bien no son suficientes para definir estructuras terciarias en una proteína, indican la tendencia relativa de la cadena polipeptídica a adoptar estructuras secundarias de α -hélice, lámina- β u otras más extendidas en diferentes regiones de la secuencia primaria [213]. A partir de estos valores de $\Delta\delta$, la topología de la proteína puede ser inferida y al contener información a nivel residuo-específico, es diferente y a la vez complementaria a otras técnicas para determinar estructura secundaria, como el Dicroísmo Circular, Espectroscopía FTIR y UV-Raman. No obstante, la sensibilidad de estos $\Delta\delta$ es tal, que el contenido porcentual local de estructuras secundarias canónicas transientes como α -hélice o lámina- β pueden ser cuantificadas. Es importante tener en cuenta que como los $\Delta\delta$ pueden ser muy pequeños para las proteínas desestructuradas, es necesario utilizar desplazamientos químicos “*random coil*” de referencia que sean adecuados [214]. Además, al momento de analizar estos datos se debe considerar que al ser las IDPs altamente dinámicas, parámetros como $\Delta\delta$, reportan sobre la población media y ponderada de todas las estructuras en el ensamble conformacional para cada residuo [208] y deben ser interpretadas en conjunto con otros parámetros como en el ejemplo presentado de p53.

El análisis de los $\Delta\delta$ residuo-específicos de los núcleos $^{13}\text{C}_\alpha$, $^{13}\text{C}_\beta$, $^{13}\text{C}'$, ^1H amida y ^{15}N amida previamente asignados de AdE1A.WT, se hizo en base a valores de referencia de δ compilados especialmente para secuencias IDPs en una biblioteca de desplazamientos químicos corregidos por el efecto de los residuos colidantes, llamado nclDP (*neighbour-corrected chemical shift library for IDPs*, por sus siglas en inglés), el cual posibilita un detección

más precisa de estructuras secundarias a nivel local [215]. Como los valores de δ de referencia para un aminoácido “x” no son fijos, sino que varían en función de los residuos “x-1” y “x+1”, primero fueron calculados para la secuencia aminoacídica particular con la que se trabajó. Para ello se usó una herramienta informática disponible en <http://nmr.chem.rug.nl/ncIDP/> la cual, una vez alimentada con la secuencia en formato FASTA de AdE1A.WT arrojó como resultados sus valores de δ de referencia.

Los desplazamientos químicos secundarios para AdE1A.WT fueron entonces calculados por la simple resta: $\delta_{\text{experimental}} - \delta_{\text{referencia}}$ para sus núcleos asignados $^{13}\text{C}_\alpha$, $^{13}\text{C}_\beta$, $^{13}\text{C}'$, ^1H amida y ^{15}N amida y los valores resultantes: $^{13}\text{C}_\alpha\Delta\delta$, $^{13}\text{C}_\beta\Delta\delta$, $^{13}\text{C}'\Delta\delta$, $^1\text{H}\Delta\delta$ y $^{15}\text{N}\Delta\delta$, están tabulados en la tabla A3.4 del Anexo.

La Figura 3.6 presenta los gráficos de $^{13}\text{C}_\alpha\Delta\delta$, $^{13}\text{C}_\beta\Delta\delta$, $^{13}\text{C}'\Delta\delta$, $^1\text{H}\Delta\delta$ y $^{15}\text{N}\Delta\delta$ en función de la secuencia aminoacídica para AdE1A.WT. Los desplazamientos químicos secundarios de los $^{13}\text{C}_\alpha$ y $^{13}\text{C}'$ pueden ser interpretados de manera conjunta, así valores de $\Delta\delta$ positivos para ambos núcleos indican una preferencia por estructura de tipo α -hélice, mientras que valores negativos son indicativos de una preferencia por estructuras tipo lámina β . Generalmente valores por encima de 2ppm en $^{13}\text{C}_\alpha\Delta\delta$, se considera en estructura α -hélice. La interpretación que se asume para los valores de $^{13}\text{C}_\alpha\Delta\delta$ y $^{13}\text{C}'\Delta\delta$ se invierte en el caso de los valores de $^1\text{H}\Delta\delta$, por lo que valores positivos indican preferencia por lámina β y valores negativos, una preferencia por estructura de tipo α helicoidal. En el caso de $^{13}\text{C}_\beta\Delta\delta$, valores positivos son indicativos de la presencia de lámina β en tanto que valores tendientes a cero indicarían presencia de estructuras tipo α -hélice. Valores negativos de $^{13}\text{C}_\beta\Delta\delta$ evidencian presencia de estructuras con ángulos ϕ y ψ positivos. La interpretación estructural de los $^{15}\text{N}\Delta\delta$ es menos clara y directa que la de los otros núcleos, en especial comparado con los $^{13}\text{C}_\alpha\Delta\delta$ y $^{13}\text{C}_\beta\Delta\delta$ [212, 216, 217].

Los valores de $^{13}\text{C}_\alpha\Delta\delta$, $^{13}\text{C}_\beta\Delta\delta$, $^{13}\text{C}'\Delta\delta$, $^1\text{H}\Delta\delta$ para AdE1A.WT muestran que esta proteína no presenta estructuras secundarias evidentes a lo largo de toda su extensión. Además, las variaciones de los valores de $^{13}\text{C}_\alpha\Delta\delta$ y $^{13}\text{C}'\Delta\delta$ en uno y otro sentido para la mayoría de los residuos y sus valores absolutos mas bien cercanos cero (o dicho de otra manera, cercanos a los valores medios “*random coil*” corregidos por secuencia) a lo largo de la proteína, sugieren que ésta es eminentemente desordenada en estado libre en solución, confirmando su estado “*random coil-like*”

Si bien existe evidencia experimental de que el motivo E2F en AdE1A está en conformación α -hélice en el estado unido a RbAB (Figura 2.3 D), en el estado libre dicha región

no muestra indicios de estar pre-estructurada con esta conformación. Para el caso del motivo LxCxE, -aunque no existe estructura cristalina de la interacción de este motivo en AdE1A con RbAB, sí lo hay del LxCxE homólogo de E7-HPV unido a RbAB, la cual muestra estar una conformación en lámina- β en su estado unido,- tampoco se observan desplazamientos químicos que indiquen pre-estructuración.

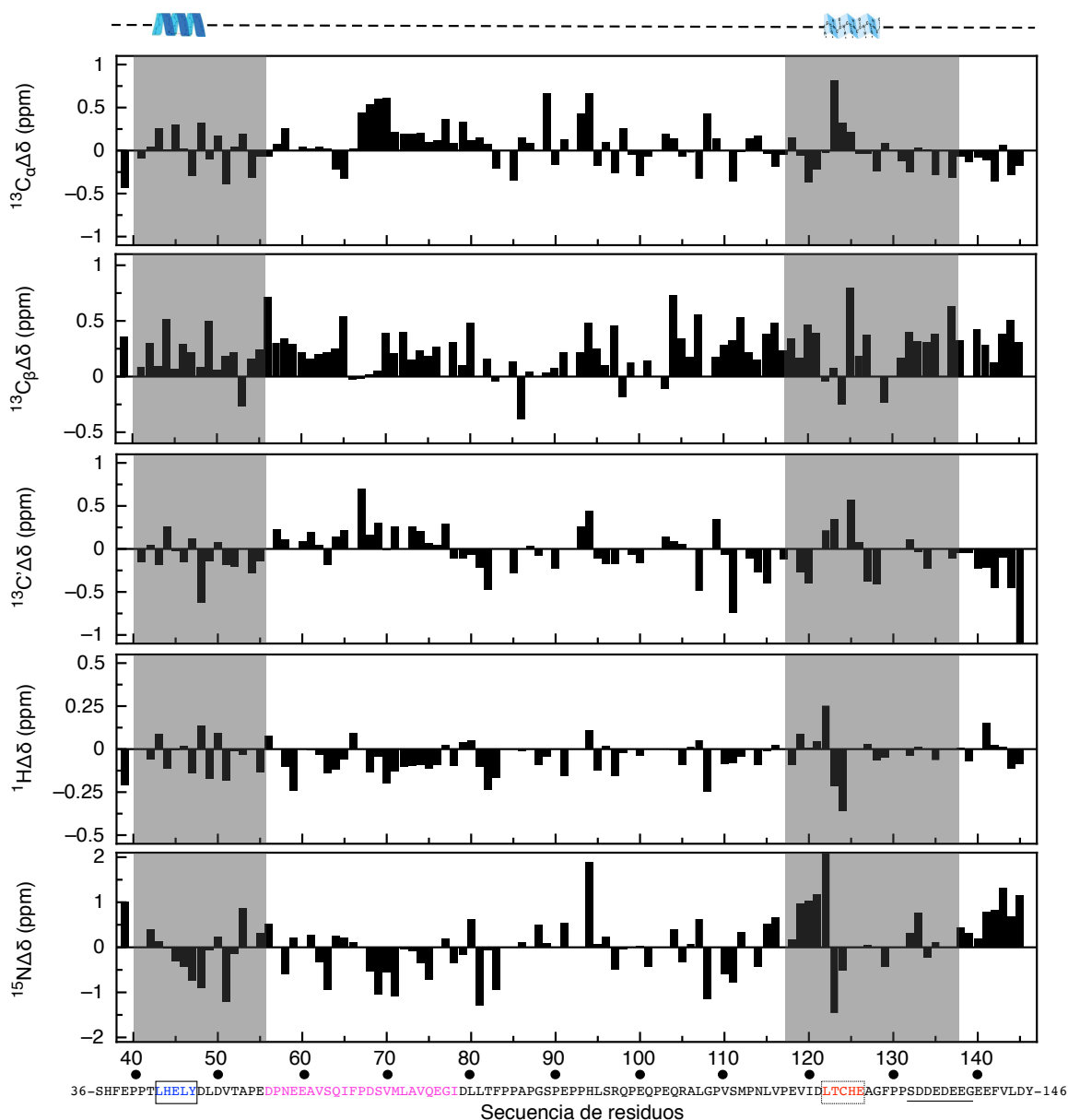


Figura 3.6: Cambio en los desplazamientos químicos de $^{13}\text{C}_\alpha$, $^{13}\text{C}_\beta$, $^{13}\text{C}'$, ^1H y ^{15}N por residuo para la proteína AdE1A.WT. Las regiones correspondientes a los motivos de interacciones y residuos flanqueantes están señalados en gris. Arriba, esquema de la conformación que AdE1A adquiere en su forma unida a RbAB, dónde el motivo E2F forma una α -hélice al unirse al sitio complementario de E2F y el motivo LxCxE, aparece plegado en lámina- β acorde al plegamiento del motivo homólogo en E7-HPV unido a RbAB (remitirse a las Figuras 2.3 D y A, respectivamente). Abajo, la secuencia de AdE1A.WT y señalados el motivo E2F (azul), el motivo TRAM-CBP (en magenta) y el motivo LxCxE (en rojo), la región ácida contigua se encuentra subrayada.

Por otro lado, se puede notar una pequeña región que parece mostrar una mínima conformación helicoidal al observar valores consecutivos positivos de $^{13}\text{C}_\alpha\Delta\delta$ entre los aminoácidos 67 a 82 y que solapa en parte con el motivo CBP-TRAM (residuos 56 a 78). Se ha visto que este motivo estructura α -hélice al unirse a los dominios TAZ-2 y NCBP de los activadores co-transcripcionales CBP/p300 [76, 82] (Figura I.11 D). No obstante a ello, la relevancia de estos valores locales de $^{13}\text{C}_\alpha\Delta\delta$ es difícil de estimar al no haber una buena correlación con las variaciones esperadas en los $^{13}\text{C}_\beta\Delta\delta$, $^{13}\text{C}'\Delta\delta$, $^1\text{H}\Delta\delta$ para un segmento que se espera tenga estructura tipo α -hélice o tendencia a ella, como sí lo muestra la región de unión a Mdm2 en p53 en el estado libre (Figura 3.5). Sin embargo, la falta de correlación entre los valores de $\Delta\delta$ para los diferentes núcleos no es algo inusual para IDPs [218], por lo que en este caso llegar a una interpretación sin ambigüedades puede resultar más difícil.

Los desplazamientos químicos secundarios para AdE1A- Δ E y AdE1A- Δ L fueron hallados siguiendo el mismo procedimiento que para AdE1A.WT: primero se calcularon sus valores de δ de referencia estándar en función de sus propias secuencias y luego se restaron los valores de δ experimentales a los de referencia. Los $^{13}\text{C}_\alpha\Delta\delta$, $^{13}\text{C}_\beta\Delta\delta$, $^{13}\text{C}'\Delta\delta$, $^1\text{H}\Delta\delta$ y $^{15}\text{N}\Delta\delta$ de ambas mutantes están tabulados en la tabla A3.5 (para AdE1A- Δ E) y A3.6 (para AdE1A- Δ L) del Anexo. La Figura A3.3 del Anexo presenta los gráficos de $^{13}\text{C}_\alpha\Delta\delta$, $^{13}\text{C}_\beta\Delta\delta$, $^{13}\text{C}'\Delta\delta$, $^1\text{H}\Delta\delta$ y $^{15}\text{N}\Delta\delta$ en función de la secuencia aminoacídica de AdE1A- Δ E, mientras que la figura A3.4 del Anexo los muestra para AdE1A- Δ L. El análisis de los valores de $^{13}\text{C}_\alpha\Delta\delta$, $^{13}\text{C}_\beta\Delta\delta$, $^{13}\text{C}'\Delta\delta$, $^1\text{H}\Delta\delta$ y $^{15}\text{N}\Delta\delta$ para ambas mutantes sugieren también que estas proteínas carecen de estructuras secundarias evidentes a lo largo de toda su extensión y son preponderantemente desordenadas en estado libre en solución.

En la Figura 3.7, se muestran además de modo comparativo y de ejemplo, los $^{13}\text{C}_\alpha\Delta\delta$ de la proteína AdE1A.WT y de las mutantes AdE1A- Δ E y AdE1A- Δ L. Como se puede observar los perfiles de los $\Delta\delta$ son muy similar entre sí, lo que permite inferir por un lado, que ni la mutación del motivo E2F en AdE1A- Δ E, ni la mutación del motivo LxCxE en AdE1A- Δ L parecerían influir en los $\Delta\delta$ de residuos adyacentes ni de otros más lejanos a los sitios mutados indicando que no favorecerían por ejemplo la formación *de-novo* de estructuras secundarias locales transientes que no están presentes en el estado de AdE1A. Por otro lado, es indicativo que conservan una conformación similar a AdE1A.WT.

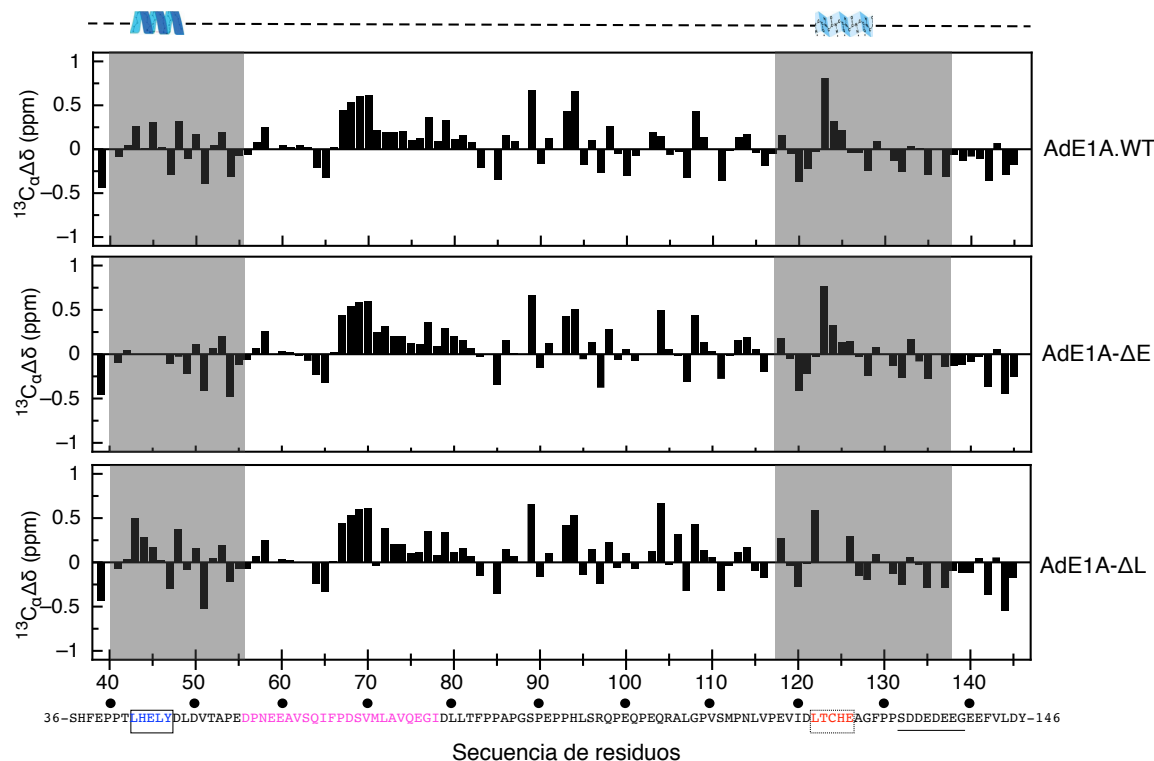


Figura 3.7. Desplazamientos químicos secundarios de $^{13}\text{C}_\alpha$ por residuo para la proteína AdE1A.WT y las mutantes AdE1A-ΔE y AdE1A-ΔL. Las regiones correspondientes a los motivos de interacciones y residuos flanqueantes están señalados en gris oscuro. Arriba, esquema de la conformación que AdE1A adquiere en su forma unida a RbAB; el motivo E2F (α -hélice) y el motivo LxCxE, aparece plegado en (lámina- β). Abajo, la secuencia de AdE1A.WT y señalados el motivo E2F (azul), el motivo TRAM-CBP (en magenta) y el motivo LxCxE (en rojo), la región ácida contigua se encuentra subrayada.

III.1.B.2. Caracterización de las propiedades conformacionales de AdE1A: determinación de la distribución de probabilidad de estructura secundaria en la proteína

Una característica de las IDPs es que presentan gran heterogeneidad conformacional haciendo difícil su descripción y representación como una única estructura, por lo que se hace necesario caracterizar sus propiedades conformacionales en términos de probabilidad de distribuciones [219]. Los desplazamientos químicos (δ) de núcleos que han sido previamente asignados para una proteína son observables que permiten esta caracterización, y pueden ser utilizados para encontrar información cuantitativa acerca de propensiones de estructura secundaria (α -hélice, lámina- β , PPII y “*random coil*”) a nivel de residuo.

La distribución de probabilidad de estructuras secundarias fue calculada para AdE1A.WT utilizando el método $\delta 2d$ desarrollado por Vendruscolo y colaboradores [219]. El método $\delta 2d$, que utiliza los δ de los núcleos $^{13}\text{C}_\alpha$, $^{13}\text{C}_\beta$, $^{13}\text{C}'$, ^1H amida y ^{15}N amida y análisis computacionales para encontrar las propensiones estructurales en una proteína, provee un mayor nivel de discriminación entre diferentes tipos de estructura secundaria en comparación

al simple análisis de los valores de $^{13}\text{C}_\alpha\Delta\delta$ [48]. No obstante, estos datos en conjunto con los $\Delta\delta$ vistos en la sección anterior, permiten hacer inferencias más acabadas acerca de las posibles estructuras secundarias presente en una IDP en su estado libre en solución.

Los valores de probabilidad de contenido de estructuras secundarias por residuo para AdE1A.WT, fueron obtenidos luego de proveer al servidor web del laboratorio del Dr. Vendruscolo (disponible en <http://www-mvsoftware.ch.cam.ac.uk>) con una tabla conteniendo los desplazamientos químicos de su secuencia. En la Tabla A3.7 del Anexo, están tabulados los valores resultantes de probabilidad por residuo para AdE1A.WT. En la Figura 3.8 se encuentran graficados dichos resultados, donde la probabilidad de estructura tipo α -hélice (en negro), lámina β (en gris oscuro) y PPII (en gris claro) se muestran apilados para todos los residuos con δ asignados. Las probabilidades remanentes no mostradas corresponden al tipo “*random coil*”. A su lado, se presentan como porcentajes las propensiones a poblaciones secundarias que asume AdE1A.WT en promedio considerando todos sus residuos.

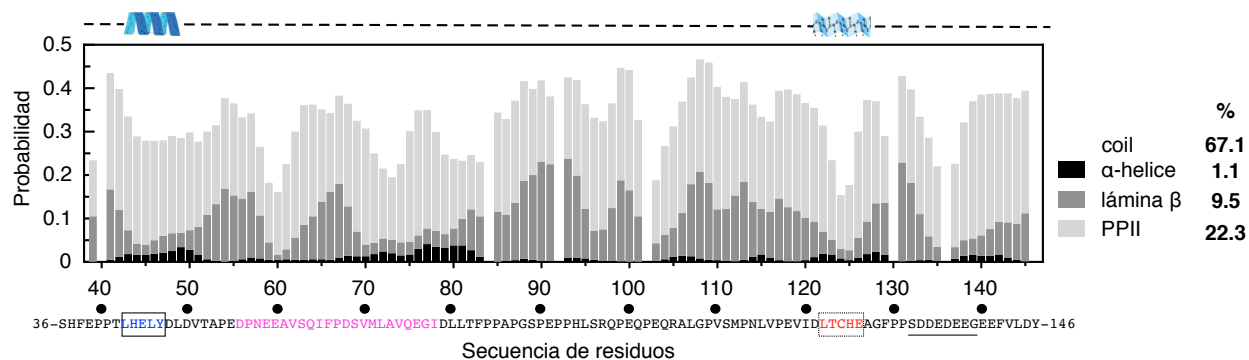


Figura 3.8. Propensión de estructura secundaria por residuo para AdE1A.WT. Representación apilada de las probabilidades por residuo de α -hélice (en negro), lámina β (en gris oscuro) y PPII (en gris claro). La probabilidad de “*random coil*” no se encuentra representada. Los valores medios para la proteína a la derecha. Arriba, esquema de la conformación que AdE1A adquiere en su forma unida a RbAB; el motivo E2F (α -hélice) y el motivo LxCxE, aparece plegado en (lámina- β). Abajo, la secuencia de AdE1A.WT y señalados el motivo E2F (azul), el motivo TRAM-CBP (en magenta) y el motivo LxCxE (en rojo), la región ácida contigua se encuentra subrayada.

En la Tabla 3.3, se presentan los valores porcentuales de las propensiones poblacionales para las diferentes regiones en la secuencia de AdE1A.WT trabajada.

Tabla 3.3. Propensiones a poblaciones secundarias en porcentajes, para distintas regiones de la secuencia de AdE1A.

Regiones en la secuencia de AdE1A	%			
	"random coil"	α -hélice	lámina- β	PPII
Motivo E2F y residuos flanqueantes (39-55)	68.53	1.38	8.02	22.07
Motivo CBP-TRAM y algunos residuos extras (56-83) "linker 1"	71.49	1.66	6.87	19.98
Región rica en prolinas (84-117) "linker 2"	62.65	0.57	13.83	22.95
Motivo LxCxE y residuos flanqueantes incluyendo región ácida (118-138)	68.79	0.88	8.25	22.08
Región C-terminal (139-146)	61.59	1.10	7.10	30.21

Como se puede observar, la probabilidad media de que AdE1A.WT se encuentre asumiendo estructura secundaria tipo α -hélice o entendido de otra manera, el porcentaje de residuos en la proteína con dicha conformación, es de apenas un 1.1% según este análisis. Éste último resultado es algo inferior al valor encontrado por CD para la misma secuencia en similares condiciones experimentales (6.96%), aunque esto puede entenderse por el hecho de que se están reportando poblaciones de estructura secundaria calculadas a partir de observables independientes (δ y elipticidad molar) y a que si bien se ha demostrado ambas técnicas muestran sorprendente aproximación ($r=0.99$), esta correlación disminuye ($r=0.93$) para proteínas cuyo contenido total de estructura secundaria es inferior al 30% [219] como lo es en este caso.

Por su parte en el análisis por regiones, la región comprendida entre los residuos 58 a 83 (que comprende al motivo CBP-TRAM) y se llamó "linker 1", es la que presenta mayor propensión conformacional en α -hélice (1.5 veces mayor a la media de 1.1%), coincidiendo con la región anteriormente descrita con valores positivos consecutivos de $^{13}\text{C}_\alpha\Delta\delta$. Sin embargo tales coincidencias pueden ser esperables teniendo en cuenta que para ambos análisis se utilizan como datos de entrada los valores de desplazamientos químicos (δ) de AdE1A.WT. La región del motivo E2F y residuos flanqueantes -que se estructura en α -hélice tras interactuar con RbAB- también presenta un valor superior a la media (1.38%). No obstante, tanto para esta región, como para "linker 1" estos valores que no superan el 2% son poco significativos en el total de estructura.

El contenido de lámina- β y PPII es mayor que el de α -hélice, ascendiendo a valores medios 9.5% y 22.3% respectivamente. La región “*linker 2*” que corresponde a la región rica en prolinas, presenta valores superiores a la media para estas dos conformaciones, que llegan a representar en conjunto a un 37% para dicha región. Justamente lámina- β y PPII corresponden a las conformaciones más “extendidas”, dotando quizás esta región “*linker 2*” de un importante grado de extensión que pudiera ser clave en el comportamiento hidrodinámico descrito para AdE1A.

Parecería que hay un mínimo de tendencia estructural, pero la media poblacional del 67% para la conformación tipo “*random coil*”, sustenta el alto grado de desorden de esta proteína en toda su extensión.

Para las mutantes AdE1A- Δ E y AdE1A- Δ L también fue calculada la distribución de probabilidad de estructuras secundarias usando el método δ 2d. En las Tablas A3.8 y A3.9 del Anexo, están tabulados los valores resultantes de probabilidad por residuo para AdE1A- Δ E y AdE1A- Δ L respectivamente. La Figura 3.9 presenta dichos resultados por residuo y el valor porcentual medio de propensiones para cada proteína.

En el caso de AdE1A- Δ E, la propensión a estructura secundaria tipo α -hélice no supera una media del 1.1%, valor que está por debajo de la estimación de contenido de α -hélice por CD (6.80%), aunque esta diferencia relativa se mantiene similar al caso de AdE1A.WT. La propensión media de lámina β y de PPII son en cambio mayores a la de α -hélice (9.7% y 21.6% respectivamente), sin embargo la conformación “*random coil*” es dominante a lo largo de toda la extensión (67.6%) (Figura 3.9 A). Las propensiones conformacionales medias de AdE1A- Δ L son muy similares (al igual que las diferencias relativas entre su propensión a α -hélice reportado por δ 2d y CD) y no varían en más de 2 puntos porcentuales respecto de las de AdE1A- Δ E (Figura 3.9 B).

Estos resultados son indicativos que ni AdE1A- Δ E ni AdE1A- Δ L tienen tendencias conformacionales claras a estructuras secundarias canónicas y son más bien desordenadas, por ende las dos mutantes no se diferencian conformacionalmente de AdE1A.WT.

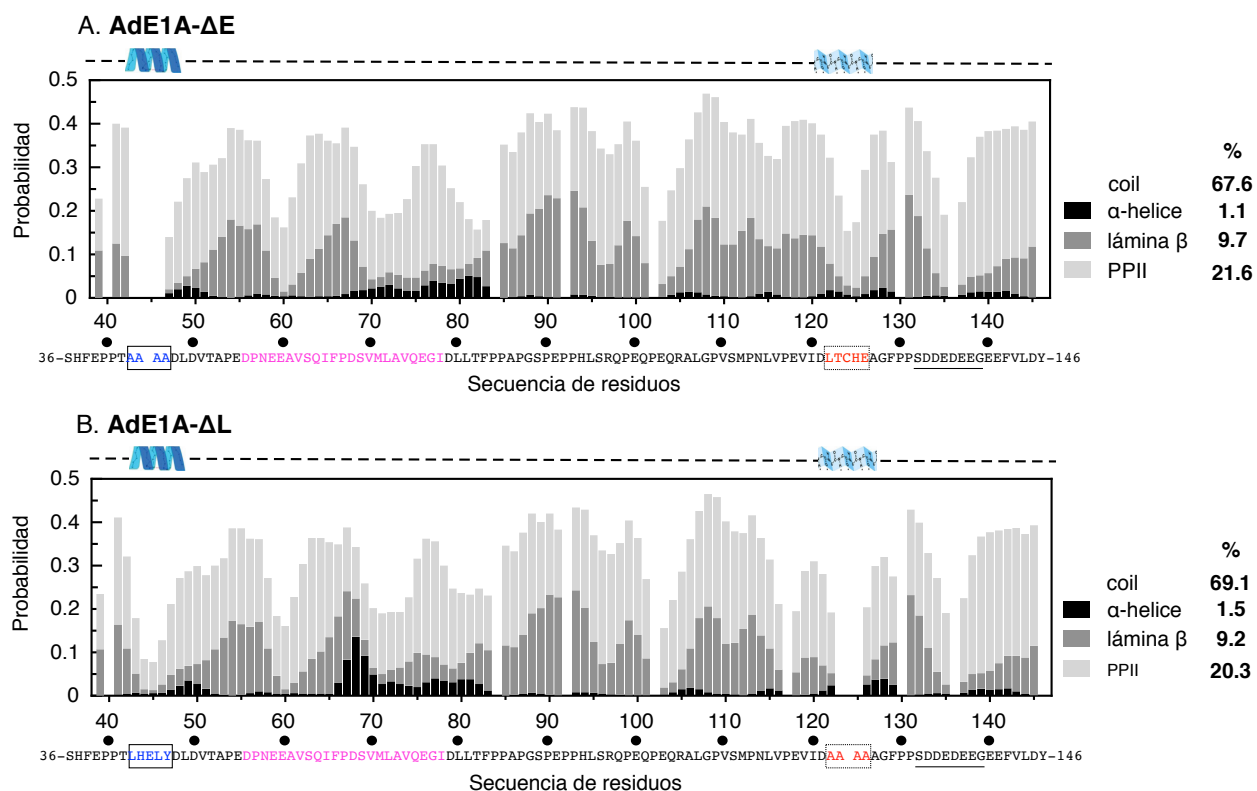


Figura 3.9. Propensión de estructura secundaria por residuo para AdE1A-ΔE (panel A) y AdE1A-ΔL (panel B). Representación apilada de las probabilidades por residuo de α-hélice (en negro), lámina β (en gris oscuro) y PPII (en gris claro). La probabilidad de *random coil* no se encuentra representada. Los valores medios para la proteínas, a la derecha. Arriba de cada gráfico, el esquema de la conformación que AdE1A adquiere en su forma unida a RbAB; el motivo E2F (α-hélice) y el motivo LxCxE, aparece plegado en (lámina-β). Abajo, la secuencia de AdE1A y señalados los motivos o sus correspondientes mutaciones.

En conclusión, a pesar del contenido de estructuras secundarias que pudiera poblar AdE1A, esta proteína es notablemente desordenada, con una tendencia claramente preponderante a conformación tipo “random coil-like”. El resultado general obtenido del análisis por esta última metodología es concordante con el obtenido a partir de los desplazamientos químicos secundarios para AdE1A.WT, demostrando que su tendencia por estructuras secundarias canónicas es muy baja y es más bien una proteína desordenada en toda su extensión. Ninguno de los sitios que median interacciones (ni otros sitios potenciales) con RbAB, presentan entonces estructuras pre-formadas ni tendencias poblacionales marcadas, abonando a un modelo de ajuste inducido para dichas interacciones.

III.1.C. Dinámica conformacional de AdE1A en estado libre. Experimentos de relajación.

Para entender las funciones de las proteínas es necesario describir sus estructuras pero también caracterizar sus movimientos, ya que por ejemplo la flexibilidad de las conformaciones proteicas particularmente para las IDPs es esencial para que estas cumplan sus roles [220].

Las variaciones locales en la dinámica del esqueleto carbonado de una proteína que resultan de la constricción o liberación de movimientos, ocurren en diferentes escalas de tiempo, que son accesibles por distintos estudios de RMN en solución [221] (Figura 3.10). Los experimentos RMN de relajación de espín nuclear, censan la *dinámica conformacional y molecular interna* de las proteínas, así como la *flexibilidad y comportamiento hidrodinámico* de IDPs en el rango de pico a microsegundos, y que incluyen movimientos de la cadena carbonada, rotaciones de cadenas laterales, “*tumbling*” molecular, movimientos de partes interna proteicas e incluso más lentos como los plegamientos [214]. Estos experimentos miden la relajación de los núcleos ^{15}N o de ^{13}C , que están gobernadas predominantemente por interacciones dipolo-dipolo con los hidrógenos a los que están unidos y en menor extensión por el mecanismo de anisotropía de desplazamiento químico, y pueden ser interpretados sin un conocimiento detallado de las propiedades estructurales de la proteína en cuestión [222].

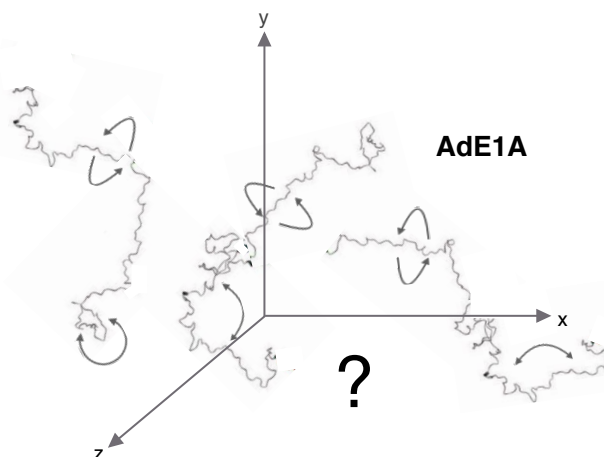
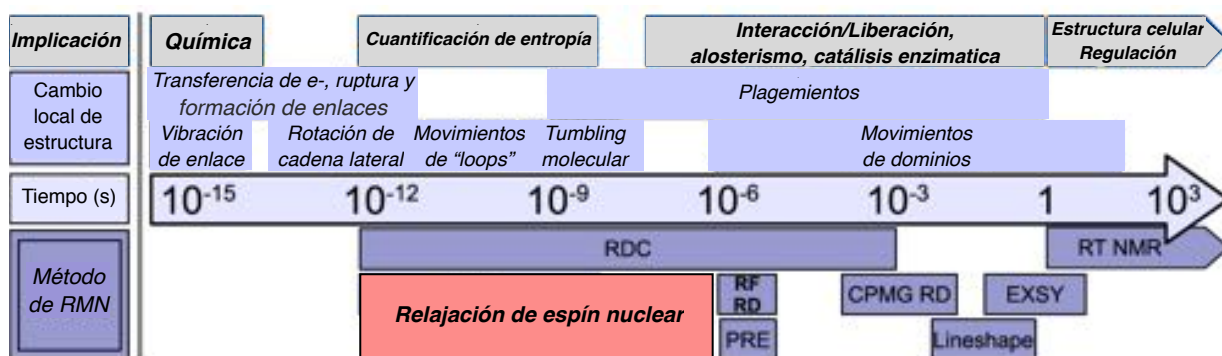


Figura 3.10. Estudio de la dinámica conformacional proteica por métodos de RMN. Arriba, cambios proteicos comprendidos en una escala amplia de tiempo y los métodos RMN disponibles para estudiarlos. Estos procesos dinámicos pueden ser estudiados por: RDC: Acoplamiento residual dipolar; RT RMN: RMN en tiempo real; RF/RD: Dispersión-relajación rotacional; CPMG RD: Dispersión-relajación Carr-Purcell Meiboom-Gill; EXSY: Espectroscopía de intercambio (o intercambio zz); PRE: Potenciación de relajación paramagnética; análisis Lineshape y Experimentos de relajación del espín nuclear (en rojo claro) (Adaptado de [221]). Abajo, AdE1A, una IDP altamente desordenada de la cual se esperaría una elevada flexibilidad conformacional para que pueda cumplir sus múltiples funciones como proteína “hub”.

Ya se ha probado el alto grado de desorden de AdE1A, pero también es necesario caracterizar su dinámica conformacional y flexibilidad al estado libre. Dada de la ausencia de estructuras consolidadas en esta proteína y a que se trata de una proteína “*hub*”, que puede interactuar con múltiples blancos celulares es esperable que la misma presente una alta flexibilidad al estado libre sería esperable para esta proteína (Figura 3.10, abajo).

Para estudiar estas dinámicas en proteínas, los experimentos de relajación, usualmente se miden sobre una proteína uniformemente marcada con ^{15}N , una serie de experimentos basados en métodos 2D-HSQC, donde la intensidad de los picos son modulados como una función de tiempos de retraso (configurados en la secuencia de pulsos) cuando el proceso de relajación está activo. Estos experimentos son:

- a. Tiempo de relajación ^{15}N T1 o relajación espín-red o longitudinal R1 ($1/T1$) útil para evaluar movimientos en la escala de pico a nanosegundos. R1 varía generalmente entre 0.5 y 5 Hz (s^{-1}) de proteínas grandes a pequeñas, se incrementa con “*tumblings*” mas rápidos y decrece con la flexibilidad asociada a una proteína.
- b. Tiempo de relajación ^{15}N T2 o relajación espín–espín o transversal R2 ($1/T2$) el cual es sensible a movimientos más lentos como el plegado de proteínas o intercambios químicos incluso hasta en la escala de los microsegundos. R2 varía típicamente entre 1 y 50 Hz (s^{-1}) de proteínas pequeñas a grandes, incrementa con “*tumblings*” más lentos y decrece con una mayor flexibilidad.
- c. Tiempo de relajación ^{15}N T1 ρ o relajación rotacional R1 ρ ($1/T1\rho$) para monitorear a nivel de residuos fenómenos de intercambio conformacional lento que ocurren en escala de microsegundos;
- d. Simultáneamente también se realizan experimentos $^1\text{H}^{15}\text{N}$ NOE (efecto nuclear Overhauser), que son útiles para analizar las dinámicas del esqueleto carbonado en una escala de pico a nanosegundos y que además proveen información parcial acerca de la presencia de elementos de estructuras secundarias e interacciones de largo alcance. Estos varían entre -3.5 hasta casi 1 (≈ 0.8) y sus valores se reducen en presencia de flexibilidad interna.

Con el objetivo de caracterizar la dinámica conformacional de AdE1A.WT en estado libre en solución, se llevaron a cabo estos experimentos en un equipo Varian VNMRS de 600 MHz para una muestra uniformemente marcada con ^{15}N a 420 μM . Los resultados de estas adquisiciones, luego de procesados y ajustados (Ver Materiales y Métodos) se presentan aquí como valores R1, R2, R1 ρ y $^1\text{H}^{15}\text{N}$ NOE graficados en función del número de residuos de la secuencia (Figura 3.11). Estos valores y sus correspondientes errores están tabulados en la

tabla 3.10 de Anexos. En la Tabla 3.4 están dadas las cifras promedio para la proteína. En general el error de los ajustes estuvo dentro del 10%, lo que es esperable para datos con buena calidad. Un total de 85 resonancias no-prolinas fueron reportadas.

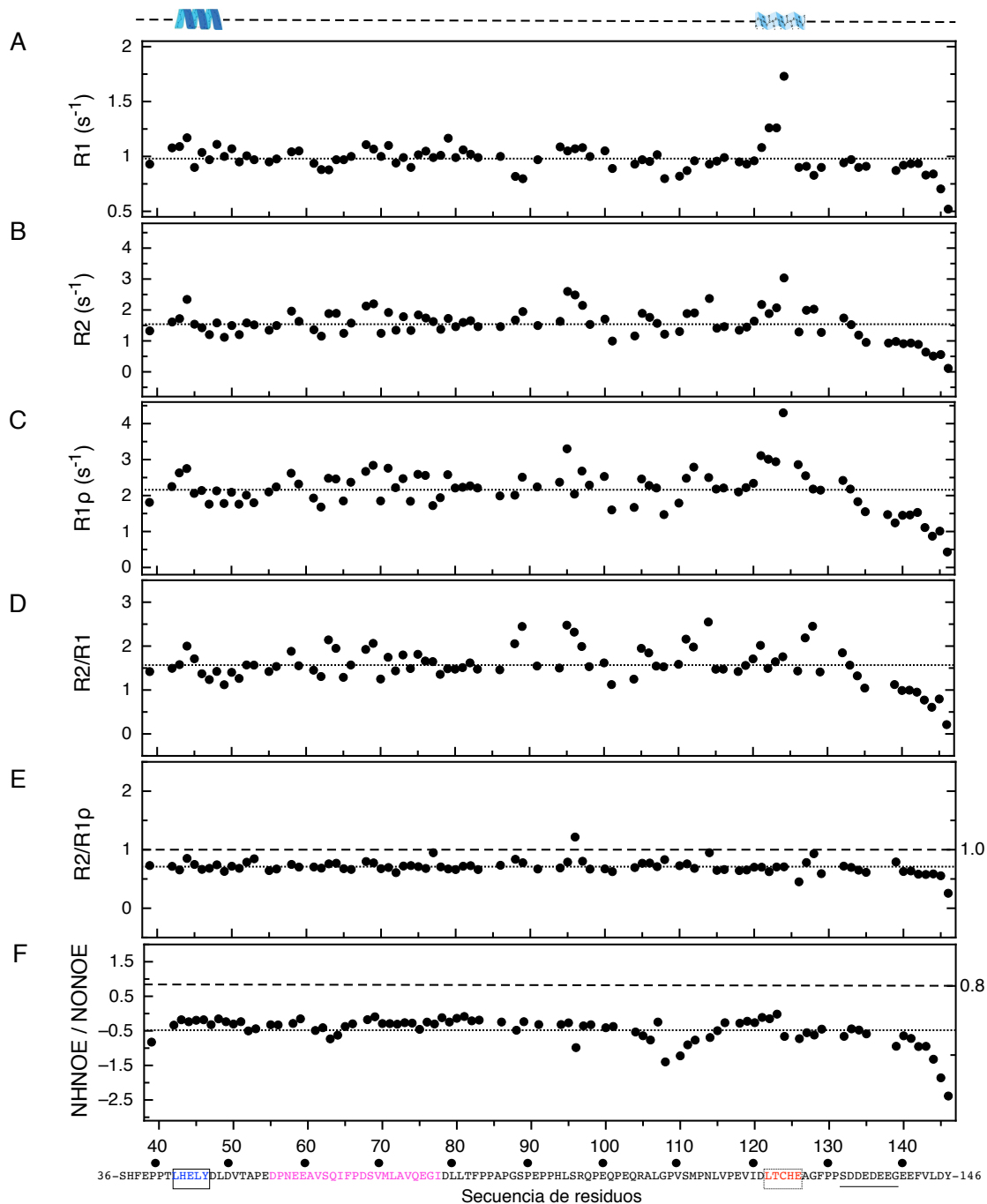


Figura 3.11. Parámetros de relajación por residuo para AdE1A.WT. A) Gráfico de relajación espín-red (R1) B) Gráfico de relajación espín-espín (R2) C) Gráfico de relajación rotacional (R1ρ) D) Gráfico de razón R2/R1. E) Gráfico de razón R2/R1ρ. Los valores por residuo se comparan con respecto a 1 (línea de rayas). F) Gráfico de NHNOE/NONOE. El valor 0.8 (línea de rayas) es el máximo esperado para conformaciones plegadas o estructuradas. Los valores promedio de cada uno de los parámetros (presentados en la Tabla 3.3) están señalados con la línea de puntos. Arriba, esquema de la conformación que AdE1A adquiere en su forma unida a RbAB; el motivo E2F (α-hélice) y el motivo LxCxE, aparece plegado en (lámina-β). Abajo, la secuencia de AdE1A.WT y señalados el motivo E2F (azul), el motivo TRAM-CBP (en magenta) y el motivo LxCxE (en rojo), la región ácida contigua se encuentra subrayada.

Tabla 3.4. Valores medios de los parámetros de relajación de AdE1A.WT

R1 (s ⁻¹)	0.98 ± 0.03
R2 (s ⁻¹)	1.54 ± 0.06
R1ρ (s ⁻¹)	2.16 ± 0.08
R2/R1	1.57 ± 0.09
R2/R1ρ	0.71 ± 0.05
¹ H ¹⁵ N NOE	-0.48 ± 0.03

Para AdE1A.WT, la mayoría de los residuos asumen valores entre 0.9 y 1.2 s⁻¹ para R1. Esta variabilidad es mayor para R2 y R1ρ los cuales asumen valores entre 1 y 2 s⁻¹ y entre 1.5 y 3.3 s⁻¹ respectivamente. Para estos parámetros, valores alejados de la media para la proteína reflejan movimientos internos sitio-específicos, no obstante en este caso no se alejan de sus valores promedios. Las excepciones lo constituyen aproximadamente los últimos 10 residuos en el extremo C-terminal, cuyos valores de relajación por debajo de las medias (o Tiempos de relajación mayores, recordando que R=1/T) sugieren aun mayor flexibilidad y movilidad, lo que es esperable para los extremos de una cadena polipeptídica desordenada [222]. Otra excepción la constituye la Cisteína 124 de la secuencia, un aminoácido conservado del motivo LxCxE, que asume valores de R1= 1.73 s⁻¹, R2= 3.04 s⁻¹ y R1ρ= 4.3 s⁻¹ y los residuos contiguos L122 y T123 cuyos R1= 1.26 s⁻¹. Esto se debería a que C124, que es altamente reactiva, está mediando la dimerización de AdE1A.WT en una muestra parcialmente oxidada (como fue demostrado en el Capítulo I), lo que la llevaría a asumir tiempos de relajación superiores a la media, influenciando también a residuos colindantes. Como se verá más adelante, la mutante AdE1A.ΔE -que conserva dicha cisteína- no presenta el mismo comportamiento probablemente debido a mejores condiciones reductoras. Esto sugiere que el esqueleto carbonado de los residuos del motivo LxCxE, al igual que los del motivo E2F (43-LHELY-47) no están involucrados en procesos dinámicos distintos de los del resto de la cadena polipeptídica cuando la proteína está libre.

La razón R2/R1 para AdE1A.WT, asumió un valor medio de 1.57 (Tabla 3.4). Esta razón es proporcional a τ_c, el tiempo de correlación o rotación isotrópica global (aquel en el que todas las orientaciones moleculares se muestrean por el mismo período de tiempo). Para una proteína globular rígida del peso molecular de AdE1A se esperaría teóricamente un τ_c de ≈ 8.2 nseg, es decir una razón R2/R1 de ≈ 8.2 [223], pero como el valor la razón que asume es mucho mas bajo, esto es indicativo de su elevada flexibilidad interna.

La razón R2/R1 ρ también fue calculada y asumió un valor promedio de 0.71 para AdE1A.WT (Tabla 3.4). Los valores individuales por residuo no se alejan de este valor promedio. Valores de R2/R1 ρ alejados significativamente de 1, -como a mas del doble- sugieren que la proteína experimenta cambios conformacionales importantes en su estado libre [224], lo que no estaría ocurriendo para el caso en estudio.

Por otro lado, también fueron medidos los HN-NOE por residuo para AdE1A.WT. Este parámetro sugiere la presencia de conformaciones plegadas si toma valores positivos y de hasta un máximo +0.8, aunque por sí mismo no es indicativo de regiones con estructura secundaria y debe interpretarse a la luz de otros datos -como los desplazamientos químicos secundarios [222]. En cambio, valores HN-NOE cercanos a cero y negativos de hasta un mínimo de -3.5, son indicativos de flexibilidad y son característicos de IDPs [225]. AdE1A asumió valores negativos en toda su extensión con una media de -0.48 (Tabla 3.4), lo que demuestra su naturaleza desordenada y flexible.

La Tabla 3.5 reúne a modo de ejemplo, los valores asumidos para estos parámetros de relajación para 4 proteínas o dominios: 2 de ellos IDPs [226, 227] y 2 globulares [224, 228] con pesos moleculares cercano al de AdE1A.WT. Como se puede observar, los valores promedios para estos parámetros en AdE1A.WT son consistentes con valores esperados para proteínas intrínsecamente desordenadas y flexibles, en tanto se alejan de valores para proteínas globulares con peso molecular equivalente.

Tabla 3.5. Comparación de los valores medios de los parámetros de relajación para distintas proteínas desordenadas y globulares de peso molecular aproximado

	IDRs/IDPs			Dominios/Proteínas globulares	
	AdE1A.WT	p53 TAD ₁₋₇₃	PAGE5 ₁₋₁₁₀	PED/PEA15 DED ₁₋₈₀	FGF2 Apo ₃₀₋₁₅₂
PM (kDa)	12.5	8.8	13.2	9.6	14.6
R1 (s ⁻¹)	0.98 ± 0.03	1.74 ± 0.22	1.93 ± 0.25	1.54 ± 0.14	1.89 ± 0.13
R2 (s ⁻¹)	1.54 ± 0.06	3.16 ± 0.56	3.44 ± 0.78	13.3 ± 2.01	12.51 ± 1.24
R1 ρ (s ⁻¹)	2.16 ± 0.08	3.81 ± 0.66	<i>sd</i>	<i>sd</i>	12.26 ± 1.19
R2/R1	1.57 ± 0.09	1.82 ± 0.41	1.77 ± 2.08	8.7 ± 1.6	6.62 ± 0.79
R2/R1 ρ	0.71 ± 0.05	0.82 ± 0.19	<i>sd</i>	<i>sd</i>	1.02 ± 0.14
¹ H ¹⁵ N NOE	-0.48 ± 0.03	-0.64 ± 0.95	0.03 ± 0.15	0.79 ± 0.12	0.77 ± 0.05
		<i>Vise 2005</i>	<i>Hellman 2011</i>	<i>Farina 2010</i>	<i>Pagano 2012</i>

Todas las mediciones fueron realizadas a 25°C. Las mediciones para AdE1A.WT, p53 TAD y PED/PEA15 fueron realizadas en un equipo Varian 600 MHz. Para PAGE5 en un equipo Varian 600 MHz y para FGF2 en un equipo Bruker DMX 500 MHz. *sd* sin datos

Los experimentos de relajación también fueron hechos sobre muestras uniformemente marcadas con ^{15}N de la mutante AdE1A- ΔE a 690 μM y de la mutante AdE1A- ΔL a 680 μM . Los resultados de estas adquisiciones, luego de procesados y ajustados están graficados en función de la secuencia:

- Para AdE1A- ΔE en la Figura A3.5 de Anexos. Los valores y sus correspondientes errores están tabulados en la tabla 3.11 de Anexos.
- Para AdE1A- ΔL en la Figura A3.6 de Anexos. Los valores y sus correspondientes errores están tabulados en la tabla 3.12 de Anexos.

En la Tabla 3.6 están dadas las cifras promedio de los parámetros de relajación para las dos mutantes. En ambos casos, los errores de los ajustes estuvieron dentro del 10%. Un total de 79 resonancias no-prolinas fueron reportadas para AdE1A- ΔE y un total 80 resonancias en el caso de AdE1A- ΔL .

Tabla 3.6. Valores medios de los parámetros de relajación para las variantes mutantes de AdE1A

	AdE1A.WT	AdE1A- ΔE	AdE1A- ΔL
R1 (s^{-1})	0.98 \pm 0.03	0.98 \pm 0.03	0.90 \pm 0.05
R2 (s^{-1})	1.54 \pm 0.06	1.40 \pm 0.18	1.74 \pm 0.05
R1 ρ (s^{-1})	2.16 \pm 0.08	2.35 \pm 0.06	1.63 \pm 0.12
R2/R1	1.57 \pm 0.09	1.40 \pm 0.19	1.91 \pm 0.11
R2/R1 ρ	0.71 \pm 0.05	0.58 \pm 0.08	1.07 \pm 0.09
$^1\text{H}^{15}\text{N}$ NOE	-0.48 \pm 0.03	-0.52 \pm 0.02	-0.55 \pm 0.02

Para ambas mutantes, con excepción nuevamente de sus extremos C-terminal, los valores individuales por residuo de todos los parámetros de relajación no se alejan de sus valores promedios poblacionales (Figura A3.3 y A3.4 de Anexos). Cabe aclarar que en el caso de AdE1A- ΔE , la mutante que conserva el motivo LxCxE de interacción a pRb, no se observa el mismo comportamiento para la cisteína 124 que se había visto en la variante .WT; esto sostiene la conclusión de que dicha particularidad se debería a un problema de oxidación parcial de la muestra de AdE1A.WT. Al igual que para la variante natural, los valores promedios para todos los parámetros de relajación en AdE1A- ΔE y AdE1A- ΔL son consistentes con valores esperados para proteínas desordenadas en toda su extensión y altamente flexibles.

En resumen, AdE1A presenta una dinámica de su esqueleto carbonado característica de una proteína desordenada altamente flexible. Además, como no se observan diferencias en los parámetros de relajación entre diferentes zonas o regiones de la proteína, esto evidencia que la misma es uniformemente flexible. Estas características que posee AdE1A al estado libre, serían clave en la interacción con RbAB porque gracias a ellas, la región “linker” permitiría los contactos de ambos motivos de interacción E2F y LxCxE simultáneamente.

III.1.D. Efecto de las mutaciones sobre la estructura y conformación nativa de AdE1A

Para evaluar el efecto de las mutaciones del *motivo E2F* en AdE1A-ΔE (43-LHELY-47Δ43-AA-AA-47) y de las mutaciones del *motivo LxCxE* en AdE1A-ΔL (122-LTCHE-126Δ122-AA-AA-126), sobre la estructura y propensiones conformacionales de AdE1A.WT, se calcularon las perturbaciones en los desplazamientos químicos (δ) combinados de los núcleos ^1H y ^{15}N por residuo. La fórmula utilizada (ver Materiales y Métodos) permite comparar residuo por residuo cada mutante respecto de la proteína nativa. La Figura 3.12 muestra dichos perturbaciones (como Δppm) en función de las secuencias mutantes.

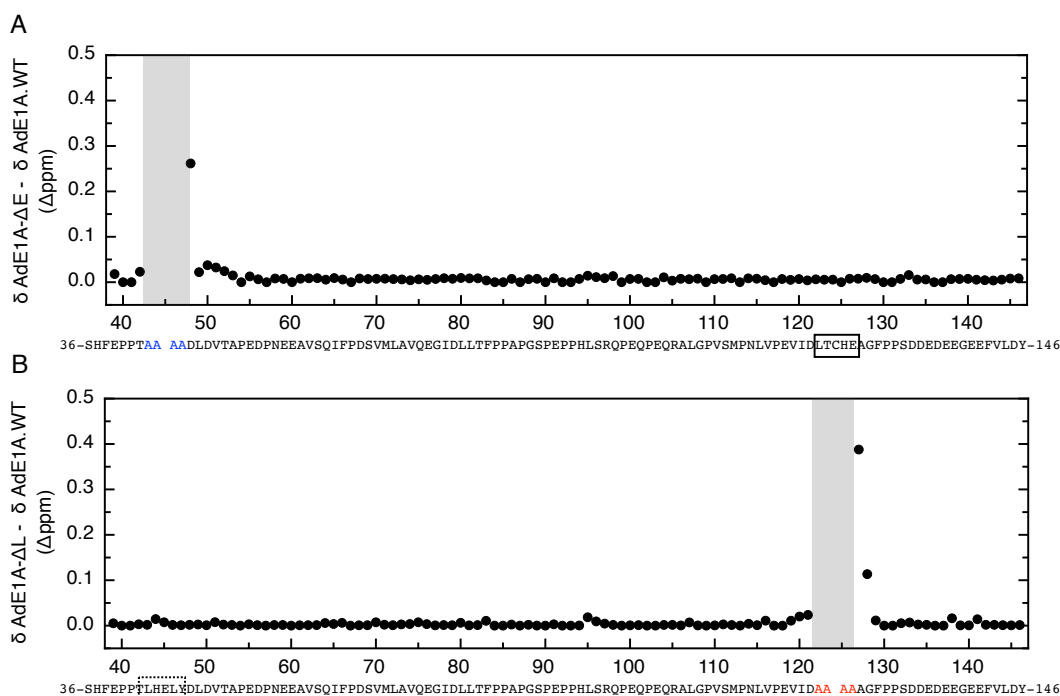


Figura 3.12. Perturbaciones en los desplazamientos químicos (δ) combinados de ^1H y ^{15}N por residuo producido por las mutaciones sobre AdE1A.WT. Los cambios de los desplazamientos químicos de AdE1A.WT vs la secuencia mutante de AdE1A-ΔE (A) y de AdE1A-ΔL (B). Las regiones mutadas están marcadas en gris claro.

Como puede observarse los cambios mas grandes en los desplazamientos químicos se encuentran cercanos a las regiones mutadas (áreas en gris). La mutación en AdE1A- Δ L produjo un mayor cambio para los residuos A127 y G128 de su secuencia, en tanto que la mutación en AdE1A- Δ E, lo produjo sobre D48. Ahora bien, 76 de 81 residuos (93% de los residuos asignados) asumen cambios en los desplazamientos químicos menores a 0.02 Δ ppm para AdE1A- Δ L (Figura 3.12 A) en tanto que 75 de 81 residuos (93% de los residuos asignados) asumen valores inferiores a 0.02 Δ ppm para AdE1A- Δ E (Figura 3.12 B). Estos resultados pueden ser visualizados claramente (Figura 3.13) al observar que los espectros $^1\text{H}^{15}\text{N}$ HSQC de las proteínas mutantes son prácticamente superponibles al espectro $^1\text{H}^{15}\text{N}$ HSQC de la proteína nativa, salvo por las alaninas por las que se mutaron cada uno de los sitios y los residuos mencionados más arriba. Ello indica claramente que no hay cambios significativos en los desplazamientos químicos producidos por las mutaciones sobre AdE1A.WT.

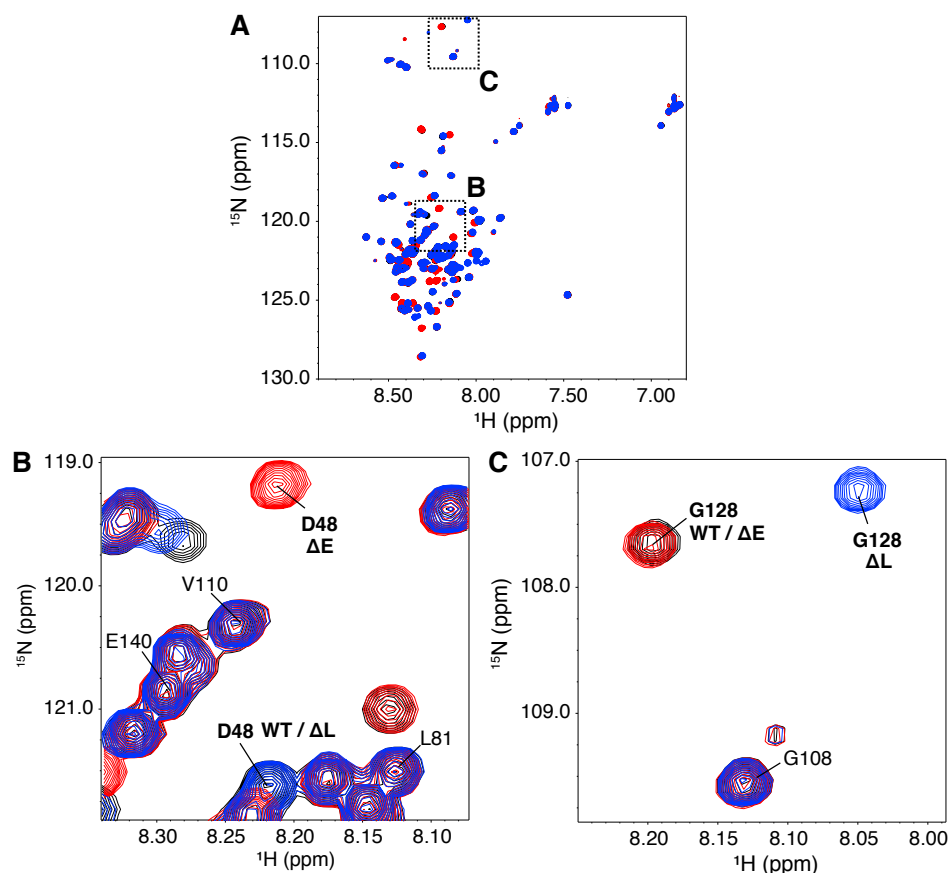


Figura 3.13. Perturbaciones en los desplazamientos químicos producidas por las mutaciones sobre AdE1A.WT visualizadas en los espectros HSQC. A) Superposición de espectros HSQC de AdE1A.WT (en negro), AdE1A- Δ E (en rojo) y de AdE1A- Δ L (en azul). Los tres espectros superponen casi por completo, excepto por los residuos mutados y residuos vecinos a dichas mutaciones. B) Detalle mostrando como D48 superpone en AdE1A.WT y AdE1A- Δ L, pero presenta mayor perturbación en AdE1A- Δ E. C) Detalle mostrando complementariamente como G128 se encuentra perturbada en AdE1A- Δ L.

En conclusión, el análisis de las perturbaciones en los δ de ^1H y ^{15}N conjuntos, la semejanza de los valores de $^{13}\text{C}_\alpha\Delta\delta$ y de las propensiones poblacionales medias analizadas por $\delta 2d$, en conjunto con la similitud en los valores medios de los parámetros de relajación, demuestran que las mutaciones diseñadas solamente para inactivar sitios funcionales de interacción con pRb de AdE1A, no provocan cambios estructurales, conformacionales ni dinámicos significativos en las proteína AdE1A.

OBJETIVO III.2.

Análisis estructural del complejo [RbAB:AdE1A.WT]

III.2.A. Análisis de la interacción entre AdE1A y RbAB. Experimentos TROSY para proteínas libres y en complejo con RbAB.

La aplicación de espectroscopia de RMN en solución a moléculas de alto peso molecular y complejos macromoleculares de importancia biológica, es de interés para obtener de ellos información estructural y dinámica de alta resolución. Dado el modelo de interacción bipartita entre AdE1A y RbAB, con estos experimentos que otorgan información a nivel de residuos, se buscó por un lado caracterizar el estado conformacional del “*linker*” en el complejo, el cual podría estar completamente desordenado, parcialmente estructurado (y no haber sido detectado por métodos como el CD), mediando múltiples interacciones transientes o débiles sin aporte energético considerable, incluso si hay contactos no revelados hasta el momento o una combinación de estos estados (Figura 3.14 A). Por otro lado, conocidas las interacciones mediadas por los motivos mínimos E2F y LxCxE con RbAB, se investigó qué residuos median los contactos en los motivos, si los residuos flanqueantes también median contactos y si estos persisten en ausencia del núcleo del motivo (Figura 3.14 B).

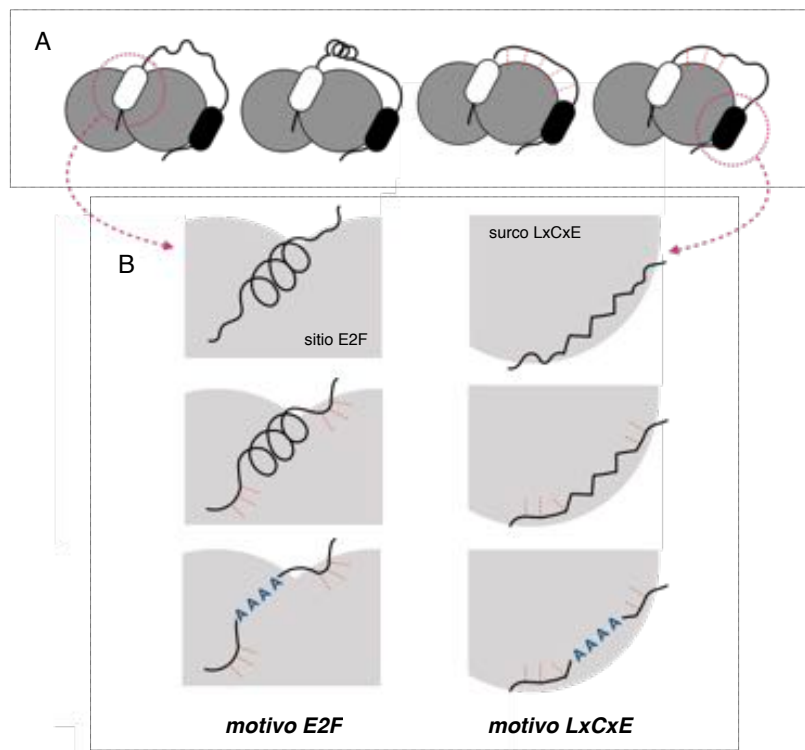


Figura 3.14. Aproximación estructural al complejo [RbAB:AdE1A.WT] por RMN. A) Evaluación de los posibles estados conformacionales y de interacción del “*linker*”. B) Estudio de los residuos interactuantes en los motivos E2F y LxCxE y sitios mutados..

Particularmente, el RMN permite estudiar la formación de complejos entre proteínas, a partir del mapeo de efectos residuo-específicos producidos como consecuencia de la unión. Para ello es necesario adquirir espectros HSQC de la proteína isotópicamente marcada en ausencia y en presencia de un compañero de interacción en su forma sin marcar, trabajando con condiciones y concentraciones adecuadas en cada caso (el compañero debe agregarse estando en el mismo tampón y en concentraciones altas –muy por encima de su K_D de interacción, si se conoce- que aseguren que $\approx 100\%$ de las proteínas estén en su forma acoplada y generalmente en exceso de compañero de interacción) de modo que al comparar ambos espectros superpuestos se puedan obtener conclusiones certeras acerca del proceso de interacción. Como cada pico en un espectro HSQC está definido por su posición (δ , el desplazamiento químico), intensidad (la concentración de cada población en ese estado) y el ancho de la señal (inversamente relacionado con los tiempos de relajación transversal $-T_2-$), las perturbaciones en dichos parámetros como consecuencia de la formación de un complejo, proveen información ligada a dicho proceso [204].

Cuando se está analizando la interacción de una proteína IDP y se siguen los cambios en sus señales, al momento de adquirir el espectro en presencia del compañero de interacción que no está marcado, se pueden presentar distintos escenarios posibles, representados en la Figura 3.15:

- a. Puede que no se produzca cambio alguno, producto de la ausencia de la interacción, o que las condiciones –tampón, pH, temperatura, etc.- no favorezcan la interacción entre ambas proteínas (Figura 3.15, 2.1).
- b. Si pocos picos cambian de posición y/o intensidad, la IDP probablemente esté interactuando vía SLiMs con su compañero y el resto de la proteína permanezca desordenada (Figura 3.15, 2.2).
- c. Ante cambios globales en la posición/intensidad de los picos, la IDP podría estar experimentando procesos de plegamiento-acoplado-a-interacción a gran escala. Cuando ocurre esto, generalmente hay un cambio en el aspecto del espectro porque los picos se mueven sobre todo en la dimensión 1H , lo que es indicativo de que adoptan estructura (Figura 3.15, 2.3).
- d. Si las señales globalmente se ensanchan y desaparecen, el proceso de interacción podría asociarse a tasas de relajación incrementada (T_2), lo que ocurre generalmente con el agregado de proteínas globulares de elevado peso molecular (≥ 50 kDa). También puede asociarse a procesos de oligomerización o agregación (Figura 3.15, 2.4).

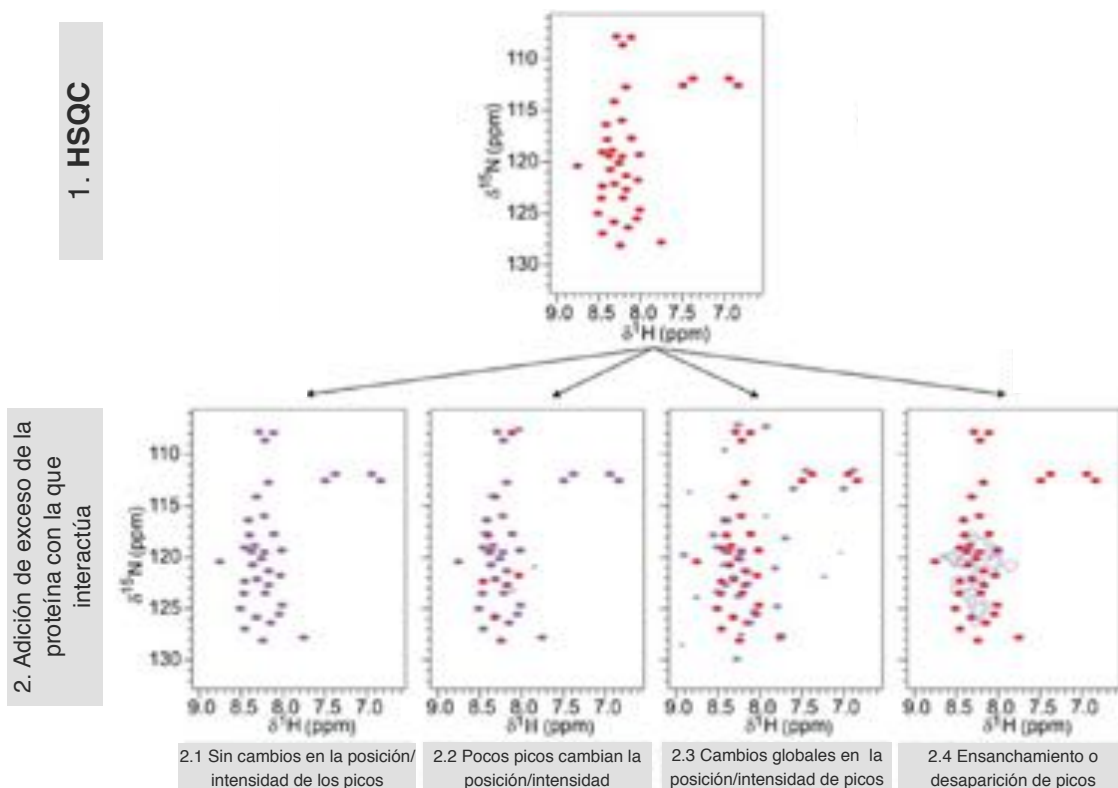


Figura 3.15. Análisis de la formación de complejos de IDPs. Espectro HSQC de la forma libre (1) y los posibles resultados a obtener producto del agregado en exceso del compañero de interacción (2) (Adaptado de [204])

Particularmente frente a esta última situación en el que se evidencia ensanchamiento de las líneas espectrales y baja o incluso nula sensibilidad en el espectro HSQC, una alternativa viable es recurrir de la Espectroscopía de Relajación Transversal Optimizada (TROSY, por sus siglas en inglés) [229, 230]. El TROSY reduce la relajación transversal de los núcleos permitiendo obtener ancho de señales de resonancia y sensibilidad satisfactorias para proteínas (y en este caso complejos) de elevado peso molecular [231].

AdE1A.WT libre y acomplejado con RbAB

La caracterización estructural en alta resolución del complejo entre AdE1A y RbAB, en principio se pensaba realizar utilizando experimentos HSQC. No obstante, como en este caso el peso molecular de RbAB es de 42.1 kDa (y el complejo pesa 54.6 kDa) y otros investigadores habían observado ensanchamiento generalizado de picos en el espectro HSQC y agregación a altas concentraciones de complejo, trabajando con RbAB y AdE1A (1-139) [76] - similar al fragmento de AdE1A utilizado en este trabajo (36 a 146)- se decidió trabajar adquiriendo espectros TROSY.

Para ello, se adquirieron los espectros TROSY de la forma libre de AdE1A y -luego de una mezcla con cantidad equimolar de RbAB- el espectro de la forma acomplejada. Se trabajó a concentraciones de proteína de 525 μM a volumen final 600 μl y en el equipo Varian VNMR 800 MHz (Ver Materiales y Métodos). Aquí es importante destacar tres cosas: la primera es que al trabajar a 525 μM de complejo, se considera que el 100% de las proteínas están unidas, ya que la K_D de interacción es 25 pM, por tanto 20 millones de veces menor a esta concentración de trabajo. Por otro lado, no se observó agregación en la forma acomplejada, incluso trabajando a tan alta concentración. Por último, se obtuvo un espectro TROSY del complejo de buena calidad e interpretable.

El resultado de la asignación (para lo que se usaron los espectros $^1\text{H}^{15}\text{N}$ -HSQC previamente asignados como “moldes”) y superposición de los espectros se presenta en la Figura 3.16. La superposición de espectros TROSY mostró que para la mayoría de los picos, no hubo asociados cambios apreciables en su posición, intensidad y el ancho de señal, asociados al proceso de interacción de AdE1A con RbAB. No obstante se observó clara desaparición de algunos picos puntuales y simultáneamente la aparición de otros nuevos -en posiciones distintas- tras la formación del complejo. Los nuevos picos que se formaron tras la formación del complejo no se asignaron, porque implicaba asignar de cero a AdE1A.WT doblemente marcada (^{15}N - ^{13}C) en su forma acomplejada con RbAB, lo que no fue posible por el esfuerzo requerido y el tiempo limitado de máquina disponible durante el período de 3 meses en el que se realizaron estos experimentos de RMN. No obstante a ello, algunos picos que aparecían tras formarse el complejo, lograron asociarse a formas unidas de residuos de los motivos E2F y LxCxE, como se explica más adelante para la Figura 3.18.

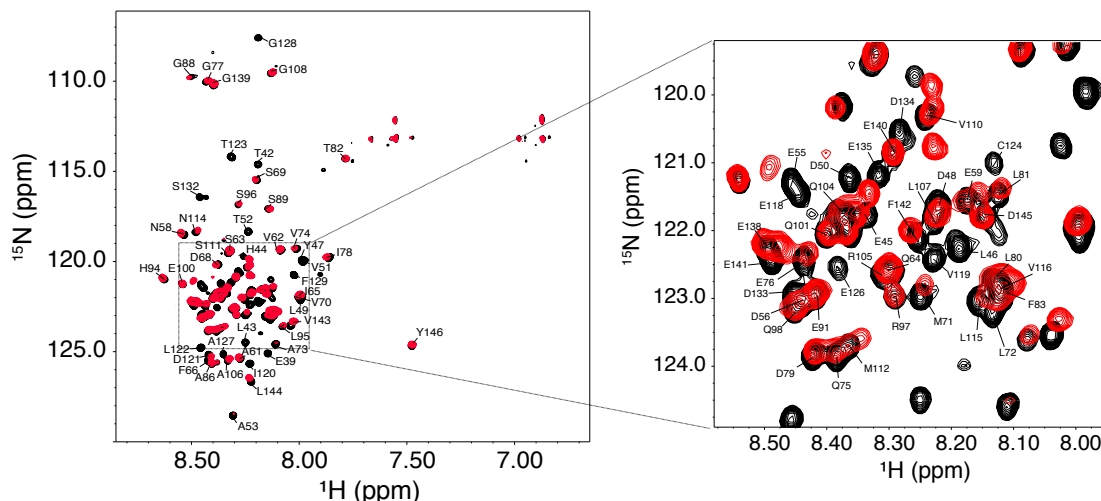


Figura 3.16. Superposición de espectros TROSY de AdE1A.WT libre a 525 μM (negro) y del complejo [RbAB:AdE1A.WT] 1:1 a 525 μM (rojo) y sus correspondientes asignaciones. A la derecha, detalle de la región central del espectro con mayor densidad de picos y sus asignaciones.

El estudio de los cambios producidos por la interacción se hizo fundamentalmente a partir del análisis de la *disminución* en las intensidades de los picos, es decir que no se analizaron nuevos picos formados. A todos los picos asignados (85 en total) le correspondió un valor de intensidad de resonancia en el espectro TROSY de la proteína libre (intensidad inicial, I_0) y otro en el espectro del complejo (intensidad final, I).

Con los valores de intensidad se calculó la razón de intensidad por residuo. En la Figura 3.17 se encuentran graficados las razones de intensidad por residuo para AdE1A.WT. Estas razones de intensidad I/I_0 , se interpretaron cuidadosamente teniendo en cuenta que sus disminuciones están asociadas a interacciones sitio-específicas con la proteína RbAB, pero también pueden ser debidas a cambios en la dinámica (por ej. afectados por el tamaño del sistema) como cambios en el entorno químico de las resonancias producto de interacciones transientes-inespecíficas [204]. Para esto, también se tuvieron en cuenta las caracterizaciones cuantitativa/termodinámicas analizadas para cada fragmento de AdE1A en su interacción con RbAB.

Como se observa en la Figura 3.17 A, los residuos con valores de $I/I_0 \cong 0$, correspondieron a los residuos del núcleo del motivo E2F (43-LHELY-47) y a los residuos del motivo LxCxE (122-LTCHE-126) –sin contar H125 que no fue asignado–.

Además, los residuos de las regiones flanqueantes a ambos motivos también experimentaron desaparición completa de las señales, presentando valores de I/I_0 de 0. Los residuos que experimentaron este comportamiento en torno al motivo E2F fueron: E39, T42, D48, L49, D50, V51, T52, A53, E55. Y los que lo presentaron entorno al motivo LxCxE fueron: E118, V119, I120, D121, A127, G128, F129 y también los residuos S132, D133, D134 y E135 que son residuos de la región ácida contigua al motivo LxCxE de AdE1A. Claramente, esta desaparición de las señales para tales residuos es consistente con los datos termodinámicos de asociación con RbAB, que mostraron que los motivos mínimos se asocian individualmente con K_D 's ≈ 100 nM.

Para el caso del motivo E2F, del que se dispone estructura cristalina en su forma acomplejada con RbAB (Figura 2.3 D), se observan que los residuos L43, H44, L46, L47 y D48 establecen contactos directos con residuos del sitio complementario a E2F en RbAB. Ahora bien, estos resultados de RMN, podrían estar indicando que los residuos desde E39 hacia el N-terminal y hasta el E55 hacia el extremo C-terminal están involucrados en la interacción. Estos residuos flanqueantes serían aquellos que modularían la afinidad del motivo E2F por RbAB.

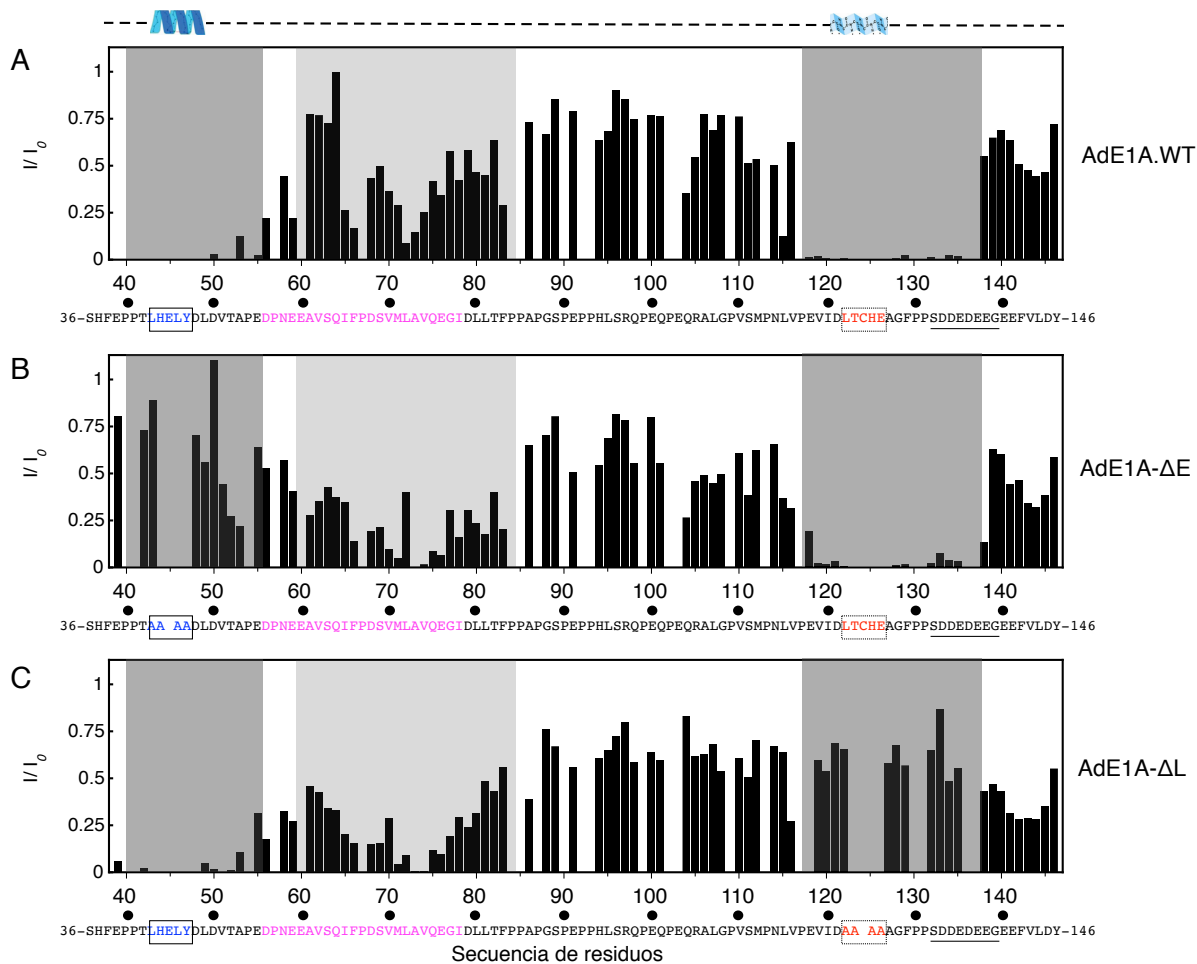


Figura 3.17. Valores I/I_0 entre AdE1A.WT y mutantes libres y unidos a RbAB. A) para AdE1A.WT B) para la mutante AdE1A-ΔE C) para la mutante AdE1A-ΔL. Las áreas sombreadas en gris muestran razones de intensidad bajas y corresponden a residuos que están interactuando con RbAB. Las áreas en gris oscuro corresponden a los motivos de interacción E2Fy LxCxE y residuos flanqueantes. La región CKII-acídica de AdE1A se señala subrayada en la secuencia. En tanto el área en gris claro corresponde a un segmento de la región *linker*, entre residuos E60 a F83, evidenció bajos valores de I/I_0 en los tres casos. Arriba de todos los gráficos, el esquema de la conformación que AdE1A adquiere en su forma unida a RbAB; el motivo E2F (α -hélice) y el motivo LxCxE, aparece plegado en (lámina- β). Abajo de cada gráfico, las secuencia de AdE1A y señalados sus motivos E2F (en azul) y LxCxE (en rojo) y/o correspondientes mutaciones. El motivo TRAM-CBP (en magenta) y la región acídica contigua al LxCxE (subrayada).

En el caso del motivo LxCxE de AdE1A, no existe estructura cristalográfica disponible, sólo la del homólogo LxCxE de E7-HPV unido a RbAB (Figura 2.3 A) que muestra un plegamiento en lámina- β . En la primera parte del capítulo anterior se había visto que el residuo hidrofóbico en posición (-1) –en este caso D121- podría tener efecto modulador. En este experimento de RMN, se ve también que los residuos hasta E118 hacia el extremo N-terminal de la secuencia, podrían estar involucrados activamente en la interacción con RbAB. Hacia el extremo C-terminal, el experimento TROSY confirma también la interacción de los residuos del parche hidrofóbico, que ya se habían visto participaban de la interacción con RbAB en los experimentos cuantitativos/termodinámicos y que presentaban un rol modulador, el cual

también ya se conocía para otras proteínas con el motivo LxCxE (E7-HPV, LT-SV40, HDAC1 celular).

Por otro lado, en el extremo C-terminal presentó un valor I/I_0 medio de 0.57, y particularmente los residuos F142, V143, L144, D145 un valores I/I_0 medio de 0.45. No hay estudios particulares en esta región, aunque al tratarse de residuos de carácter hidrofóbico (incluido Y146) esta disminución podría ser resultado de interacciones inespecíficas con la superficie de RbAB.

Respecto del comportamiento de los picos presentes en la región “*linker*”, es conveniente valerse nuevamente de la sub-clasificación presentada en la Tabla 3.3, en la cual “*linker 1*” correspondía al Motivo CBP-TRAM y algunos residuos extras (56-83) y “*linker 2*” a la región rica en prolinas (84-117). Para la región “*linker 1*”, fue llamativo que los residuos aquí comprendidos presentasen un valor medio I/I_0 bajo (≈ 0.43), particularmente para los residuos entre V70 y E76 (Figura 3.17 A, región gris clara). Este comportamiento ya había sido observado por Ferreon y colaboradores, quienes analizando el espectro HSQC del fragmento AdE1A (27-91) titulado con RbAB (en relación molar 1:2) notaron la casi completa desaparición de las señales M71 a E76 [76]. Los datos de la termodinámica, han demostrado no obstante, que la región “*linker*” contribuye muy poco energéticamente a la unión con RbAB (el $\Delta\Delta G$ era de apenas 0.6 kcal/mol si se comparaban las interacciones con RbAB de AdE1A- ΔE y de AdE1ACR2_{AC}—que se diferencian sólo por la presencia del “*linker*” en el primer fragmento—), lo que implica que estas disminuciones en I/I_0 se deberían a interacciones transientes muy débiles e inespecíficas con RbAB mediadas por estos residuos con carácter hidrofóbico.

Para la región “*linker 2*” de AdE1A en cambio, los valores I/I_0 se mantuvieron altos, alrededor de ≈ 0.66 (con máximos en 0.85), es decir con valores de intensidad escasamente alterados por la formación del complejo. Esta región es la que según el análisis de la distribución de propensiones poblacionales, tendría más probabilidad de adoptar conformaciones extendidas en su estado libre. En este experimento, claramente comprenden residuos que no muestran interacción con RbAB. Que mantengan tan alta relación de intensidades implicaría que están en una conformación muy dinámica aun formado el complejo, es decir que mantendrían la elevada flexibilidad con la que se caracterizó previamente al estado libre de esta proteína.

De igual manera que para AdE1A.WT, se procedió a hacer la adquisición de los espectros TROSY de las mutantes:

- AdE1A-ΔE uniformemente marcada con ^{15}N a concentración 690 μM libre y en complejo con RbAB en relación molar 1:1 a concentración 690 μM de complejo final.
- AdE1A-ΔL uniformemente marcada con ^{15}N a concentración 680 μM libre y en complejo con RbAB en relación molar 1:1 a concentración 680 μM de complejo final.

Para ambos casos, las altas concentraciones de los complejos aseguran que cerca del 100% de las proteínas estén justamente en su forma unida, ya que dichas concentraciones son cerca de 5000 veces superiores a las $K_D \cong 100$ nM, estimada para estas interacciones. Las asignaciones y el procedimiento para obtener los valores utilizados para calcular los valores I/I_0 por residuo se hicieron como está descrito en Materiales y Métodos.

AdE1A-ΔE libre y acomplejado con RbAB

En la figura A3.5 de Anexos se muestran el resultado de la asignación y superposición de los espectros TROSY de AdE1A-ΔE libre y en su forma acomplejada con RbAB. Los valores I/I_0 por residuo están graficados en la Figura 3.17 B.

Como se puede observar los residuos con valores $I/I_0 \cong 0$, correspondieron a los residuos del núcleo del motivo LxCxE (122-LTCHE-126) y sus residuos flanqueantes incluyendo la región acídica, los cuales permanecieron sin mutar en esta variante, mostrando el mismo comportamiento que el observado para AdE1A.WT.

Los residuos de la región “*linker 1*” (Figura 3.17 B, en gris claro), también presentaron valores I/I_0 bajos como en la variante natural, pero asumieron valores sensiblemente menores en todo este segmento en comparación (un valor medio $I/I_0 \cong 0.19$, la mitad del valor medio presentado por la variante AdE1A.WT, $I/I_0 \cong 0.37$).

Ahora bien, en la región conteniendo la mutación del núcleo del motivo E2F (43-LHELY-47) por el parche de tetra-alaninas, los valores I/I_0 para los residuos que pudieron ser asignados en esta variante, permanecieron altos (E39, T42, A43, D48, L49, D50, V51, T52, A53, E55), con un valor medio $I/I_0 \cong 0.64$. Esto indicaría que la interacción de dicho sitio se encuentra efectivamente anulada por la mutación y que los residuos flanqueantes que estarían también interactuando con RbAB en la variante natural de AdE1A, lo harían sólo si los residuos del núcleo del motivo están presentes e interactuando con RbAB, reafirmando su rol modulador pero no central en la interacción.

AdE1A- Δ L libre y acomplejado con RbAB

Para AdE1A- Δ L, la figura A3.6 de Anexos muestra el resultado de su asignación y superposición de los espectros TROSY libre y en complejo y los valores I/I_0 por residuo están en la Figura 3.17 C.

Para las regiones que permanecieron sin modificación en su secuencia, se observaron resultados similares a los presentados para AdE1A.WT. Así, los residuos con valores $I/I_0 \approx 0$, correspondieron a los residuos del núcleo del motivo E2F (43-LHELY-46) y sus residuos flanqueantes. En tanto, los residuos de la región “*linker 1*” (Figura 3.17 C, en gris claro), presentaron las mismas variaciones vistas para AdE1A- Δ E respecto de AdE1A.WT, es decir que los valores I/I_0 para estos residuos fueron también notoriamente menores en comparación a los mismos residuos en AdE1A.WT ($I/I_0 \approx 0.25$ para AdE1A- Δ L vs $I/I_0 \approx 0.43$ para AdE1A.WT).

En la región conteniendo la mutación del núcleo del motivo LxCxE (122-LTCHE-126) por el parche de tetra-alaninas, los valores de razones de intensidad para los residuos asignados en esta variante, permanecieron altos (V119, I120, D121, A122, A127, G128, F129 y S132, D133, D134 y E135 que son residuos de la región acídica), con valor I/I_0 medio de 0.60. Como en el caso de AdE1A- Δ E, esto indicaría que la interacción de dicho sitio se encuentra efectivamente anulada por la mutación y que los residuos flanqueantes que estarían también interactuando con RbAB en AdE1A.WT, lo harían sólo si los residuos centrales del motivo están mediando la unión a RbAB (rol modulador no central en la interacción).

La Tabla 3.7 a continuación resume los valores medios de I/I_0 por región para las tres variantes de AdE1A analizadas (la natural y las mutantes de motivos). Como se puede observar, hubo básicamente tres comportamientos de los residuos acorde a los valores de I/I_0 : los que asumieron valores $I/I_0 \approx 0$, aquellos con valores $0.2 \leq I/I_0 \leq 0.5$ y aquellos con valores $I/I_0 \geq 0.6$. Mientras los primeros representaron regiones de la secuencia de AdE1A con interacciones fuertes con RbAB, confirmadas por análisis cuantitativos/termodinámicos (motivos y regiones flanqueantes), el segundo grupo representa a residuos con posibles interacciones transientes-débiles e inespecíficas, probablemente de carácter hidrofóbico, con aportes energéticos despreciables acorde a los experimentos cuantitativos/termodinámicos (región “*linker1*”, tal vez región C-terminal). El tercer grupo corresponde a regiones sin interacción con RbAB y que permanecen altamente desordenadas y flexibles tras la formación de los complejos con RbAB (región “*linker 2*” y áreas conteniendo a las mutaciones).

Tabla 3.7. Valores promedio de I/I_0 por región para las tres variantes de AdE1A en interacción con RbAB

Regiones en la secuencia de AdE1A	AdE1A.WT	AdE1A-ΔE	AdE1A-ΔL
Motivo E2F y residuos flanqueantes (39-55)	0.01	0.64	0.03
Motivo CBP-TRAM y algunos residuos extras (63-83) "linker 1"	0.43	0.25	0.25
Región rica en prolinas (84-117) "linker 2"	0.66	0.57	0.62
Motivo LxCxE y residuos flanqueantes incluyendo región ácida (118-138)	0.07	0.07	0.60
Región C-terminal (139-146)	0.57	0.47	0.37

Por último, ya se había mencionado si bien no se realizó la asignación de los picos de AdE1A en su forma unida a RbAB, algunos de ellos que aparecían tras formarse los complejos, lograron asociarse a formas unidas de residuos de los motivos E2F y LxCxE. Los picos correspondientes a la forma unida del motivo LxCxE se revelaron en la superposición de los espectros TROSY de [RbAB:AdE1A.WT] y [RbAB:AdE1A-ΔE], pero no se encontraron en el espectro TROSY de [RbAB:AdE1A-ΔL]. En el mismo sentido, el único pico que se encontró fácilmente para la forma unida del motivo E2F se logró al superponer los espectros TROSY de [RbAB:AdE1A.WT] y de [RbAB:AdE1A-ΔL], y no se encontró en el espectro TROSY del complejo [RbAB:AdE1A-ΔE] (Figura 3.18). Estos resultados ayudarán a la comprensión del experimento de titulación presentado a continuación.

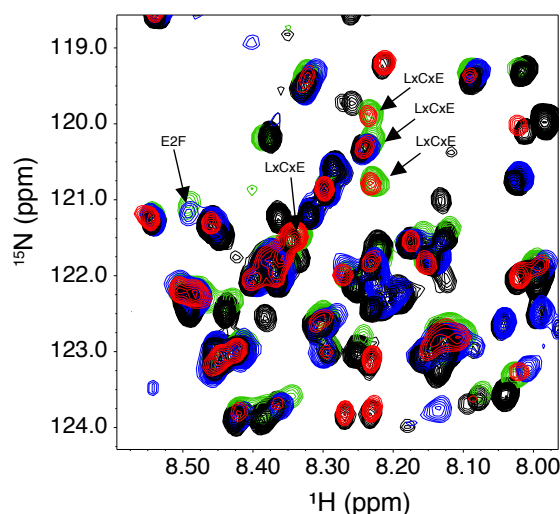


Figura 3.18. Picos correspondientes a los residuos de los motivos E2F y LxCxE en su forma unida. Se muestra la superposición de todos los espectros TROSY de AdE1A variante natural y mutantes en sus formas libres y en complejo 1:1 con RbAB. Los espectros de las formas libres se muestran en color negro, en tanto que el espectro TROSY del complejo [RbAB:AdE1A.WT] está en color verde claro, el del complejo [RbAB:AdE1A-ΔE] en rojo y el del complejo [RbAB:AdE1A-ΔL] en azul. Los picos en el estado unido correspondientes a cada motivo están señalados.

En conclusión, estos experimentos demuestran por un lado que gran parte de la región “linker” permanece desordenada y flexible tras la formación del complejo y que habrían interacciones muy débiles y transientes mediadas por una región que solapa con el motivo TRAM-CBP. Curiosamente, como los valores I/I_0 para esta región “linker 1” en las variantes mutantes son menores que en el caso de AdE1A.WT, esto podría sugerir que en estos casos dicha región establece interacciones inespecíficas probablemente “mas estables” (aunque sin contribución energética) con RbAB. Esto se debería a que el “linker” flexible puede adoptar nuevas conformaciones y establecer un conjunto diferente de interacciones relativamente “más estables” debido a la ausencia de la restricción conformacional impuesta por los dos motivos de interacción en los extremos de AdE1A unidos a RbAB. Por otro lado, los motivos E2F y LxCxE median interacciones con RbAB a través de residuos del núcleo y de residuos flanqueantes a ambos lados de los motivos, por lo que estas interacciones se extienden más allá de los motivos mínimos. No obstante, los residuos flanqueantes presentan sólo contactos accesorios en la unión ya que al no estar presente los residuos centrales, no logran unirse a RbAB.

III.2.B. Determinación de los regímenes de intercambio de los residuos que interaccionan con RbAB. Experimento de titulación del complejo [RbAB:AdE1A.WT]

La naturaleza dinámica de determinados procesos estudiados por RMN influencia la apariencia de los espectros que de ellos se puedan obtener. Para ganar conocimiento acerca de propiedades dinámicas asociadas a una interacción proteína-proteína y la formación de un complejo, es útil llevar a cabo un experimento de titulación de la proteína marcada isotópicamente con el agregado de cantidades crecientes de su compañero de interacción. En este experimento, se van adquiriendo los espectros a medida que progresa la titulación y la información que pueda ser obtenida dependerá de cómo lucen los cambios en estos espectros. Los posibles escenarios se presentan en la Figura 3.19:

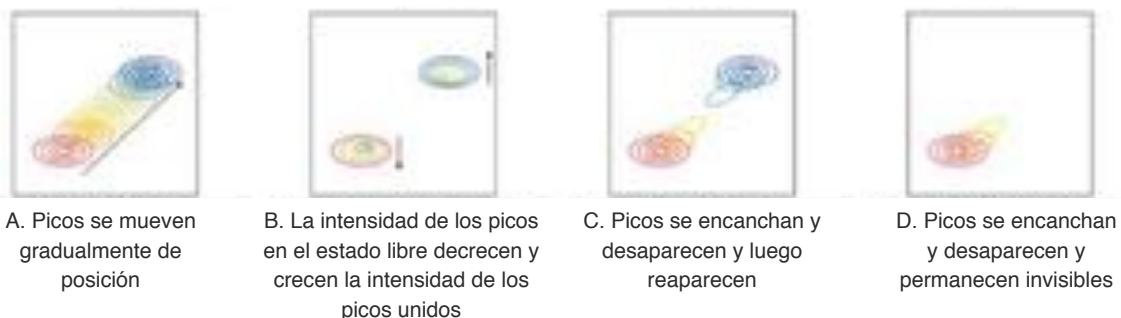


Figura 3.19. Posibles cambios en los picos o señales de resonancia que pueden resultar de un experimento de titulación. Se presentan 4 escenarios factibles (Adaptado de [204]).

- a. Se produce un movimiento gradual de los picos con concentraciones crecientes de ligando, característico de complejos con una razón de intercambio rápido ($k_{int} = k_{on} [L] + k_{off}$) entre el estado libre y en complejo en comparación con la diferencia de sus frecuencias de sus resonancias ($k_{int} > \Delta\omega$). La posición de estos picos en *régimen de intercambio rápido* en cada punto de la titulación es un promedio ponderado de las posiciones de los picos en estado libre y en el estado unido. En general aunque no rigurosamente, si la interacción posee una K_D mayor a 10^{-5} M (baja afinidad), los picos estarán en este régimen de intercambio rápido [232] (Figura 3.19 A).
- b. Los picos del estado libre desaparecen gradualmente sin cambiar de posición mientras aparecen y crecen en intensidad picos correspondientes al estado unido. En este caso, los picos están en *régimen de intercambio lento* ($k_{int} < \Delta\omega$) y pueden ser observados individualmente. Los picos cambian de intensidad pero no de posición con el progreso de la titulación. Si la interacción tiene una K_D igual o menor 10^{-8} M (alta afinidad), los picos involucrados se encontrarían en intercambios lentos [232] (Figura 3.19 B).
- c. Si las señales se ensanchan y/o desaparecen y luego reaparecen y se agudizan gradualmente con la titulación, entonces estos picos están en un *régimen de intercambio intermedio* ($k_{int} \approx \Delta\omega$). Generalmente desaparecen con la adición sub-estequiométrica del ligando en la titulación y reaparecen cerca de la saturación. Aquellas picos involucrados en interacciones con K_D entre 10^{-6} M y 10^{-7} M generalmente presentan un régimen de intercambio intermedio [232] (Figura 3.19 C).
- d. Las señales se ensanchan y/o desaparecen y permanecen invisibles en concentraciones saturantes del ligando. Esto puede ocurrir en los casos en los que los procesos de relajación transversal están acelerados, por ejemplo debido al gran tamaño del complejo que se forma [204] (Figura 3.19 D).

Con el fin de caracterizar la dinámica de interacción y los regímenes de intercambio para cada residuo en AdE1A.WT a medida que progresa la interacción con RbAB, se llevó a cabo un experimento de titulación TROSY. Para ello se prepararon dos muestras correspondientes a:

- a. AdE1A.WT marcada uniformemente con ^{15}N en estado libre a $525 \mu\text{M}$ en $600 \mu\text{l}$ volumen final.
- b. AdE1A.WT marcada uniformemente con ^{15}N acomplejado con RbAB en relación molar 1:1, a concentración de complejo $525 \mu\text{M}$ en $600 \mu\text{l}$ volumen final.

La adquisición de espectros TROSY se hizo para estas dos muestras considerándolas punto inicial (muestra a) y punto final de la titulación (muestra b). Para los puntos intermedios se adquirieron sus espectros tras hacer cuidadosas mezclas cruzadas. Se adquirieron 14 espectros TROSY correspondientes a cada punto de la titulación (Ver Materiales y Métodos). El resultado de la superposición de los 14 espectros de la titulación de AdE1A.WT con RbAB, se muestra en la Figura 3.20. La calidad de los espectros obtenidos en esta titulación fue alta, debido en parte a las concentraciones de trabajo alcanzadas por los reactivos.

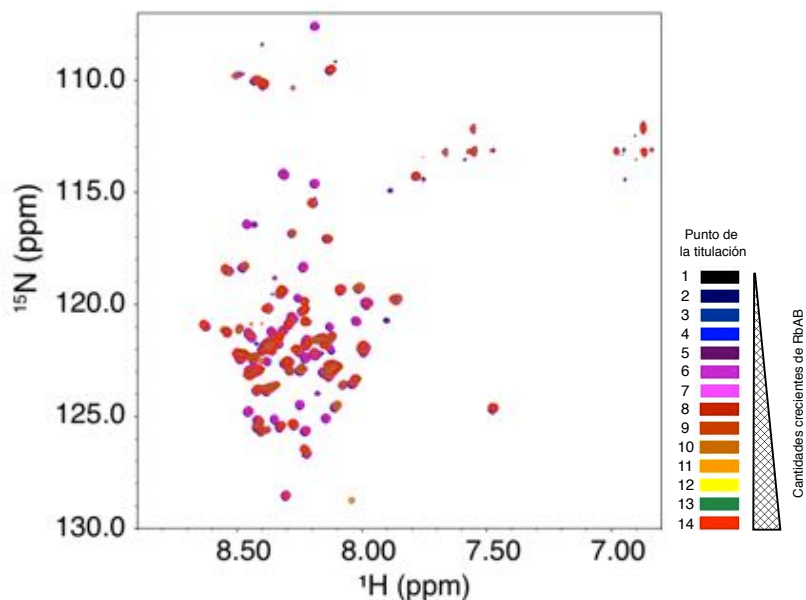


Figura 3.20. Superposición de espectros TROSY de AdE1A.WT uniformemente marcada con ^{15}N libre (en negro, detrás de todos los espectros) y con concentraciones crecientes de RbAB hasta llegar a una relación equimolar -100% unido- a $525\ \mu\text{M}$ final de complejo (en rojo, en la capa superior). A la derecha, el código de colores utilizados para los espectros TROSY de cada punto de la titulación.

En la Figura 3.20 no se consignaron las asignaciones, pero las mismas fueron hechas para 85 residuos en la secuencia, siguiendo el procedimiento aplicado para asignar el espectro TROSY (explicado en Materiales y Métodos). Se tomaron los valores de intensidad de resonancia para cada pico en cada uno de los 14 espectros TROSY (datos tabulados en la Tabla A3.13 de Anexos) y fueron normalizadas a 1, con la relación I_x/I_0 , donde I_x es la intensidad de la señal en cada punto de la titulación e I_0 es el valor de intensidad inicial de la señal, correspondiente al adquirido para la proteína AdE1A.WT en estado libre.

El análisis de los regímenes de intercambio y de la variación de las intensidades de resonancia se hizo residuo por residuo para todos aquellos residuos asignados siempre que no presentase solapamiento considerable con otras señales. Ellos fueron:

- De la región central y flanqueante del motivo E2F, 9 residuos: E39, T42, L43, H44, L46, Y47, V51, D50 y T52. Y del motivo LxCxE, 12 residuos: V119, I120, L122, T123, C124, E126, A127, G128, F129, S132, D134, E135.
- De la región “*linker*”, 21 residuos: 13 residuos de “*linker 1*”: A61, V62, D68, S69, M71, A73, V74, E76, G77, I78, D79, L81 y T82; y 8 residuos del “*linker 2*”: H94, L95, S96, R97, A106, G108, V110, N114.

En el Anexo, bajo el título *Detalles de espectros TROSY por pico para AdE1A.WT y mutantes*, se encuentran en 11 figuras (de Figura A3.7 a Figura A3.17) los espectros TROSY obtenidos en la titulación de AdE1A.WT para estos 42 picos y comparativamente los espectros TROSY por pico adquiridos para AdE1A- Δ E y AdE1A- Δ L en estados libre y acomplejados con RbAB (de los experimentos hechos para calcular sus I/I_0).

Análisis de los residuos del motivo E2F

La Figura 3.21 presenta para un residuo del núcleo del motivo E2F (L43), en el panel izquierdo: las razones de intensidad de las resonancias punto a punto de la titulación en función de la razón molar $[RbAB]/[AdE1A.WT]$ y en el derecho: el detalle del espectro TROSY para el residuo con sus respectivas proyecciones de intensidad de resonancia en el eje X.

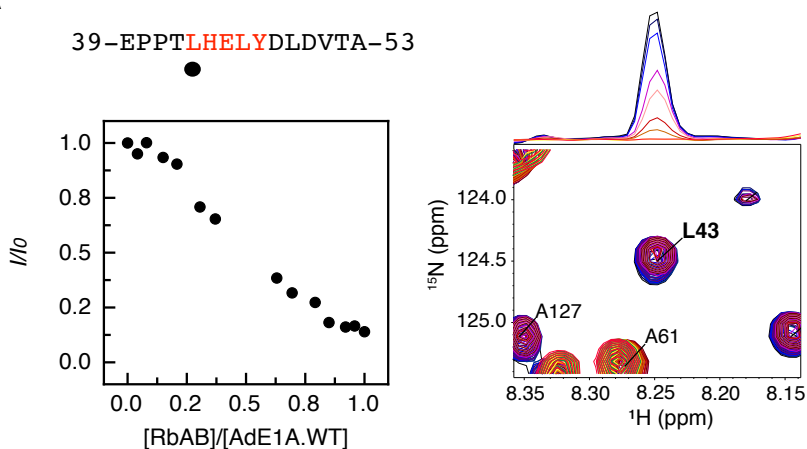


Figura 3.21. Titulación de AdE1A.WT por residuo para L43 del núcleo del motivo E2F. A la izquierda: gráfico de razón de intensidades de resonancias en función de la razón molar $[RbAB]/[AdE1A.WT]$. A la derecha: detalle de la superposición de espectros TROSY para el residuo y por arriba, las proyecciones de intensidad de resonancia a medida que progresa la titulación. El código de colores en la Figura 3.19.

Como es posible observar, para este residuo las intensidades son decrecientes y no hay perturbaciones en la posición del pico a lo largo de la titulación. Además, se observa un comportamiento estequiométrico con la titulación: la intensidad decrece en forma lineal con la relación molar creciente hasta llegar a 1:1, en la que la misma se completa. Esto último es esperado, dado a que la condición estequiométrica (1:1 a 525 μ M) es \approx 5000 veces superior a

la K_D global del motivo (119 nM). Estos resultados son indicativos de que el residuo L43 está en un *régimen de intercambio lento* debido a una alta afinidad por RbAB. Este comportamiento también se presentó para el residuo L46 que es otro residuo del núcleo del motivo E2F.

En el caso de H44 -otro residuo del núcleo del motivo E2F- se observa en la Figura 3.22 A, que el pico pareciera no titular completamente, sin embargo esto es solo aparente, puesto que existe un solapamiento con las resonancias de un “nuevo pico” que aparece tras la titulación y que no es H44 unido a RbAB. Esto se dedujo puesto que: el mismo residuo (H44) en la mutante AdE1A- Δ L desaparece completamente al formarse el complejo equimolar con RbAB y el “nuevo pico” no aparece en este caso (Figura 3.22 B). Además, en la misma región del espectro de la mutante AdE1A- Δ E -que carece del residuo H44- este “nuevo pico” sí aparece, sólo tras formarse el complejo equimolar con RbAB (Figura 3.22 C). Este “nuevo pico” corresponde a uno de los identificados en la Figura 3.18.

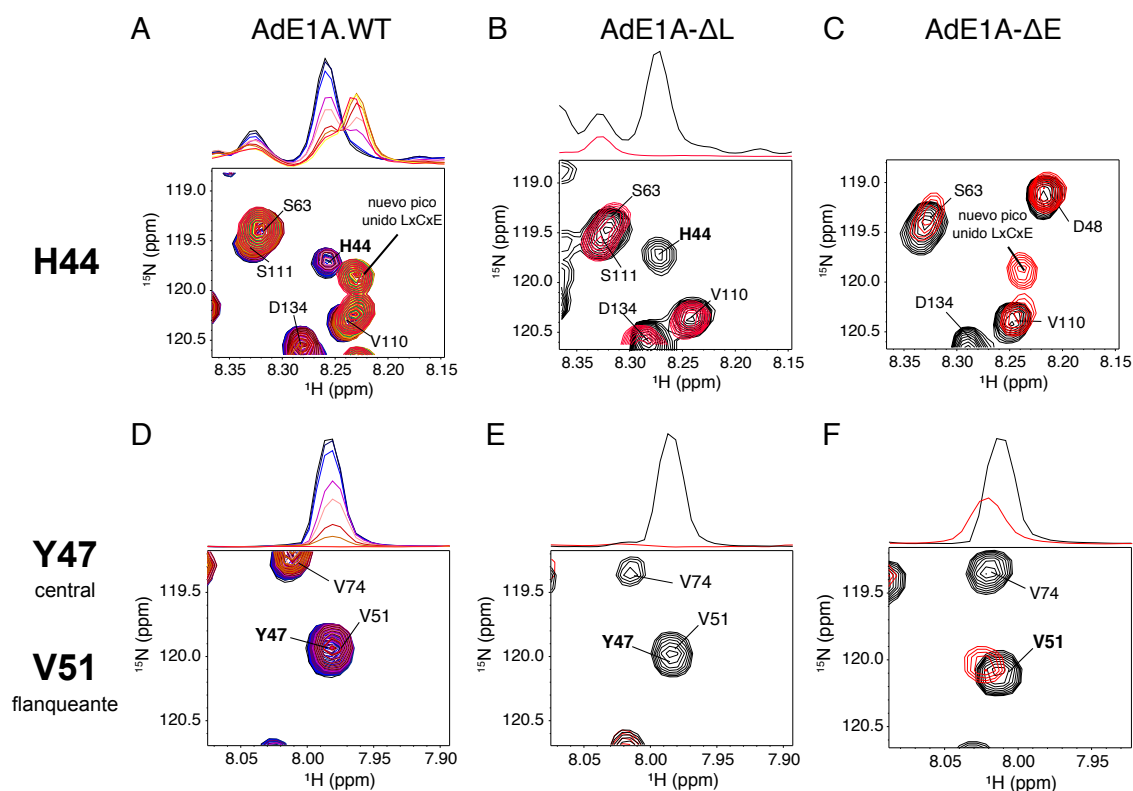


Figura 3.22. Titulación de AdE1A.WT por residuo para H44 e Y47 del núcleo y V51 de la región flanqueante del motivo E2F. A los fines comparativos se muestran para dichos picos los espectros TROSY en las formas libres (en negro) y acomplejadas en relación molar 1:1 con RbAB (en rojo) de la mutante AdE1A- Δ L (segunda columna, B y E) y de la mutante AdE1A- Δ E (tercer columna, C y F).

Otro caso particular es el de Y47 (residuo central del motivo E2F) y V51 (residuo flanqueante) los cuales si bien tienen resonancias solapadas comparten el mismo

comportamiento en la titulación, presentando ambos un régimen de intercambio lento (Figura 3.21'2 D). Y47 y V51 fueron identificados en el mismo pico ya que su valor I_0 duplicó la media para valores I_0 de otros picos individuales (6 vs 3.5 aproximadamente, ver Tabla A3.11 de Anexos) además de que en la mutante AdE1A- Δ E, que no tiene el residuo Y47 por la mutación por Alaninas, el residuo V51 fue identificado e individualizado en la misma zona del espectro (Figura 3.22 F).

En la Figura 3.23 se muestra el comportamiento de un residuo flanqueante (T52) del motivo E2F frente a la titulación de AdE1A.WT con cantidades crecientes de RbAB. Como se puede observar, al igual que L43 del núcleo del motivo E2F, hay un comportamiento estequiométrico con la titulación. Este residuo flanqueante, también presenta un *régimen de intercambio lento*. El mismo comportamiento se presentó para los residuos E39, T42 y D50 flanqueantes del motivo E2F.

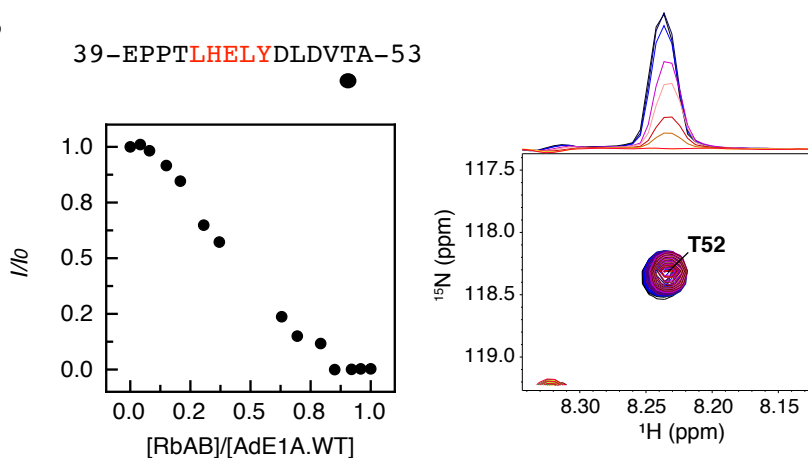


Figura 3.23. Titulación de AdE1A.WT por residuo para T52 del motivo E2F. A la izquierda: gráfico de razón de intensidades de resonancias en función de la razón molar $[RbAB]/[AdE1A.WT]$. A la derecha: detalle de la superposición de espectros TROSY para el residuo y por arriba, las proyecciones de intensidad de resonancia a medida que progresa la titulación. El código de colores en la Figura 3.19.

Estos resultados muestran que las señales de resonancia tanto de los residuos centrales como de los residuos flanqueantes del motivo E2F, pierden intensidad pero no cambian de posición con de la progreso titulación con RbAB y siguen un comportamiento estequiométrico en la titulación. Además, si bien no se han identificado los nuevos picos que aparecen con la formación del complejo $[RbAB:AdE1A.WT]$, al menos uno de ellos correspondería a la forma unida a RbAB de alguno de estos residuos del motivo E2F (Figura 3.18). Probablemente otros picos unidos para esta región pudieron no identificarse fácilmente debido al solapamiento de señales. En su conjunto, esto evidencia que tanto residuos del núcleo como flanqueantes del motivo E2F se encuentran en un *régimen de intercambio lento* lo

que es consistente con los experimentos termodinámicos, que revelaron una afinidad nanomolar ($K_D = 119 \text{ nM}$ -por EF-) para la interacción de dicho motivo con RbAB.

Análisis de los residuos del motivo LxCxE

En la Figura 3.24 se observa el comportamiento frente a la titulación de un residuo del núcleo del motivo LxCxE (C124) en AdE1A.WT. Al igual que el residuo L43 del núcleo del motivo E2F, aquí también se observa un comportamiento estequiométrico y los resultados son indicativos de que este residuo se encuentra en un *régimen de intercambio lento*. Esto es indicativo de una alta afinidad del motivo LxCxE por RbAB. El mismo comportamiento lo presentan L122, T123, E126 que constituyen los otros residuos núcleo de este motivo.



Figura 3.24. Titulación de AdE1A.WT por residuo para C124 del núcleo del motivo LxCxE. A la izquierda: gráfico de razón de intensidades de resonancias en función de la razón molar $[\text{RbAB}]/[\text{AdE1A.WT}]$. A la derecha: detalle de la superposición de espectros TROSY para el residuo y por arriba, las proyecciones de intensidad de resonancia a medida que progresa la titulación. El código de colores en la Figura 3.19.

Para los residuos en la región flanqueante a este motivo, se aprecia el mismo comportamiento antes descrito para C124. En la Figura 3.25 se aprecia a modo de ejemplo, cómo el residuo flanqueante S132 pierde intensidad de manera lineal con la titulación, la cual se completa en la relación molar 1:1 con RbAB. Para la misma no se observa perturbación en la posición del pico. Esto es indicativo de que también se encuentra en *régimen de intercambio lento*, como también ocurre para los residuos V119, I120, A127, G128, F129, D134 y E135. Aquí se demuestra que estos residuos flanqueantes al motivo de interacción, siempre que estén presentes los residuos centrales LxCxE, también median una interacción de alta afinidad con RbAB. Esto es consistente con los resultados de los análisis termodinámicos que

mostraron que el motivo LxCxE se une con una $K_D = 104\text{nM}$ a RbAB y que dicha afinidad se incrementa levemente en presencia de los residuos de la región acídica ($K_D = 72\text{ nM}$).

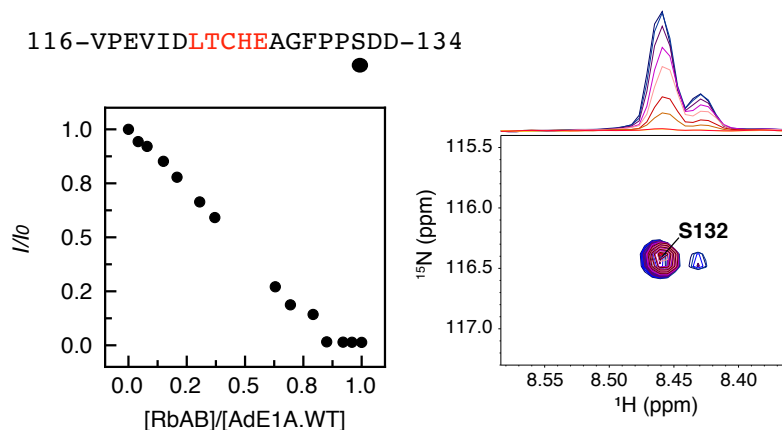


Figura 3.25. Titulación de AdE1A.WT por residuo para S132 flanqueante al motivo LxCxE. A la izquierda: gráfico de razón de intensidades de resonancias en función de la razón molar $[RbAB]/[AdE1A.WT]$. A la derecha: detalle de la superposición de espectros TROSY para el residuo y por arriba, las proyecciones de intensidad de resonancia a medida que progresa la titulación. El código de colores en la Figura 3.19.

Como se observa en la Figura 3.18, varios picos emergen en condiciones de saturación cuando AdE1A forma el complejo con RbAB, y algunos de ellos corresponden a las formas unidas o de los residuos centrales y/o flanqueantes al motivo LxCxE -aunque no se hayan identificado específicamente a cuáles corresponden-. En sumatoria, y como se concluyó para el motivo E2F, estos resultados evidencian que tanto residuos del núcleo como flanqueantes del motivo LxCxE se encuentran en un *régimen de intercambio lento* lo que es consistente con los resultados obtenidos de los análisis cuantitativo/termodinámicos que demuestran su alta afinidad por RbAB.

Análisis de los residuos de la región “linker” en AdE1A

En el análisis de los comportamientos de las señales de resonancia frente a la titulación de AdE1A.WT con RbAB, se vio que los correspondientes a residuos de la región “linker 1”, evidenciaban lo siguiente:

- Mayores perturbaciones en la posición (δ) del pico en comparación a los residuos en los motivos lineales de interacción. Estos corrimientos en el espectro ocurren en forma gradual a medida que progresa la titulación.
- Una disminución parcial en las intensidades de las señales hacia la condición 1:1. Para estos residuos por lo tanto, no se observa una titulación completa en la condición equimolar entre RbAB y AdE1A.WT ($I_f \approx 0.5 I_0$).

En la Figura 3.26 se presenta el caso del residuo D68, el cual forma parte de esta región de “*linker 1*”. En las proyecciones de intensidad de resonancia se observa claramente como la señal se desplaza con el progreso de la titulación, a la vez que disminuye levemente en la intensidad. A su vez no se observan nuevos picos que pudieran corresponder a formas unidas a RbAB. Estas observaciones también se replican para residuos en la misma región como residuos A61, V62, S69, M71, A73, V74, E76, G77, D79, L81 y T82 (Figuras A3.12 a 15 de Anexos), e indican que dichos residuos se encontrarían en un *régimen de intercambio intermedio a rápido*. Los datos obtenidos de los análisis cuantitativos/termodinámicos señalan que el “*linker*” no aporta significativamente a la interacción, por lo que estas observaciones a nivel estructural corresponderían a interacciones transientes o débiles.

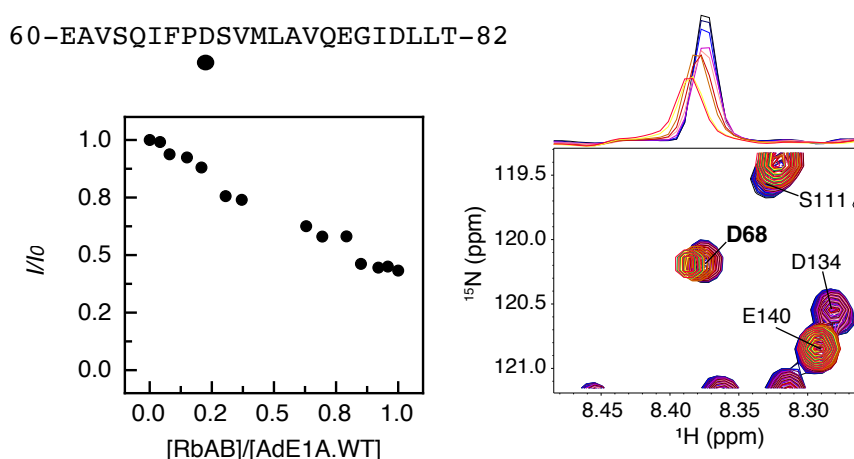


Figura 3.26. Titulación de AdE1A.WT por residuo para D68 de la región “*linker 1*”. A la izquierda: gráfico de razón de intensidades de resonancias en función de la razón molar [RbAB]/[AdE1A.WT]. A la derecha: detalle de la superposición de espectros TROSY para el residuo y por arriba, las proyecciones de intensidad de resonancia a medida que progresa la titulación. El código de colores en la Figura 3.19.

Para los residuos de la región “*linker 2*” que pudieron ser individualizados y analizadas sus comportamientos con la titulación (H94, L95, S96, R97, A106, G108, V110 y N114), se encontró que, aunque experimentaron disminuciones en la intensidad de sus señales con el progresivo agregado del titulante, las variaciones fueron mínimas y despreciables en comparación con el comportamiento de otros residuos como los de los motivos de interacción y los de la región “*linker 1*”. Un ejemplo se presenta en la Figura 3.27, que muestra la mínima variación de la señal de resonancia correspondiente a S96 con la titulación. Para esta región se demuestra por lo tanto una clara ausencia de todo de tipo de interacción con RbAB. (Ver Figuras A3.16 y 17 de Anexos).

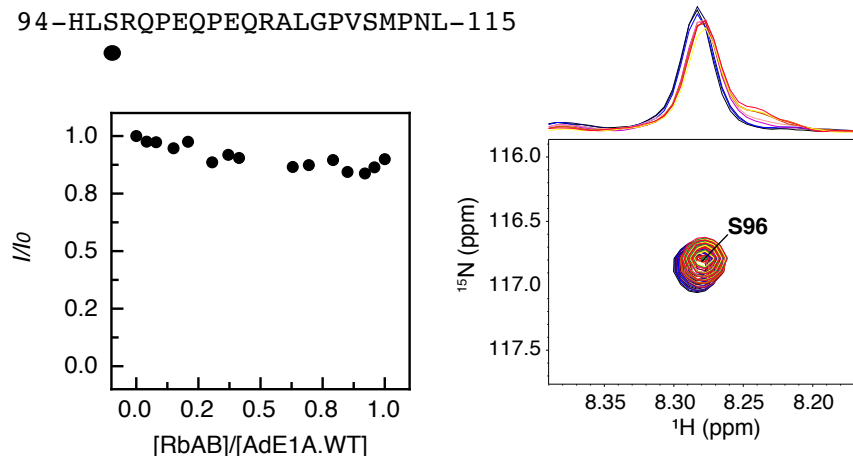


Figura 3.27. Titulación de AdE1A.WT por residuo para S96 de la región "linker21". A la izquierda: gráfico de razón de intensidades de resonancias en función de la razón molar $[RbAB]/[AdE1A.WT]$. A la derecha: detalle de la superposición de espectros TROSY para el residuo y por arriba, las proyecciones de intensidad de resonancia a medida que progresa la titulación. El código de colores en la Figura 3.19.

En conclusión, los residuos de la región "linker" muestran dos comportamientos frente a la titulación, mientras la región "linker 2" se muestra totalmente libre de interacciones, la región "linker 1" evidenciaría interacciones transientes o débiles. No obstante a esto último, las contribuciones energéticas de esta región a la interacción son despreciables, como se ha visto por los resultados de la termodinámica.

En el caso de las mutantes AdE1A- ΔE y AdE1A- ΔL , no fue posible establecer propiedades dinámicas de los residuos en la formación de sus complejos con RbAB porque no se hicieron experimentos de titulación sobre ellos. Para estas mutantes solo se obtuvieron los espectros TROSY de las especies libres y acomplejadas en relación molar 1:1 con RbAB, como fue presentado en la sección anterior. Sin embargo, en perspectiva con los resultados del experimento de titulación de AdE1A.WT, es posible hacer una serie de inferencias:

- Como los valores I/I_0 de las señales correspondientes a los residuos del núcleo y áreas flanqueantes al motivo E2F (en la mutante AdE1A- ΔL) y al motivo LxCxE (en la mutante AdE1A- ΔE) son cercanos a cero y los picos no parecen registrar perturbaciones en sus posiciones (δ) (Figura A3.5 y A3.6 de Anexos), éstos residuos también estarían bajo un *régimen intercambio lento* con RbAB, implicando interacciones de alta afinidad con esta proteína.
- Para los residuos correspondientes al "linker 1" en ambas mutantes, también se observaron mínimos desplazamientos en las posiciones de los picos aunque presentaron valores I/I_0 más bajos comparados a los valores I/I_0 en el punto final de la

titulación de AdE1A.WT (totalmente unido a RbAB). Esto podría indicar que los mismos se encuentren en un régimen de intercambio intermedio donde esas interacciones débiles y transientes puedan ser mas estables, aunque no aporten energéticamente a la interacción con RbAB.

- Los residuos de la región “*linker 2*” en ambas mutantes, presentaron comportamientos similares a los de AdE1A.WT tras producirse la interacción con RbAB. Ninguno de ellos experimentó un cambio significativo en intensidades de señal ni su posición, indicativos de que los mismos no interactúan y permanecen en una conformación desordenada tras la formación de los complejos.

CONCLUSIONES III

En este capítulo se presentó un análisis estructural de alta resolución por RMN de la proteína AdE1A, en sus variantes natural y mutantes.

En primer lugar los experimentos 3D permitieron la asignación de más del 70% de los residuos en los tres casos. Para las tres variantes, el espectro HSQC evidenció el desorden intrínseco en estas proteínas. El análisis de los desplazamientos químicos secundarios ($\Delta\delta$) reveló que AdE1A carece de estructuras secundarias evidentes a lo largo de toda su extensión y es preponderantemente desordenada en estado libre en solución. Los resultados del análisis δ_{2d} fueron concordantes, demostrando que sus propensiones a estructuras secundarias canónicas es muy baja. Como se conoce que AdE1A adopta estructuras consolidadas en su forma unida a RbAB (el motivo E2F forma una α -hélice y el motivo LxCxE formaría una lámina- β) y otras proteínas (el motivo TRAM-CBP forma α -hélices unida a CBP/p300 [76, 82]), estas mismas regiones podrían presentar elementos de estructura secundaria pre-estructurados, sin embargo no hay evidencia de ello. Esto permite comprender que los mecanismos de interacción sitio-específica mediados por los motivos E2F y LxCxE de AdE1A con RbAB, acuerdan a un modelo de ajuste inducido, rindiendo interacciones energéticamente muy estables. Por otro lado, con el estudio de la dinámica conformacional sobre AdE1A, se pudo caracterizar esta proteína como una IDP con un carácter altamente flexible a lo largo de toda su extensión. Esta característica es sumamente importante porque permite comprender como el “*linker*” puede ser capaz de lograr acomodarse para que se logre el contacto simultáneo de los dos motivos de interacción en AdE1A sobre sus sitios complementarios en distintas caras de la superficie de RbAB.

En segundo lugar, se planteó como objetivo el análisis estructural a nivel atómico de la interacción entre AdE1A y RbAB. Los experimentos TROSY confirman los dos sitios conocidos en AdE1A que interactúan con RbAB (los motivos E2F y LxCxE), cuyos residuos centrales establecen contactos con RbAB, como también residuos flanqueantes a ambos lados de dichos motivos, definiendo así secuencias de interacción más largas que los motivos mínimos. Sin embargo las interacciones mediadas por residuos flanqueantes tendrían sólo rol modulador, puesto que éstos no se unen a RbAB cuando los residuos centrales del motivo están ausentes. Como la interacción de un motivo a RbAB, se mantiene en ausencia del motivo complementario, se evidencia su carácter independiente en la interacción. Por otra parte, se observó que la región del “*linker*” que comprende el motivo TRAM-CBP, establecería interacciones con RbAB de carácter débil o transitorio, que no aportan energéticamente, como

lo demostraron los experimentos de termodinámica. Dicho comportamiento había sido observado por Ferreon y colaboradores, quienes atribuyeron a posibles contactos hidrofóbicos débiles con el “pseudo-LxCxE” formado por los residuos 71-MLAVQE-76 y cuya interacción era fácilmente competida por el dominio TAZ2 de CBP [76]. Es llamativo como en las variantes mutantes, dicha región mediaría interacciones transientes “más estables” con RbAB probablemente por la falta de restricciones conformacionales sobre el “*linker*” de AdE1A que le permiten mayor movilidad.

Con el experimento de titulación, se evidenció que tanto los residuos del núcleo como flanqueantes a los motivos E2F y LxCxE, presentan regímenes de intercambio lento, mediando interacciones de alta afinidad, consistente con los datos obtenidos del análisis cuantitativo-termodinámico. La región “*linker 1*”, en tanto se encuentra en un régimen de intercambio intermedio a rápido, evidenciando interacciones transientes con RbAB con aporte energético poco significativo. No obstante a lo observado para el “*linker*” sobre el motivo TRAM-CBP, es destacable como la mayor parte del “*linker*” permanece prácticamente inalterado tras la formación del complejo, manteniendo su grado de alto desorden y flexibilidad, tanto en el complejo con la variante natural y con las mutantes. Esta característica permite describir al complejo [RbAB:AdE1A] como un complejo “*fuzzy*”, lo que le aportaría ventajas funcionales a AdE1A como proteína “*hub*”.

CAPITULO IV

Modelado de la interacción entre AdE1A y su blanco específico RbAB

En una somera recopilación de resultados hasta ahora presentados y que permiten echar luz sobre el mecanismo global de interacción de AdE1A con RbAB y en consecuencia del desplazamiento efectivo de E2F celular por parte de la proteína viral, se ha visto que:

- a. AdE1A es una proteína eminentemente desordenada y extendida, sin contenido distintivo de estructura aunque caracterizada por una gran plasticidad conformacional. Esta sumatoria de cualidades dotan a AdE1A de una gran flexibilidad.
- b. AdE1A interacciona con RbAB con una estequiometría 1:1, por una interacción bipartita mediada por dos motivos de interacción conocidos E2F y LxCxE, dando lugar a un complejo en el que AdE1A no experimenta grandes cambios conformacionales.
- c. Cada motivo de interacción es capaz de unirse de manera independiente y con afinidad nanomolar a RbAB. En AdE1A, ambos motivos unidos por la región “*linker*” interaccionan en simultáneo con RbAB con una afinidad potenciada 4000 veces en comparación a la interacción de los motivos individuales, y sólo de esta manera AdE1A es capaz de desplazar a E2F2 de su complejo endógeno con RbAB. La potenciación de la afinidad observada, no está mediada por contactos energéticamente estabilizantes en la región “*linker*” ni por un mecanismo alostérico entre los sitios de interacción.
- d. Si bien los motivos y sus residuos flanqueantes en AdE1A interaccionan fuertemente con sus superficies complementarias en RbAB, el desorden y la flexibilidad de la proteína viral se mantienen en gran medida sobre la región “*linker*” tras formarse el complejo con RbAB.

Con esta serie de conclusiones presentes, se puede hipotetizar que la región “*linker*” es responsable de generar la potenciación en la afinidad global del sistema y que su desorden y flexibilidad, son sus elementos clave que le permiten cumplir dicho rol en el complejo. En esto se sustenta el tercer modelo planteado en esta tesis para explicar la interacción entre AdE1A y RbAB. Este modelo (Modelo C) establece que una vez unido el primer motivo de AdE1A a RbAB, el “*linker*” produce un aumento efectivo de la concentración del segundo motivo en la región cercana a su sitio de interacción, favoreciendo su interacción (Figura 4.1)

El enfoque planteado en este último capítulo permitirá confluír de modo sencillo todas las evidencias aportadas a lo largo de esta tesis que permitan entender las bases mecánicas de la interacción en estudio y extensivamente aporten conocimientos acerca de los mecanismos que utilizan los virus para subvertir la maquinaria en una célula infectada.

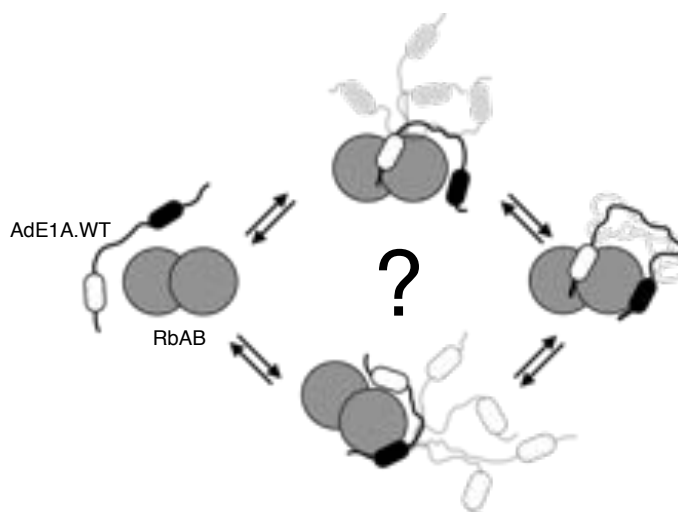


Figura 4.1: Modelo de interacción entre AdE1A y RbAB a validar en el presente capítulo.

OBJETIVOS ESPECIFICOS:

- I. Presentar modelos teóricos obtenidos por simulación para los complejos [RbAB:AdE1A.WT], [RbAB:AdE1A- Δ E] y [RbAB:AdE1A- Δ L], apuntando a generar mayor nivel de comprensión en ellos pero también a probar si la simplificación de los sistemas (en términos de los parámetros tenidos en cuenta para realizar modelados) es útil para describir sus comportamientos.
- II. Analizar la estructura del complejo [RbAB:AdE1A.WT] por SAXS, sumando una última evidencia experimental, que permita evaluar la consistencia de las evidencias biofísicas, termodinámicas y estructurales en alta resolución (RMN) acerca del sistema en estudio.
- III. Presentar un modelado simplificado de la interacción entre AdE1A.WT y RbAB que de cuenta de las características principales del complejo y permita probar esta hipótesis central a fin de comprender el mecanismo de interacción entre AdE1A y RbAB.

Los experimentos presentados en este capítulo se llevaron a cabo en colaboración multilateral. Las muestras fueron proporcionadas por este laboratorio en tanto que la recolección de los señales de dispersión SAXS fueron realizados en las instalaciones del Laboratorio Europeo de Biología Molecular (EMBL) en Hamburgo, Alemania y análisis de los datos de SAXS fueron hechos por el del Dr. Pau Bernadó del *Centre de Biochimie Structurale* de la *Université de Montpellier*. Francia . Los modelados teóricos fueron elaborados por el Dr. Juan Cortes y su grupo del *LAAS-CNRS* de la *Université de Toulouse*, también en Francia. Una presentación más detallada de los experimentos se encuentra en Materiales y Métodos en tanto aquí se presentan los datos y conclusiones contextualizados en el presente trabajo.

IV.1. Modelado teórico híbrido de los complejos entre AdE1A.WT y mutantes con RbAB

Muchos complejos proteicos, sobre todo los compuestos por IDPs, son altamente dinámicos y esta característica es esencial para que medien los procesos biológicos en los que están involucrados. Entender como están conformados, es por lo tanto el primer paso para comprender sus funciones.

El desorden constitutivo y alta flexibilidad de una IDP, que puede preservarse tras formar un complejo (así como también las rápidas fluctuaciones en proteínas globulares y las reconfiguraciones moleculares internas lentas a gran escala) implican dinámicas que deben ser tenidas en cuenta al momento de describir estos sistemas. Un concepto muy utilizado para hacer descripciones de sistemas dinámicos es el de “*ensamble de conformaciones*” [233] o “*estructuras súper-terciarias*” [234]. Con el fin de generar un *ensamble de conformaciones* que se acerque a una serie de conformaciones realistas que estos complejos proteicos pueden estar adquiriendo, en situaciones en las que no se dispone de información completa en alta resolución acerca de los mismos, es a partir del modelado teórico (simulación) de los mismos. De esta manera se genera un “*pool*” de complejos cubriendo el espacio conformacional disponible de los que es factible obtener información acerca de sus propiedades [235]. Estas propiedades del ensamble pueden contrastarse con datos del complejo obtenidos por métodos experimentales (RMN, RX, CD, SEC-SLS, etc.), los cuales deben corresponderse para generar confiabilidad en el modelo teórico desarrollado.

Dependiendo de la información disponible al momento de construir estos modelos, el modelado puede considerarse: *modelado ab initio*, donde ninguna información estructural es conocida *a priori* para ninguno de los componentes de un complejo; o un *modelado híbrido*, en el que se utilizan estructuras pre-existentes de alta resolución de dominios o subunidades proteicas, o de alguno de los componentes del complejo proteico [236].

Se generaron modelos teóricos por simulación de los complejos [RbAB:AdE1A.WT], [RbAB:AdE1A- Δ E] y [RbAB:AdE1A Δ L] para evaluar la hipótesis de que es posible modelar dichos sistemas a partir de información estructural limitada, pero que genere resultados consistentes con los aportes biofísicos y termodinámicos hallados en sus caracterización previas.

Aquí se utilizó un modelado híbrido de para generar los ensamblajes de estos complejos. Para RbAB, que es una proteína globular altamente ordenada, se utilizó la estructura cristalina

disponible (PDBID: 3POM) [191]. En cambio, para modelar AdE1A se utilizó como dato de entrada, las secuencias reales de las tres variantes y las estructuras cristalinas de los motivos unidos a RbAB, el de E2F (PDBID: 2R7G) [84] y del homólogo LxCxE-E7 de HPV (PDBID: 1GUX)[135]. La información estructural codificada en estas secuencias reales se tuvieron en cuenta para generar distintas conformaciones posibles evitando choques estéricos. Para generar el complejo [RbAB:AdE1A.WT], los dos motivos de interacción E2F y LxCxE se consideraron fijos y estáticos a RbAB, en tanto el “linker” y los pequeños segmentos libres en los extremos N- y C-terminal se consideraron cadenas polipeptídicas que adaptan su conformación a las restricciones impuestas por los motivos unidos y por aquellas que provienen de bibliotecas de fragmentos peptídicos –tomando en consideración la secuencia real de AdE1A- (ver Materiales y Métodos). Para las variantes mutantes, un solo motivo se consideró unido y el resto de la cadena se dejó libre a fluctuaciones que muestrean conformaciones posibles. Ningún modelo consideró, efectos de solventes, interacciones electrostáticas o hidrofóbicas fuertes o débiles más allá de las posiciones que se fijaron para los motivos. Los ensambles de estos tres complejos de los que se modelaron 10000 estructuras posibles, se representan en la Figura 4.2, mostrando 30 conformaciones elegidas azarosamente:

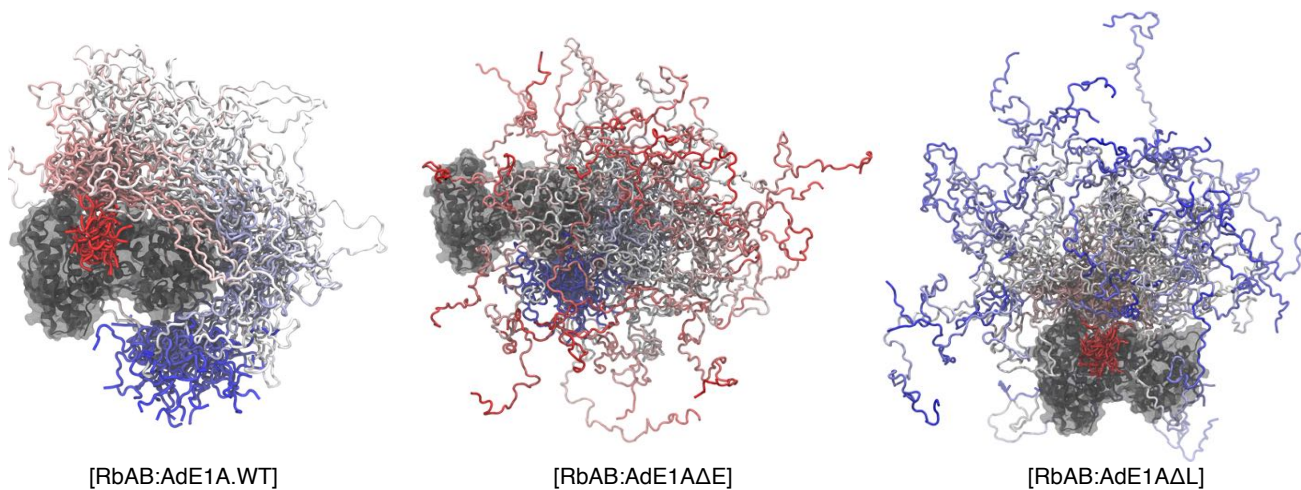


Figura 4.2: Ensamblajes conformacionales modelados de los complejos de AdE1A.WT y mutantes con RbAB. Se representa una sola estructura para RbAB (gris) y 30 copias de AdE1A en distintas conformaciones para el mismo ensamble elegidos azarosamente. En color rojo el extremo N-terminal por donde se ubica el motivo E2F y en color azul el extremo C-terminal de AdE1A, donde está ubicado el motivo LxCxE.

De los mismos se calcularon radios hidrodinámicos (R_H) medios (Materiales y Métodos) de 1000 conformaciones al azar para cada ensamble a 20 y 25°C. Para [RbAB:AdE1A.WT] se seleccionaron otras 1000 conformaciones extras elegidas por un Método de Optimización de Ensamblajes –EOM-, los cuáles constituyen una sub-población cuyas regiones “linker” se encuentran en un estado conformacional más extendido que la media de todas las estructuras

simuladas (y que fueron las conformaciones seleccionadas por datos de SAXS- ver sección siguiente-) a ambas temperaturas. La Tabla 4.1 muestra los datos obtenidos y los presenta comparativamente con los datos experimentales de R_H obtenidos por SEC (Capítulo II).

Tabla 4.1 Radios Hidrodinámicos experimentales y simulados a 2 temperaturas para complejos de AdE1A con RbAB

	SEC		Modelado teórico			
	PM _{TEO} (kDa)	R_H (nm) ¹	R_H (nm) ²	R_H (nm) ² EOM	R_H (nm) ³ EOM	
[RbAB:AdE1A.WT]	54.6	3.25	3.71	3.79	3.27	3.36
[RbAB:AdE1A-ΔE]	54.2	3.32	4.12	-	3.64	-
[RbAB:AdE1A-ΔL]	54.3	3.40	4.07	-	3.60	-

¹ el radio hidrodinámico experimental se calculó de la relación $\log R_H = -0.204 + 0.357 \log PM_{apSEC}$, estimados por SEC a temperatura de 25°C

² el radio hidrodinámico fue calculado a 20 °C

³ el radio hidrodinámico fue calculado a 25 °C

Como se puede observar en la Tabla 4.1, entre los R_H calculados para los modelos se observa un mayor grado de compactación de los complejos con mayor temperatura. No obstante estas diferencias son menores (del 12% entre sí). Mas allá de esta observación, es notable el excelente grado de concordancia para los valores de R_H calculados a 25°C y el valor de R_H experimental obtenido por SEC también llevado a cabo a la misma temperatura para el complejo [RbAB:AdE1A.WT] (menos del 4% de diferencia entre los valores mas alejados entre sí), lo que valida a este modelado que considera al linker una cadena polipeptídica flexible y sin interacciones entre el “*linker*” y RbAB, como lo sugieren los análisis termodinámicos (Figura 4.3).

Para los complejos formados por las mutantes, también hay un muy buen acuerdo entre los valores de R_H calculados para los modelos a 25°C y los valores experimentales (diferencias menores al 10%). En comparación al caso del complejo formado por AdE1A.WT, la concordancia para los complejos mutantes es levemente menor, siendo los R_H experimentales menores a los calculados. Quizás estas diferencias podrían radicar en que los “*linkers*” en estos complejos podrían estar mediando interacciones transientes “mas estabilizadas” dado el mayor grado de libertad conformacional del mismo, como se había observado en los experimentos de RMN. La media poblacional para el R_H de estos complejos podría pensarse entonces más desplazada hacia formas más compactas, que las que son muestreadas azarosamente para calcular los valores R_H de los modelos simulados, los cuales a su vez no consideran ningún tipo de interacción extra a los establecidos por los motivos E2F o LxCxE (Figura 4.3). Estos

experimentos validan la idea de que estas “interacciones transientes” vistas en los complejos mutantes, no se encuentran presentes o son aún más insignificantes al estar unidos ambos motivos en el complejo con la variante natural.



Figura 4.3. Modelo de interacción entre AdE1A y RbAB. En la interacción bivalente de AdE1A.WT y RbAB, un “linker” de carácter desordenado y flexible extendido, permite el contacto de ambos motivos. El “linker” no establecería interacción alguna con RbAB y se encuentra extendido preservando sus propiedades entrópicas en el complejo formado. En los modelos de complejos mutantes, la libertad conformacional del “linker” a consecuencia de la ausencia de un segundo motivo de interacción a RbAB, le permitiría adoptar múltiples conformaciones poblando particularmente aquellos estados más compactos probablemente dados por interacciones muy transientes-inespecíficas, pero sin aporte energético.

Concluyendo, estos modelados teóricos han demostrado ser una herramienta útil a la hora de describir y entender el comportamiento de los complejos en estudio, validando un comportamiento “entrópico” de la región “linker”. En el caso de [RbAB:AdE1A.WT] esto es esencial, pues el mismo será utilizado para analizar los datos obtenidos de SAXS para este complejo.

IV.2. Análisis de SAXS del complejo [RbAB:AdE1A.WT]

La técnica de Dispersión de Rayos X de ángulo pequeño (SAXS, por sus siglas en inglés) constituye un método biofísico de resolución media que es útil para estudiar la forma general y tamaño de macromoléculas en solución. Esta técnica es poderosa al permitir ajustar los datos obtenidos, a perfiles derivados de modelos basados en estructura ya sea de RX o RMN o de simulaciones de estas macromoléculas (o complejos) en solución que están siendo estudiados. El SAXS permite una caracterización cuantitativa de la dispersión conformacional que pueden presentar en solución moléculas completa o parcialmente desordenadas, incluyendo IDPs o proteínas multi-dominios con regiones “linker” flexibles –por caso complejos- y llegar a parámetros geométricos simples que los describan, así como probables configuraciones en solución [237, 238].

En este caso, la técnica de SAXS fue utilizada para conocer y describir el estado y forma global que adquiere en solución el complejo [RbAB:AdE1A.WT]. Las evidencias

aportadas por el SAXS permitirán por caso, elucidar si las conclusiones hasta ahora arribadas por los otros experimentos provistos (biofísicos, termodinámicos y estructurales en alta resolución por RMN) así como el modelado teórico de la interacción, son consistentes con una estructura en la que el “*linker*” permanecería altamente desordenado, flexible, extendido y sin interacciones aun cuando AdE1A.WT se encuentra unido a RbAB a través de sus dos motivos de interacción.

En primer lugar, se realizó la adquisición de los datos de dispersión para una serie de muestras de concentraciones crecientes de RbAB y del complejo [RbAB:AdE1A.WT]. La preparación y procedimientos realizados se encuentran detallados en Materiales y Métodos. Luego, los perfiles de SAXS obtenidos se compararon, en el caso de RbAB, con su estructura cristalina (PDBID: 3POM) y con una versión optimizada de la misma. En el caso del complejo [RbAB:AdE1A.WT], los perfiles experimentales se compararon con el ensamble de conformaciones (el “*pool*” o total de 10000 conformaciones) obtenido por simulación (sección anterior) y con una sub-población seleccionada de ese ensamble. La Figura 4.4 resume gráficamente los resultados de estos análisis y los presenta de modo comparativo.

Para el caso de RbAB, la estructura cristalina monomérica de la que se disponía ajustó muy bien a los perfiles de SAXS experimentales (Figura 4.4 A), especialmente luego de haberle adicionado los fragmentos que no estaban resueltos –por tener densidades electrónicas perdidas (regiones enlazantes y segmentos N- y C- terminales)- y optimizándola con una corrida corta de dinámica molecular (MD) empleando el servidor Allos-Mod (Figura 4.4 D). El RMSD entre la estructura cristalográfica y la estructura refinada al que se ajustaron los perfiles de SAXS fue de 1.7 Å. Por tanto RbAB en solución, es equivalente (en la resolución de SAXS) a la estructura presente en su forma cristalizada.

Para el complejo [RbAB:AdE1A.WT] (Figura 4.4 B), dado que no se disponía de estructuras de alta resolución (RX o RMN), para tratar de ajustar los perfiles de SAXS se utilizaron: el total del ensamble de conformaciones generado por simulación y sub-ensambles de un tamaño de 20 o 50 conformaciones por ensamble, seleccionados por el Método de Optimización de Ensamblados (EOM). El análisis EOM de los datos experimentales de SAXS indicó que el complejo [RbAB:AdE1A.WT] ajusta en un perfecto acuerdo sólo a una sub-población del ensamble de conformaciones total, en el que la región “*linker*” de AdE1A.WT se encuentra moderadamente expandida (R_g sub-población ≈ 3.02 nm $>$ R_g medio del “*pool*” de estructuras modeladas ≈ 2.87 nm, Figura 4.3 C) a la vez que está unida por los dos motivos de interacción a RbAB (Figura 4.4 E).

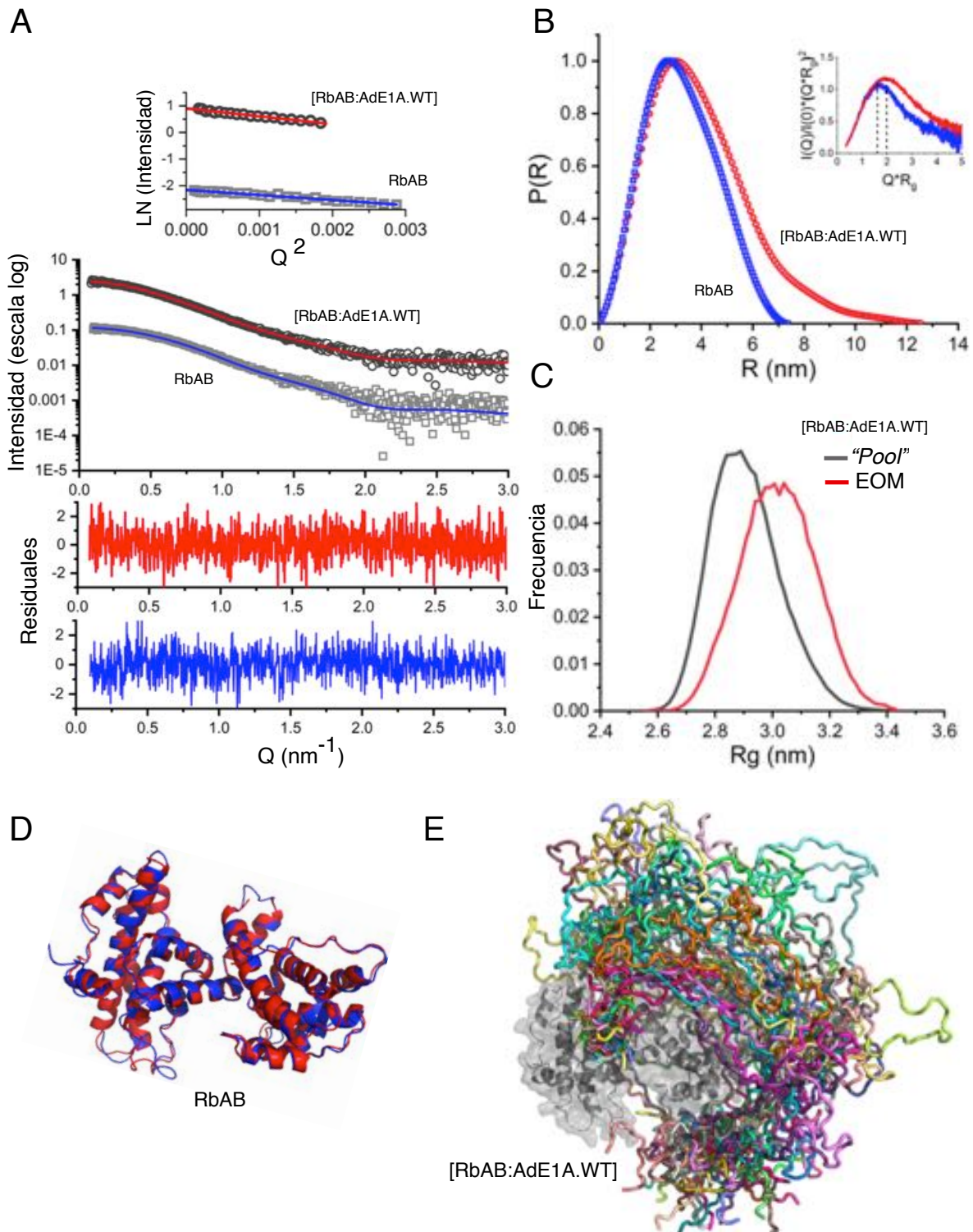


Figura 4.4. Análisis de SAXS para RbAB y el complejo [RbAB:AdE1A.WT]. A) Arriba, los gráficos de Guinier para RbAB y el complejo y sus ajustes lineales. En el medio, perfiles de SAXS para RbAB (cuadros grises) y para el complejo [RbAB:AdE1A.WT] y los perfiles teóricos de SAXS ajustados para la estructura de RbAB optimizada por MD (línea azul) y para la sub-población del ensamble del complejo optimizado por EOM (línea roja). Los residuales de los ajustes están presentados abajo. B) Gráfico de distribución de probabilidades para RbAB y [RbAB:AdE1A.WT] generados por la transformada de Fourier de los perfiles de SAXS. Los gráficos de Kratky están en el inserto. Las líneas punteadas marcan las posiciones de los picos de las funciones. C) La distribución de Rg para el “pool” total de ensambles modelados para el complejo [RbAB:AdE1A.WT] (línea negra) y para la sub-población seleccionada por EOM (línea roja). D) Estructura cristalina de RbAB (PDBID: 3POM, en rojo) y la estructura a la que se adicionaron los bucles perdidos de la densidad electrónica usando Allos-Mod (en azul). E) Las estructuras del complejo [RbAB:AdE1A.WT] seleccionados por EOM, con RbAB (en gris) y varias conformaciones distintas de AdE1A (en colores).

En resumen, como las estructuras modeladas del complejo [RbAB:AdE1A.WT] no consideraban interacciones electrostáticas ni hidrofóbicas “extra”-motivos y se observó un buen ajuste con los datos experimentales de SAXS, esto corroboró las evidencias previas que sugieren que la región “linker” de AdE1A no presenta interacciones permanentes con la superficie de RbAB (o si las hubiese serían muy débiles y no modifican la forma general del complejo) en tanto que permanece desordenada y flexible una vez formado dicho complejo (“linker” entrópico).

IV.3. Modelado “WLC” de la interacción entre AdE1A.WT y RbAB

El incremento en la afinidad en unas 4000 veces observado cuando AdE1A.WT -que posee ambos motivos de interacción en la misma molécula unidos por la región “linker”- se une a RbAB, en comparación con la afinidad de motivos individuales, implicaría que los motivos actuarían en forma conjunta y cooperativa en la unión con RbAB. Las probables explicaciones de que este efecto cooperativo estuviese mediado por un mecanismo alostérico clásico o interacciones estabilizantes mediadas por el “linker” fueron descartadas y esto es validado por el presente análisis.

En un mecanismo alternativo cooperativo (Modelo C, Figura 4.1), las propiedades conformacionales de AdE1A serían críticas en la determinación de su afinidad global por RbAB. Así, el “linker” en AdE1A que une ambos motivos de interacción a los extremos de la proteína viral, siendo desordenado y altamente flexible, sería funcionalmente importante porque modularía la distancia entre los motivos de interacción en AdE1A y la disponibilidad de los mismos en el espacio cercano a sus sitios complementarios en la proteína RbAB.

Dado que plantear un modelo de la interacción entre AdE1A y RbAB que satisfaga todas sus características y propiedades reales demostradas experimentalmente es complicado, se plantea la hipótesis de que es posible tomar una cualidad central del sistema, como el carácter entrópico del “linker”, para poder llegar a una descripción realista del mecanismo de interacción en estudio. Dicho esto, el comportamiento de este “linker” -como de otras regiones espaciadoras flexibles- puede ser descrito y predicho utilizando un modelo sencillo de polímero flexible como lo es el *modelo WLC* (“*Wormlike-chain*” en inglés) [111, 239]. Si bien este modelo omite detalles atómicos del sistema, composición de residuos de la cadena polipeptídica, incluso efectos de volúmenes excluidos y efectos electrostáticos, provee de una expresión analítica de concentración efectiva (C_{eff}) que es difícil de obtener a partir de

modelados mas explícitos (como el de las simulaciones) y que puede ser sumamente útil para entender cómo funciona la interacción entre AdE1A y RbAB.

El modelo WLC, que fue inicialmente propuesto por Zhou y colaboradores, describe a polímeros semiflexibles como una IDP, en forma de un cilindro continuo con un radio de curvatura fijo que cambia la direccionalidad de sus movimientos en solución de manera azarosa. En este modelo, las únicas dos variables de las que depende su descripción son la longitud de persistencia (l_p) y la longitud de contorno (l_c), los cuáles valen en este caso $l_p = 3 \text{ \AA}$ y $l_c = 3.8 \text{ \AA}$ multiplicado por el número de residuos de la región “*linker*” (Ver Materiales y Métodos).

Se utilizó el modelo WLC para calcular la distribución de probabilidad de distancias “*fin-a-fin*” ($\rho(r)$) y la concentración efectiva (C_{eff}) para AdE1A y su interacción con RbAB. Para ambos cálculos, sus respectivas ecuaciones están presentadas en Materiales y Métodos. $\rho(r)$ es la probabilidad de que los dos sitios extremos en la misma molécula de AdE1A.WT (el motivo E2F y el motivo LxCxE), estén separados por una distancia radial r , expresada así:

$$\rho(r) = 4\pi r^2 \left(\frac{3}{4\pi l_p l_c} \right)^{\frac{3}{2}} \exp\left(\frac{-3r^2}{4l_p l_c} \right) \zeta(r, l_p, l_c)$$

La fórmula desarrollada se presenta en Materiales y Métodos. Con una separación de 71 residuos entre el final del motivo E2F y el inicio del LxCxE, $\rho(r)$ mostró un máximo a unos 42 \AA de separación (Figura 4.5 A).

También fue calculada la concentración efectiva (C_{eff}) en función de la separación (distancia en residuos) entre los dos motivos en la misma molécula de AdE1A. C_{eff} está relacionado con $\rho(r)$ cuando la región “*linker*” se fija a una distancia radial dada (r_0). Para este cálculo se utilizó la distancia euclidiana entre el fin del motivo E2F y el inicio del motivo LxCxE, valiéndose de la superposición de las estructuras cristalográficas disponibles en el que RbAB se encuentra unido a E1A-CR1 (PDBID: 2R7G) y al motivo homólogo HPV16-E7 LxCxE (PDBID: 1GUX) (Figura 4.5 B), la cual asumió un valor de 49.02 \AA y se tomó como valor fijo r . Dentro de la aproximación del modelo WLC, la C_{eff} presentó el máximo de la función entre 70 y 90 residuos, cayendo abruptamente por debajo de 60 residuos y más gradualmente por arriba de 100 residuos (Figura 4.5 C). Además, C_{eff} asumió un valor de 0.95 mM cuando la longitud del “*linker*” se tomó como 71 residuos. En principio, estos resultados sugieren que el “*linker*” flexible de 71 residuos en AdE1A.WT presenta una longitud aproximada a la óptima entre los motivos E2F y LxCxE para mantener alta y constante su C_{eff} al unirse a RbAB. La forma del

gráfico puede entenderse así: longitudes de “linker” mas cortas serían insuficientes para permitir el contacto simultáneo de los dos motivos con sus surcos complementarios en RbAB, en tanto que para longitudes mas largas del “linker” disminuye la probabilidad de producirse el segundo contacto al tener permitido explorar/muestrear más áreas disponibles en el espacio.

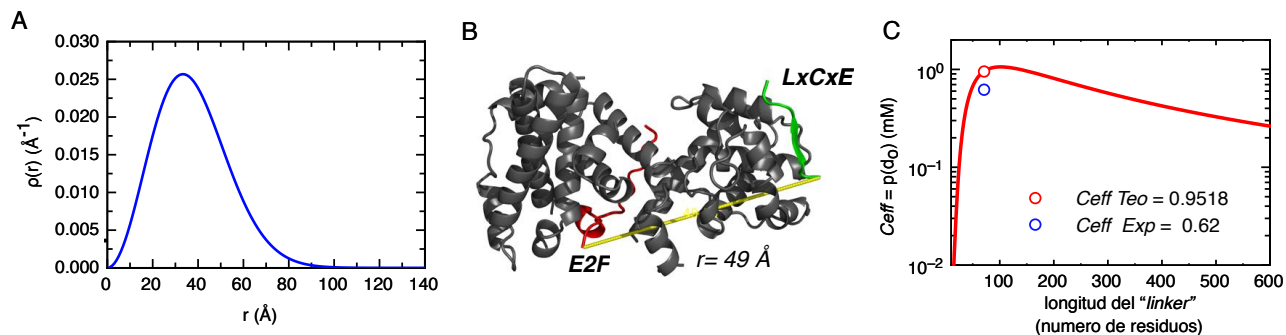


Figura 4.5. Modelado WLC para el “linker” de AdE1A.WT interactuante con RbAB. A) Distribución de probabilidad con parámetros $lp= 3 \text{\AA}$ y $lc= 3.8 \times 71 \text{ residuos} = 269.8 \text{\AA}$. B) Estructura cristalina de RbAB, unido en complejo a E1A-CR1 (E2F) de PDBID: 2R7G y unido en complejo a HPV16-E7-CR2 (LxCxE) PDBID: 1GUX, y la distancia de 49\AA tomada entre fin e inicio de un sitio de interacción y otro. C) Concentración efectiva del “linker” de AdE1A, asumiendo $lp= 3 \text{\AA}$, $lc= 269.8 \text{\AA}$ y $r= 49 \text{\AA}$, para longitudes de regiones espaciadoras variables. La línea roja describe el comportamiento de la función con respecto a la longitud del linker. El valor de la C_{eff} teórico calculado para el sistema en estudio está señalado con un círculo rojo, en tanto que el valor C_{eff} obtenido experimentalmente se señala en azul.

Para interacciones bivalentes, como la de AdE1A con RbAB, la afinidad de interacción global es proporcional a la C_{eff} . Esto puede ser visualizado si se considera la reacción de interacción entre AdE1A.WT con RbAB en un proceso de dos pasos (Figura 4.6):

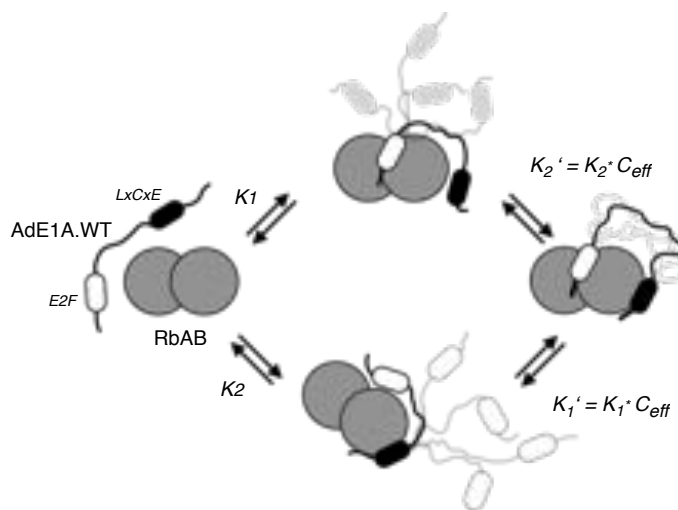


Figura 4.6. Modelo de interacción bivalente de AdE1A.WT con RbAB. La interacción de cualquiera de los dos motivos en AdE1A puede darse en primer lugar, en tanto que la segunda interacción es parte de un proceso unimolecular. El resultado global de la interacción propuesta es indistinto al orden de las interacciones. K , representa la constante de asociación (K_A) al equilibrio.

Para este modelo de interacción bivalente, una vez unido el primer motivo de AdE1A a RbAB, la segunda interacción avanza como un proceso unimolecular, en la que la constante de equilibrio puede escribirse como el producto de la K del motivo por C_{eff} . Así, resulta que K_{global} puede ser expresada de forma independiente al orden en que interaccionan los motivos (ya habiendo descartado procesos alostéricos), como:

$$K_{global} = K_1 \times K'_2 = K_2 \times K'_1 = K_1 \times K_2 \times C_{eff}$$

La K_{global} fue experimentalmente obtenida a partir de la constante de disociación de la interacción entre AdE1A.WT con RbAB ($K_{global} = 1/K_D$; $K_D = 2 \times 10^{-11}$ M). En tanto que K_1 fue obtenida a partir de la K_D de la interacción entre AdE1A-CR1 (motivo mínimo E2F) con RbAB ($K_1 = 1/K_D$; $K_D = 1.04 \times 10^{-7}$ M) y K_2 a partir de la K_D de la interacción de AdE1A-CR2 con RbAB ($K_1 = 1/K_D$; $K_D = 1.19 \times 10^{-7}$ M). Con estos datos, la C_{eff} experimental hallada fue de 0.62 mM (Figura 4.5 C). Como se puede observar, el valor experimentalmente obtenido de C_{eff} , acuerda muy bien con el valor derivado de aplicar el modelo WLC a un “linker” de 71 residuos ($C_{eff} = 0.95$ mM). Esto es sorprendente, sobre todo tendiendo en cuenta que el modelo WLC representa una reducción importante del sistema en estudio, lo que a su vez evidencia que el carácter entrópico o desordenado y altamente flexible del “linker” de AdE1A es determinante en la formación del complejo con RbAB.

En conclusión, el modelo WLC (Modelo C, Figura 4.6) permite explicar la potenciación de la afinidad de AdE1A.WT por RbAB por un incremento en la concentración del segundo motivo sobre su sitio complementario en RbAB, producido por un “linker” de carácter entrópico y longitud adecuada. Aunque sea una aproximación muy simplificada o reduccionista, es consistente con las evidencias biofísicas, termodinámicas y estructurales (RMN) y permite comprender el mecanismo de la interacción entre AdE1A y RbAB.

CONCLUSIONES IV

En este último capítulo de la tesis se trabajaron con modelos de la interacción de RbAB con AdE1A que permitan explicar a modo integral las contribuciones energéticas y estructurales de los motivos de contacto y la región “*linker*” en la proteína viral y por tanto, el mecanismo cooperativo que media tan alta afinidad de interacción entre estas proteínas.

En primer lugar, se presentaron modelados teóricos obtenidos por simulación (en un trabajo colaborativo) para los complejos [RbAB:AdE1A.WT], [RbAB:AdE1A- Δ E] y [RbAB:AdE1A- Δ L]. Estos modelos se construyeron a partir de la estructura cristalina de RbAB y considerando sólo las secuencias reales de las variantes de AdE1A. Los R_H teóricos calculados a partir de ellos mostraron excelente acuerdo con los R_H experimentales, validando estos modelos. En el complejo [RbAB:AdE1A.WT] modelado, el “*linker*” que se muestra flexible y sin interacciones con RbAB, es concordante con los datos termodinámicos y de RMN hasta ahora alcanzados. Para los complejos mutantes, la comparación entre los R_H calculados y experimentales, evidencian que AdE1A podría estar preferencialmente adoptando ciertas conformaciones más compactas, en las que el “*linker*” mediaría interacciones transientes más estables, en concordancia con los datos que evidencia el RMN.

Por otro lado se presentaron resultados obtenidos de experimentos SAXS sobre el complejo [RbAB:AdE1A.WT], también realizados en colaboración. Para el caso de RbAB, el perfil de señales de dispersión obtenido fue equivalente a la estructura cristalina monomérica de esta proteína, indicando que la misma también se encuentra en dicho estado en solución. Tal concordancia además constituye un indicativo de la calidad de estos experimentos. Ahora bien, para determinar las propiedades globales en términos de tamaño y forma del complejo entre RbAB y AdE1A.WT, se utilizaron los modelos teóricos –el ensamble de conformaciones generados por simulación– con los que ajustaron los datos experimentales de SAXS. De posibles escenarios que ajustan a los datos de SAXS, definitivamente el más probable y consistente con los datos cuantitativos/termodinámicos, es el de un modelo que considera un sub-ensamble de conformaciones en que el “*linker*” en AdE1A.WT se encuentra siempre extendido y sin mediar interacciones con RbAB, una vez formado el complejo a partir de los dos motivos E2F y LxCxE.

Por último, se analizó el comportamiento de la región enlazante o “*linker*” entre los motivos de interacción E2F y LxCxE virales, según el modelo WLC o de cadena entrópica. El modelado fue satisfactorio prediciendo la potenciación de la afinidad de AdE1A.WT por RbAB que se había observado con los datos termodinámicos, a razón de un incremento en la concentración efectiva de los motivos en el entorno cercano a su sitio complementario en

RbAB. En la interacción de carácter bivalente entre AdE1A.WT con RbAB, la región “*linker*” es entonces clave porque se comporta aproximadamente como un polímero ideal desordenado y flexible que teniendo una longitud óptima, modula las distancias entre los motivos y permite su interacción en los sitios complementarios de RbAB. Esta descripción del mecanismo de interacción de AdE1A con RbAB, es consistente con los datos estructurales aportados por RMN, la termodinámica, así también con los datos de SAXS.

Estas características describirían al complejo RbAB:AdE1A como un sistema típico “*fuzzy*”, en el que el estado desordenado de AdE1A en el complejo juega un rol funcional clave. Además, el efecto de la región “*linker*” en la interacción con RbAB no es un excepcional, si no mas bien una característica que parece ser inherente de sistemas que presentan regiones enlazantes flexibles, como lo demuestra la aplicación satisfactoria del modelo WLC en varios sistemas no relacionados entre sí [57, 112, 240].

DISCUSION Y CONCLUSIONES GENERALES

Los virus constituyen agentes patogénicos de importancia mundial y el costo social y económico consecuente a las enfermedades que provocan, ponen de manifiesto la importancia de las investigaciones tendientes a estudiar los mecanismos por los cuáles estos patógenos invaden y toman control de las células. Se estima que al menos un 11% de los casos de cáncer humano tiene etiología viral, reforzando la necesidad de conocer los mecanismos de acción de virus oncogénicos, para así generar conocimiento que a la larga desemboque en estrategias útiles para su tratamiento.

La base de la comunicación virus-hospedador está dada por interacciones proteína-proteína entre los mismos, a partir de las cuáles los virus logran subvertir toda la maquinaria celular a su favor. Dado el pequeño tamaño de los genomas virales, el número de proteínas codificadas con el que pueden lograr este objetivo, está restringido. No obstante, la alta representación del desorden intrínseco en sus proteínas les proveería de ventajas funcionales, que serían útiles para lograr una efectiva interferencia de las redes de interacción celulares.

En esta tesis, se tomó a la interacción entre la proteína viral AdE1A de Adenovirus y el supresor de tumores Retinoblastoma, pRb, como sistema modelo para analizar mecanismos de subversión viral de la célula huésped y tratar de responder si el estudio estructural y biofísico de este complejo puede ayudar a comprender el aporte del desorden intrínseco a las funciones de las proteínas virales.

Información codificada a nivel de secuencia y conformación en AdE1A juegan un rol clave en sus propiedades y funciones

Predicciones y estudios experimentales previos a este trabajo, ya habían identificado la naturaleza desordenada de algunas regiones AdE1A, sin llevar a cabo estudios exhaustivos. No obstante, en la presente tesis se presentó un estudio biofísico-bioquímico que permitió comprender en mayor detalle el carácter IDP de esta proteína viral (Capítulo I). AdE1A mostró ser particularmente sensible a proteólisis *in vitro*, lo que está asociado frecuentemente a regiones desordenadas en proteínas, ya que para que esto se produzca, se requiere una interacción y adaptación al sitio activo de una proteasa, lo que es potenciado por la flexibilidad de la cadena polipeptídica [25]. Esta sensibilidad a proteólisis *in vitro* (que se localizó sobre argininas con un grado de conservación moderado y sin función caracterizada hasta ahora –ver más adelante-), podría estar relacionada con mecanismos de control rápido y efectivo que regulen el recambio (“*turnover*”) [132, 241] de AdE1A *in vivo*, así como también con una hipotética posibilidad de generar módulos de interacción independientes con blancos celulares a los que se une y que requieran distintas localizaciones sub-celulares y que a la vez no

demanden que AdE1A esté intacta en toda su extensión. AdE1A es una IDP “*random coil-like*”, con migración electroforética anómala y un comportamiento hidrodinámico altamente extendido en solución, que estarían dados por el sesgo composicional de su secuencia. Las IDPs se encuentran depletadas en residuos hidrofóbicos y enriquecidas en residuos polares y cargados [26]. La alta fracción de prolinas en AdE1A podría imprimirle efectos de isomerización *cis-trans* desestabilizantes de estructuras [117] y su alto contenido de residuos acídicos, un probable efecto repulsivo de fuerzas electrostáticas negativas que mantengan su conformación expandida y solvatada [29]. La agregación observada bajo pH, podría ser resultado final de una neutralización progresiva de esas cargas negativas que conlleve a estructuras cada vez más compactas, como se ha demostrado para otras IDPs, como la α -sinucleína [242]. AdE1A es una proteína que presenta gran flexibilidad conformacional y es eminentemente desordenada, careciendo de cantidades significativas de estructuras secundarias. Sin embargo, muestra gran plasticidad conformacional al poseer la capacidad de estabilizar en forma transitoria estructura α -hélice y PPII, en distintos entornos físico-químicos. Estos elementos transientes se estabilizarían en la interacción con diferentes blancos celulares [47], como lo demuestran las α -hélices formadas por el motivo E2F en complejo con RbAB [84] y el motivo TRAM-CBP en complejo con los dominios NCBP y TAZ2 de los co-activadores transcripcionales CBP/p300 [76, 82].

Los motivos E2F y LxCxE de AdE1A comparten propiedades con SLiMs de unión a pRb ya conocidos.

AdE1A es una IDP que media sus interacciones proteína-proteína con blancos celulares a través de múltiples SLiMs en su secuencia [5], los cuáles hacen mímica de motivos endógenos. La interacción entre AdE1A y pRb por caso, puede ser racionalizada en las contribuciones establecidas de manera conjunta y sinérgica por dos de estos SLiMs: el motivo E2F en CR1 y el motivo LxCxE en CR2 (y la región “*linker*” desordenada y flexible que los separa). Los experimentos termodinámicos llevados a cabo en esta tesis mostraron que en AdE1A, estos dos motivos se unen con afinidad moderada a alta a pRb, en concordancia con sus regímenes de intercambio lento, caracterizados por RMN. Estos sitios son capaces de unirse a pRb de manera individual e independiente uno respecto del otro, y han demostrado no presentar efectos alostéricos en “*trans*” recíprocos que modifique su afinidad por pRb. La afinidad del motivo viral E2F es menor al del motivo celular E2F por pRb, mientras que la afinidad del motivo viral LxCxE es mayor a la que presentan blancos celulares endógenos con el mismo motivo, como HDAC y/o RIZ [141, 150] a pRb. Dado este escenario, AdE1A requiere

del funcionamiento cooperativo entre ambos motivos, facilitado por el “*linker*” que los une, para desplazar al factor de transcripción E2F de su complejo endógeno con pRb. Esto revela una estrategia distinta para inactivar este regulador del ciclo celular, a la usada por ejemplo, por HPV-E7, la cual posee un motivo LxCxE que es 50 veces más afín que el mismo motivo en AdE1A y requeriría de una segunda molécula de E7 (dimerización) para lograr tal inactivación [100].

El incremento en 5 veces la afinidad del motivo mínimo viral LxCxE por pRb, a consecuencia de la fosforilación de su sitio CKII (S132), compara muy bien con el aumento en 4 veces en la afinidad por pRb de la región homóloga en HPV-E7, producido por la fosforilación de sus sitios CKII (S31 y S32) [100], y denota que en AdE1A: este motivo, el parche ácido y el sitio CKII (S132) se comportan como un módulo funcional, tal y como se ha visto para otras proteínas que presentan el motivo LxCxE [100, 141, 150]. No obstante, el RMN evidenció que estos sitios flanqueantes se unen sólo cuando el motivo mínimo está presente, demostrando que este último es el principal responsable de la energía de interacción con pRb y los residuos flanqueantes cumplen mas bien un rol modulador, mejorando cumulativamente su afinidad por la proteína complementaria. Para el caso del motivo E2F en AdE1A se encontró por RMN una primera evidencia de que también existen residuos flanqueantes hacia el C-terminal que podrían estar modulando la afinidad de dicho motivo. Estas evidencias en AdE1A corroboran que la secuencia aminoacídica que contextualiza un dado SLiM, tendría efectos significativos en su afinidad y funcionalidad [193].

Los mecanismos de interacción de AdE1A como IDP con otras proteínas

Tanto los estudios de CD como los de RMN, demostraron que el fragmento de AdE1A estudiado (36-146) no posee elementos de estructura secundaria ni tampoco los puebla en forma transiente. Esto implica que las interacciones que establecen los motivos E2F y LxCxE con RbAB (o en el “*linker*”, el motivo TRAM-CBP con CBP/p300) los cuales están estructurados sólo en sus formas unidas, experimentan plegamientos-acoplado-a-interacción por un mecanismo de ajuste inducido. En las interacciones entre dominios globulares, la especificidad está frecuentemente acoplada a la afinidad, pero en las interacciones que involucran transiciones desorden-a-orden, la entropía puede convertirse en la principal contribución de la energía libre de unión, de modo que la disminución de la entropía conformacional desacopla la especificidad de la afinidad [243]. Así, estas interacciones individuales pueden ser altamente específicas pero reversibles [21]. Por un lado, esta característica, sumado a los múltiples y pequeños SLiMs en la secuencia de AdE1A, la harían propensa a actuar cambiando

rápidamente de blanco de interacción y formar complejos multivalentes cruciales para re-direccionar las respuestas celulares en su beneficio. Por otro lado, la forma de actuar conjunta y cooperativa de múltiples SLiMs (y por caso, segmentos como un “*linker*”), le pueden conferir afinidad global muy alta por blancos celulares que son cruciales inactivar para tomar control celular, como a pRb en una etapa temprana de la infección.

La importancia funcional de la composición y el desorden en AdE1A se evidencia en su alto grado de conservación evolutiva

La presencia de elementos funcionales en la secuencia de AdE1A por medio de los que interactúa con pRb: motivos E2F y LxCxE y la región “*linker*”, no son el resultado de una evolución neutral -en ausencia de selección natural-, como ha sido propuesto para algunas IDPs [244], ni tampoco presentaría un grado menor de conservación de secuencias y co-evolución de residuos, como usualmente se cree al comparar tasas de conservación evolutiva de IDPs vs proteínas globulares [244, 245]. En un primer reporte, Avvakumov y colaboradores [246] y posteriormente Glavina y colaboradores, en un trabajo más completo basado en el análisis de 116 secuencias de AdE1A de *Mastadenovirus* [75], lograron demostrar que muchas posiciones en la secuencia presentan de medio a alto grado de conservación (Figura D.1 A). Los residuos que componen el motivo E2F y el módulo funcional: motivo LxCxE + parche ácido + sitio CKII (S132), por ejemplo, están presentes en casi todas estas secuencias estudiadas. Otros motivos como TRAM-CBP, que media contactos con co-activadores transcripcionales CBP/p300 [76, 82] o MYND con el represor transcripcional BS69 [247], ambos presentes en la región “*linker*” estudiada en esta tesis, están representados en más del 10 y 50% de las secuencias, respectivamente. La moderada a alta conservación incluso de posiciones sin actividad biológica conocida (Figura D.1, residuos en gris), sugeriría la presencia de nuevos SLiMs funcionales aún no caracterizados [75].

Por otro lado, la predicción del desorden tras el alineamiento de este “*pool*” de secuencias, mostró que esta característica también está conservada evolutivamente (Figura D.1 B) [75]. Las prolinas, considerados residuos disruptores de estructuras muestran también un sorprendente grado de conservación sobre la región “*linker*”, revelando la importancia de mantener el desorden en esta región de AdE1A. Estos resultados reflejan la presión evolutiva para mantener el desorden y múltiples motivos de interacción, a pesar de las altas tasas de mutación genómica viral [9]. Esto sería necesario para permitirle a esta proteína interactuar con una plétora blancos celulares, optimizando funcionalmente la longitud de su secuencia, constreñida como consecuencia de su tamaño genómico.

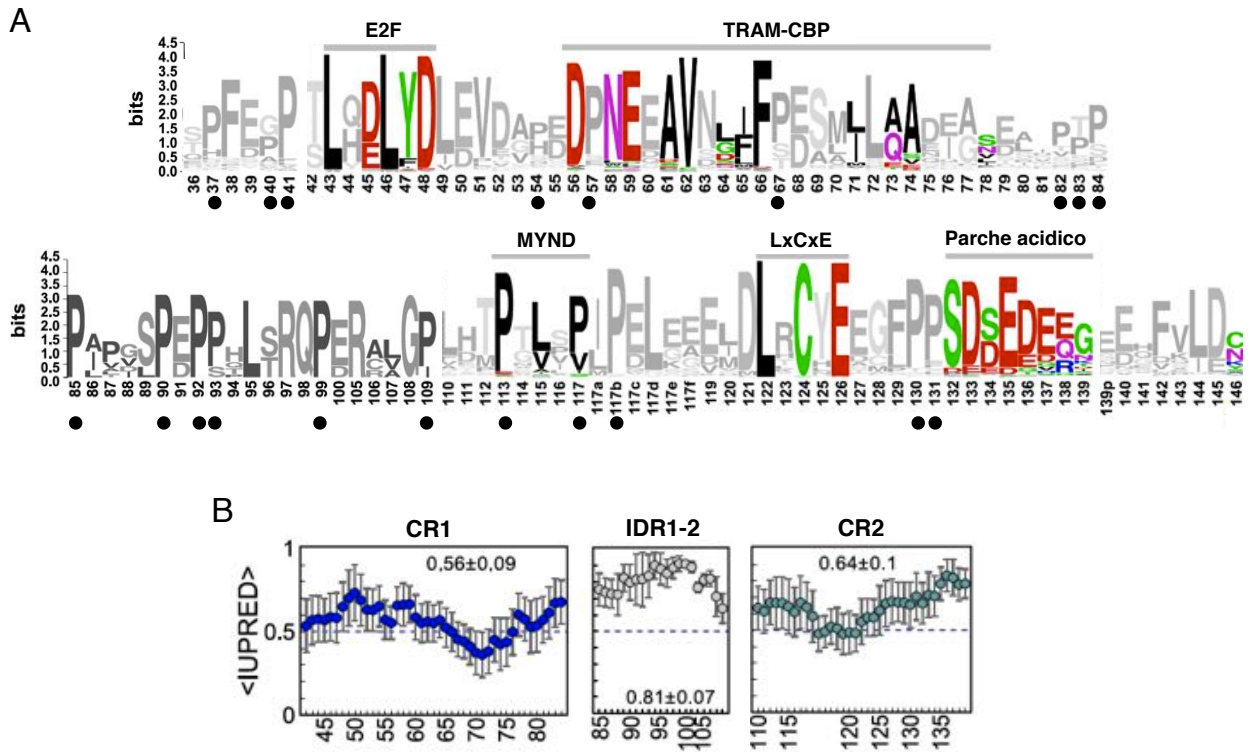


Figura D.1: Conservación evolutiva de secuencia y desorden en AdE1A A) La representación de logos mostrando el alto grado de conservación (valores de información iguales o superiores a los 2 bits) de los residuos en 116 secuencias alienadas de AdE1A de *Mastadenovirus*. La numeración es coincidente con la secuencia usada en esta tesis (Ad5E1A). Las posiciones repetidas representan inserciones para alinear aquellas secuencias de AdE1A más largas. Los motivos están señalados y nombrados, con sus posiciones fijas coloreadas. Las prolinas se muestran señaladas con puntos negros. B) Predicción de desorden IUpred [85], para cada residuo en los dominios y región estudiados. El puntaje superior a 0.5 indica desorden y por debajo, estructura ordenada. Se muestran valores promedio y desvíos estándar (Adaptado de [75]).

Las propiedades conformacionales del “linker” en AdE1A también están conservadas

En esta tesis se demostró que la región “linker” en AdE1A no establece contactos directos estables con RbAB, pero tampoco es un elemento pasivo. El acuerdo general entre los resultados biofísicos, termodinámicos y estructurales con el modelo WLC, sostiene la hipótesis de que el “linker” promueve la optimización de la interacción AdE1A-pRb por un efecto de concentración efectiva. Esto se da gracias a que esta región se comporta globalmente como un polímero ideal, que es desordenado, flexible y de longitud óptima para permitir el contacto de ambos motivos a pRb. Estas propiedades que son clave en el rol de este “linker”, resultaron también estar conservadas evolutivamente. De un análisis de 110 secuencias de “linkers” AdE1A de *Mastadenovirus*, (entre los motivos E2F y LxCxE), se encontró que más del 85% de sus posiciones se predicen desordenadas. Este comportamiento entrópico ideal, podría ser producto de una compensación en el sistema real entre interacciones hidrofóbicas débiles y repulsión electrostática, lo que podría contribuir a la función “hub” de AE1A (ver más adelante).

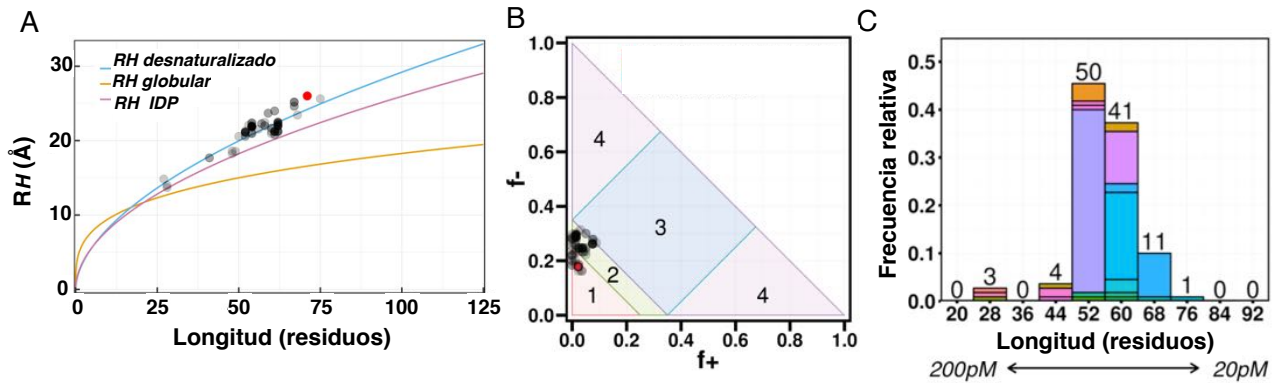


Figura D.2: Conservación evolutiva de las propiedades conformacionales del "linker" de AdE1A A) Radios hidrodinámicos predichos para "linkers" 110 secuencias de AdE1A. En rojo en "linker" de 71 residuos en estudio. Curvas de colores teóricas siguiendo las formulaciones básicas de Marsh & Forman-Kay, para la predicción de R_H en función del número de residuos, según su naturaleza IDP, globular o desnaturalizada. B) Gráfico de distribución de cargas y propensiones conformacionales, mostrando la alta fracción de cargas negativas conservadas en los "linkers" AdE1A. En rojo, la fracción de carga del "linker" en estudio. (para mayor referencia ver Introducción, Figura I.2) C) La distribución de los largos de los "linkers" muestra la conservación de un largo óptimo necesario para que se produzca la potenciación de la afinidad. Abajo, la K_D global que correspondería al sistema bivalente de interacción AdE1A-RbAB, según el largo del "linker". Los datos de la presente figura fueron generados y amablemente provistos por la Lic. Juliana Glavina.

Los valores predichos para los radios hidrodinámicos de estas secuencias, muestran que las mismas corresponden a proteínas altamente extendidas (Figura D.2 A) debido a su desorden y flexibilidad, y preservan el sesgo composicional con alto contenido de residuos acídicos (Figura D.2 B). De modo sorprendente, el largo del "linker" (en términos de residuos) se encuentra muy bien conservado entre estas secuencias analizadas, con una media de ≈ 59 residuos (Figura D.2 C), cercanas al largo óptimo predicho por el modelo WLC para permitir el contacto más afín de ambos motivos con RbAB (70-95 residuo). Este comportamiento descrito para el "linker", ha sido validado también para otros sistemas bivalentes de interacción entre IDPs y sus proteínas blanco. En estudios de Barrick y colaboradores, la interacción entre NCID (dominio intracelular Notch) y factor de transcripción nuclear CSL -involucrados en diferenciación celular en *matozoos*- fue modelado utilizando estadística de polímeros. Con ello, describieron que el "linker" de 103 residuos entre los dominios RAM y ANK de NCID, desordenado y rico en residuos cargados (35% D, E, K, R) poseía la longitud ideal para lograr que ambos se unieran a sus sitios complementarios en el complejo, prediciendo el incremento en la afinidad del sistema vía concentración efectiva [112]. En estudios posteriores, observaron que el acortamiento de esta región "linker", por delección [248] o por interacciones electrostáticas intracatenarias [110], disminuía la activación transcripcional asociada a la formación del complejo. Esto constituye evidencia *in vivo* de que la longitud y el patrón de distribución de cargas de estas regiones "linker" no es arbitraria, sino que está bajo selección evolutiva para lograr que la proteína presente niveles óptimos de funcionamiento.

AdE1A es una proteína clave en la subversión de la maquinaria celular

Para poder unirse a pRb, AdE1A utiliza dos motivos unidos por un "linker" que si bien permanece desordenado y flexible, también establecería interacciones muy transientes o débiles con la superficie de pRb, a través del motivo TRAM-CBP. En un escenario en el que las funciones proteicas virales están maximizadas, interacciones inespecíficas y/o ubicaciones arbitrarias de motivos parecen improbables, por lo que estas interacciones transientes podrían entenderse como parte de una estrategia evolutiva viral para "esconder" al motivo TRAM-CBP, hasta dar con su blanco de interacción específico. La formación del complejo ternario pRb-AdE1A-CBP/p300, que maximiza las funciones biológicas de la proteína viral y que fue caracterizado por Ferreon y colaboradores [76], sería una prueba de ello. Con la formación de estos complejos de alto orden por ejemplo, AdE1A podría eludir al sistema inmune de la célula huésped gracias a que sus epítopes virales se harían difíciles de reconocer por el sistema inmunológico, como se ha visto para otras proteínas virales [11].

En los últimos años, se ha propuesto que esta capacidad de unir simultáneamente a varios blancos, que es compartida por muchas IDPs, sería la base de un mecanismo alostérico no clásico mediado por el desorden intrínseco. El grupo de investigación de Deniz y colaboradores, ha probado que dada una primera interacción de un motivo de AdE1A con su blanco celular (pRb o CBP/p300), se dejaría a disponibilidad otros sitios de interacción en su secuencia, que dependiendo de cuáles sean, podrían conducir a efectos cooperativos que regulen finamente la formación de complejos binarios (pRb-AdE1A o CBP/p300-AdE1A) vs el ternario [180] y en consecuencia, regularían las redes de interacción en las que AdE1A se encuentra involucrada. Esto sería particularmente importante para esta proteína, que siendo una IDP que funciona como nodo o "hub", es capaz de direccionar respuestas celulares en función de sus necesidades. La alta afinidad provista por el comportamiento entrópico del linker permitiría priorizar complejos como AdE1A-pRb, sobre el cual se ensamblen complejos ternarios y cuaternarios. La formación de estos últimos podría depender de la compensación entre propiedades electrostáticas e hidrofóbicas, que permitiría codificar múltiples SLiMs hidrofóbicos en una región negativamente cargada que se repele con la superficie de RbAB.

El desorden intrínseco, alta flexibilidad y plasticidad conformacional y la presencia de múltiples SLiMs en su secuencia, hacen de AdE1A una proteína viral "hub" capaz de interactuar con más de 50 blancos celulares probados -y de manera indirecta con cerca de otros 2000 blancos [4]. Su capacidad para interactuar con tantas proteínas endógenas a pesar de su pequeño tamaño, junto su alto grado de conservación evolutiva [75] y el hecho de ser la primer proteína en ser transcripta tempranamente en el ciclo infectivo del Adenovirus, revelan la

importancia estratégica de esta proteína y sus características y propiedades, en el proceso de subversión de la célula provocado por estos virus.

Los aportes presentados en esta tesis, a partir del estudio de la interacción entre AdE1A de Adenovirus y pRb, constituyen la primera demostración del uso del mecanismo de “concentración efectiva” en el desplazamiento de blancos endógenos por una IDP viral. En este trabajo se resalta el rol del desorden intrínseco en las proteínas virales y las ventajas funcionales derivadas, y generan mayor comprensión acerca de los mecanismos que utilizan los virus para secuestrar y tomar control de la maquinaria celular, pudiendo servir de puntapié para el desarrollo de estrategias que sean efectivas para combatirlos.

Tesista

Bqco. Nicolás Sebastián González Foutel

Directora

Dra. Lucía Beatriz Chemes

MATERIALES Y METODOS

Biología Molecular

Clones de RbAB:

De la proteína Retinoblastoma (Uniprot ID P06400) se utilizó el dominio central (RbAB) (aminoácidos 372 a 787) con una delección de la región enlazante comprendida entre los aminoácidos 582 a 642 – que le otorga mayor estabilidad termodinámica– y cuyo vector de expresión fue previamente desarrollado en nuestro laboratorio [133]. Brevemente, RbAB fue sub-clonada entre los sitios de restricción *Bam*HI y *Hind*III de un vector pRSET-A (ThermoFischer Scientific) que fue “ingenierizado” para contener un sitio de corte específico de Trombina entre la marca de 6 histidinas (6xHis *tag*) y la secuencia de RbAB. Este sitio de corte de Trombina dejó como subproducto de la proteólisis los residuos GSISM hacia el extremo N-terminal de la secuencia natural de RbAB, los cuales no interfirieron con ninguna de las propiedades de dicha proteína.

Clones de AdE1A: fueron explicados en el Capítulo I.

Expresión y Purificación de Proteínas

Expresión y purificación de RbAB:

Para expresar RbAB se siguió protocolo similar al previamente desarrollado por el laboratorio y descrito en [133]. Brevemente, A. *Transformación y expresión:* se transformaron bacterias *Escherichia coli* C41-(DE3) con el plásmido pRSET- RbAB y fueron crecidas por una noche a 37 °C. Se seleccionó un único clon con el que se inoculó un Erlenmeyer con 150 ml de medio 2TY suplementado con ampicilina 1X y glucosa 1% (m/v), que se hizo crecer durante una noche a 37 °C con agitación constante (220 rpm). Con el pre-cultivo saturado se inocularon 3l de medio 2TY y las bacterias fueron crecidas hasta alcanzar una $DO_{600} = 0.6$, momento en que los cultivos fueron inducidos con 1 mM IPTG y se dejaron crecer a 28 °C ON ($DO_{600} = 4$). Las bacterias fueron recolectadas y el pellet bacteriano fue lisado por sonicación para obtener una fracción soluble en la que particiona RbAB. B. *Cromatografía IMAC:* la fracción soluble se dializó ON con tampón 25 mM Tris:Cl, 1mM PMSF pH 7.5 y se le adicionó 0.5 M NaCl y 10mM Imidazol, tras lo cual fue inyectada en una columna IMAC cargada con $NiCl_2$ a un flujo de 1ml/min. Luego se lavó la columna con 10% de tampón IMAC B (25 mM Tris:Cl pH 7.5, 1mM PMSF, 0.5 M NaCl + 0.5M Imidazol) y se aplicó un gradiente 10 a 100% por 30 min, tras lo cual se recuperó RbAB. Las fracciones que contenían RbAB fueron dializadas contra un tampón: 20mM fosfato de sodio pH 7.0, 2mM DTT y 0.1M NaCl. C. *Cromatografía de intercambio iónico:* RbAB tiene un pI teórico de 8.9, por lo que en este paso se utilizó una columna de intercambio catiónico fuerte SP Sefarosa. La muestra se inyectó en

la columna previamente equilibrada a 5% de tampón B (20mM Fosfato de Sodio pH 7.0, 2mM b-ME + 2M NaCl). Se aplicó un gradiente de 5-100% de tampón B en 20 minutos (1ml/min) tras lo que eluyó RbAB. D. *Corte con Trombina*; la muestra fue dializada en tampón de corte (20mM Fosfato de Sodio pH 7.0, 0.2M NaCl, 2mM b-ME). Se cuantificó RbAB por Abs 290nm y se concentró a 2mg/ml. Luego su extremo N-terminal conteniendo el 6xHis *tag* (un total de 38 aminoácidos) fue clivado por tratamiento con Trombina (Sigma) a razón de 1 unidad por mg de proteína a 30°C ON en el mismo tampón+ 5% glicerol. El corte fue para por adición de 2mM PMSF y la muestra fue centrifugada y filtrada. E. *Cromatografía de exclusión molecular*, en el último paso la muestra fue concentrada y pasada utilizando una columna de filtración. Se realizaron inyecciones de 2ml de muestra y se corrió una columna preparativa Superdex 75 a 2 ml/minute con tampón: 20mM Fosfato de Sodio pH 7.0, 0.2M NaCl, 2mM b-ME . La muestra de proteína eluyó como un pico homogéneo del PM esperado. La proteína fue obtenida con una pureza superior al 95% evaluada por gel SDS-PAGE y fue cuantificada por Bradford y Absorbancia a 280 nm.

Expresiones (Medio Rico y Mínimo M9) y purificación de AdE1A:

Fueron explicados en el Capítulo I. A continuación se presentan la composición de medios y tampones utilizados en estos procedimientos:

<i>Medio rico 2TY</i>	p/3 l 2TY
Triptona	48 g
Extracto Levadura	30 g
NaCl	15g
H ₂ O	csp 3l

Medio mínimo M9

Sales M9	p/1 l sales M9
Na ₂ HPO ₄	60g
KH ₂ PO ₄	30g
NaCl	5g
pH 7.0 y filtrado en esterilidad	

	p/2 l medio mínimo M9
Sales M9	200 ml
MgSO ₄ 1M	30 g
CaCl ₂ 50mM	4ml
FeCl ₃ 0.01M	2ml
Vitamina B1 5mg/ml	400µl
Glucosa 40%	10ml
(o ¹³ C-Glucosa sólida 2g)	
¹⁵ NH ₄ Cl	1g

<i>Tampón de lisis</i>	p/50 ml
50 mM Tris:Cl pH 7.5	2.5 ml Tris:Cl pH 7.5 1M
1 mM EDTA	250 ul EDTA 200mM
2 mM β -Me	7,1 ul β -Me 14M
2 mM PMSF	0.5 ml PMSF 0.2M
<i>Tampón A (Amilosa)</i>	p/1000 ml
50 mM Tris:Cl pH 7.5	50 ml Tris:Cl pH 7.5 1M
200 mM NaCl	11,68g NaCl
2 mM β -Me	143 ul β -ME 14M
2 mM PMSF	10ml PMSF 0.2M
<i>Tampón C (Corte conTrombina)</i>	p/1000 ml
50 mM Tris:Cl pH 7.5	50 ml Tris:Cl pH 7.5 1M
150 mM NaCl	8,76 g NaCl
2 mM β -Me	143 ul β -Me 14M
Sin PMSF	
<i>Tampón Q (Q-Hyper-D)</i>	p/1000 ml
QA: 30 mM Tris:Cl pH 7.5	30 ml Tris:Cl pH 7.5 1M
50 mM NaCl	12,5 ml NaCl 4M ó 2,92 g NaCl
2 mM β -Me	143 ul β -Me 14M
6M Urea	360,36 g Urea sólida
2 mM PMSF	10ml PMSF 0.2M
QB: Q + 1.2M NaCl	300 ml QA + 19,725 g NaCl
<i>Tampón S-75</i>	p/ 600 ml
10 mM fosfato pH 7.0	30 ml 0.2M fosfato pH 7.0
2 mM β -Me	85.8 ul
200 mM NaCl	7.08 g

Síntesis de Péptidos

Los péptidos utilizados en este trabajo fueron sintetizados por método de síntesis con 9-fluorenilmetoxicarbonilo (Fmoc) con un grado de pureza >95% (GenScript USA Inc. Piscataway, NJ) y liofilizados. Los mismos se resuspendieron en 20 mM fosfato de sodio pH 7.0 y los que experimentaron mayor dificultad para su disolución se resuspendieron en agua MilliQ agregando NaOH. Todos los péptidos se disolvieron a una concentración aproximada de 1 mM (\approx 2mg/ml) y fueron conservados a -80 °C hasta su uso. La siguiente constituye una lista y descripción de los péptidos utilizados:

Nombres	PM_{TEO} (Da)	Descripción	Secuencias
AdE1A-CR1	1912.09	Motivo E2F localizado en la región CR1 de AdE1A	36-SHFEPPTLHELVDLDV-51
AdE1A-CR2	1723.96	Motivo LxCxE localizado en la región CR2 de AdE1A	116-VPEVIDLTCHEAGFPP-131
AdE1A-CR2 _{AC}	2600.70	Parche ácido hacia el C-terminal contiguo al motivo LxCxE de AdE1A	116-VPEVIDLTCHEAGFPPSDEDEEG-139
AdE1A-CR2 _{ACP}	2679.68	Parche ácido hacia el C-terminal contiguo al motivo LxCxE de AdE1A, fosforilada en la Serina 132	116-VPEVIDLTCHEAGFPP (pS) DDEDEEG-139
E2F2	2641.76	Motivo E2F en E2F2 Humano	404-SPSLDQDDYLWGLEAGEGIGISDLFD-427

Marcado fluorescente de Péptidos y Proteínas

Las proteínas y péptidos fueron marcados en su extremo N-terminal con Fluoresceína 5-isotiociano (FITC, Sigma). Para esto se preparó un “stock” de FITC 15mg/ml en acetona 100%, debido a su insolubilidad en agua. Por otro lado, proteínas y péptidos fueron pesados alrededor de 3mg de muestra y marcados por adición de 1.5 mg/ml de FITC en tampón 100mM carbonato de sodio pH 8.0 a volumen final 1ml a temperatura ambiente por 2 a 16 horas, cubriendo de la luz a los tubos eppendorff donde se llevaron a cabo las reacciones y en agitación constante. Al pH utilizado (pH 8.0) FITC reacciona preferencialmente con grupo α-amino libres (amino del extremo amino-terminal) y con menor probabilidad los grupos ε-amino (como los de las cadenas laterales de los residuos lisina y arginina). Después de las horas de incubación, la reacción de marcación fue detenida por agregado de tampón Tris:Cl 50 mM a pH 7.5 por 20 minutos. Posteriormente se procedió a separar las muestras marcadas del FITC en exceso usando columnas de desalado (PD-10; GE Healthcare, Uppsala, Sweden). Para el caso de las proteínas, se realizó además un paso extra que implicó uso de cromatografía de alto rendimiento en fase reversa (RP-HPLC) para purificarlos y luego chequear por espectrometría de masa el producto final purificado (Ver Capítulo I). Esto último se hizo fundamentalmente porque las proteínas poseían hasta 2 residuos argininas en los que –aunque con baja probabilidad- se pudo haber conjugado el FITC y porque al experimentar degradación (proteólisis natural de R105) se requería separar esta especie espuria de menor PM (Ver Capítulo I). Los péptidos derivatizados con FITC no contenían lisinas ni argininas, por lo que el grupo amino terminal fue el único sitio posible para su incorporación. La doble cuantificación de estos reactivos marcados (cuantificación del total de proteína o péptido - Absorbancia a 280nm y/o cuantificación por enlaces- y cuantificación de proteína o péptido marcado –Absorbancia a

495nm-) permitió la determinación del razón FITC/Proteína (F/P), los cuales fueron para todos los casos superiores a 0.75.

Control de Calidad y Cuantificación de péptidos y proteínas

Todas las proteínas expresadas de manera recombinante y posteriormente purificadas alcanzaron una pureza superior al 95% verificadas por geles desnaturalizantes de poliacrilamida (SDS-PAGE) y por espectrometría de masa MALDI-TOF.

Cuantificación de las proteínas y péptidos

La cuantificación de las proteínas y péptidos se realizaron utilizando alguno los métodos siguientes abajo:

Método de Bradford: este método utiliza el Azul Brillante de Coomasie G-250, cuyos grupos aniónicos al reaccionar con los aminoácidos (principalmente triptófano, tirosina, histidina, fenilalanina y arginina) lo convierten de su forma leuco (color marrón=naranja) a una forma intensamente azul, la cual absorbe a 595nm [249]. Este método requiere la construcción de una curva de calibración que en este caso usó como patrón Seroalbúmina Bovina (BSA), para lo cual se armó un “stock” de 5ml a concentración 10ng/μl. Luego se armó la siguiente curva , en tubos eppendorff (por duplicado cada punto de concentración, total 12 tubos).

Tabla MM1. Curva de calibración del Método de Bradford

	Curva BSA (ng/μl)	BSA 10ng/μl (μl)	H2O MQ (μl)	Bradford (μl)
1	0	0	800	200
2	2	160	640	200
3	4	320	480	200
4	6	480	320	200
5	8	640	160	200
6	10	800	0	200

Se pipetearon los reactivos: primero el agua, después BSA y por ultimo el Bradford y se agitó. Se dejó incubar por 15 minutos y se midió en cubetas de plástico en el espectrofotómetro Uva 595nm.

Para las muestras a cuantificar se armaron tubos por triplicado con 800 μl de agua MQ, 200 μl de Bradford y 5-10 μl de una dilución de la muestra. Las muestras fueron medidas por triplicado y sus concentraciones halladas por interpolación en la curva de calibración.

Absorbancia a 280 nm/275nm: el residuo aromático Triptófano (W) presenta una coeficiente de extinción o absorptividad molar máximo a 280nm (ϵ) de $5500 \text{ M}^{-1} \text{ cm}^{-1}$ y en el caso de Tirosina (Y) a 275nm (ϵ) de $1490 \text{ M}^{-1} \text{ cm}^{-1}$. Su contribución a la absorptividad molar total de la proteína o péptido se contabiliza como $\epsilon_{\text{TOTAL}} = \# \epsilon_{\text{W}} + \# \epsilon_{\text{Y}}$, donde # es el número total de residuos W y Y en la secuencia y permite el cálculo final concentración [250].

Cuantificación por enlaces o Absorbancia a 220 nm: el enlace peptídico absorbe a 220 nm, lo que independiza a este método de cuantificación de la presencia de residuos aromáticos en la secuencia (para medir por *Absorbancia a 280 nm/275nm*) y también de interacciones entre reactivos y los residuos aminoacídicos (como en *Método de Bradford*). Este método fue especialmente útil para cuantificar las proteínas desordenadas, cuya cuantificación por Bradford resulta poco reproducible. En la misma se realizan las mediciones en condiciones de pH ácido, construyendo una curva de calibración con BSA que posee 606 enlaces y un peso molecular (MW) de 66000 g/mol. Para ello, se armó un “stock” de concentración 400 ng/ μl de BSA (0.4 ug/ μl), el cual se chequeó por Absorbancia a 280nm (BSA tiene ϵ : $48.8 \text{ mM}^{-1} \text{ cm}^{-1}$, 10uM corresponden 0.66mg/ml). Se armó la siguiente curva, en tubos eppendorff (por duplicado cada punto de concentración, total 12 tubos).

Tabla MM2. Curva de calibración del Método de cuantificación por enlaces

	HCl fumante (μl)	Volumen final(μl)	BSA stock (μl)	H2O MQ (μl)
1	5	520	0	515
2	5	520	13	502
3	5	520	27.5	487.5
4	5	520	40.5	474.5
5	5	520	53.5	461.5
6	5	520	67	440

Se pipetearon los reactivos: el agua primero, después BSA y por ultimo el HCl fumante y se agitó. Se dejó incubar *30 minutos* y se midió en cubeta de cuarzo en espectrofotómetro UV a longitud de onda 220 nm. Para las muestras a cuantificar se armaron tubos por triplicado con alrededor de 5 o 10 μl de muestra concentrada o diluida (al 1/2, 1/4, etc.), 5 μl de HCl fumante y agua csp 520 μl finales. La cantidad de muestra añadida fue tal, que el valor de la señal a 220 nm se encontró dentro de los valores medios de la curva. El rango de μM enlaces cubierto por la curva fue de 0 a 500 μM de enlaces aproximadamente, con el “stock” cercano a 400ng/ μl . La concentración de las muestras fueron halladas por interpolación en la curva de calibración.

Absorbancia a 495 nm: para el caso de los péptidos y proteínas con marca fluorescente, la concentración de FITC fue determinada a pH 7.0 a 495nm usando como coeficiente de extinción o absortividad molar $\xi = 75000 \text{ M}^{-1} \text{ cm}^{-1}$ [237], el cual permitió la determinación del razón FITC/Proteína (F/P). los cuales fueron para todos los casos superiores a 0.8.

Todos los reactivos se cuantificaron al menos por dos métodos independientes y se consideraron correctos cuando mostraban un error inter-método menor al 15%. Para el caso de RbAB, se utilizaron el método de Bradford, porque la proteína por su estructura y contenido aminoacídico reacciona bien con el colorante (Coomasie G-250), y por Absorbancia a 280nm, debido a su alta absortividad molar (Ver Tabla abajo). Para toda la serie de péptidos y proteínas que no fuesen RbAB, no se utilizó el método de Bradford ya que por el carácter desordenado y composición aminoacídica de estos péptidos y proteínas, el colorante de Bradford no se asocia bien con ellas y las cuantificaciones probadas no mostraban reproducibilidad. En el caso de las proteínas AdE1A.WT y mutantes, fue de rutina utilizar el método de cuantificación por enlaces y en los casos en dónde se sabía que se trabajaba con altas concentraciones de estas proteínas, se confió en los valores arrojados por la Absorbancia a 275 nm (sólo contenían tirosinas, con su máximo de absorción a esta longitud de onda). Para los péptidos también se utilizó de rutina el método de cuantificación por enlaces. Para el caso del péptido E2F2 también se utilizó Absorbancia a 280nm por contener un triptófano y una tirosina, y para AdE1A-CR1, Absorbancia a 275nm por contener una tirosina (Ver Tabla abajo).

Para las proteínas y péptidos con éstos residuos en su secuencia, se utilizaron las siguientes absortividades molares:

Tabla MM3. Coeficientes de extinción molar de las especies usadas en la tesis

$\xi_{275-280} (\text{M}^{-1} \text{ cm}^{-1})$	
AdE1A-CR1 (1Y)	1490
E2F2 (1 W, 1Y)	6990
AdE1A.WT (2Y)	2980
AdE1A- Δ L (2Y)	2980
AdE1A- Δ E (1Y)	1490
RbAB (3 W, 16 Y)	40340

Electroforesis en geles de poliacrilamida con dodecil sulfato de sodio (SDS-PAGE)

Todas las proteínas fueron evaluadas siendo sembradas en un gel de poliacrilamida al 15%, previamente tratadas con una solución de carga que contiene 125 mM Tris-HCl, pH 6.8; 4% SDS; 0.02% Azul de Bromofenol; 25% v/v glicerol, 200 mM β -Me y sometidas a calentamiento a 95 °C por 5 minutos. Los geles fueron teñidos con Azul Brillante de Coomassie y desteñidas con posterioridad, para revelar el patrón de corrimiento que fue evaluado por comparación con los marcadores de peso molecular [116].

Espectrometría de masa MALDI-TOF

Tanto péptidos sintéticos como proteínas recombinantes fueron analizadas por espectrometría de masa MALDI-TOF (Ionización/desorción láser asistida por matriz- Tiempo de Vuelo, por sus siglas en inglés del inglés). Los péptidos que poseen un peso inferior a 3 kDa, fueron tratados con una matriz de ácido α -ciano- 4-hidroxicinámico (HCCA) (Sigma), mientras que las proteínas fueron tratadas con matriz ácido sinapínico (SA) (Sigma), mezclando 1 μ l de muestra en 0.1% TFA con 1 μ l de solución de HCCA o SA (suspendidos en 60% acetonitrilo/0.1% TFA). El equipo utilizado para obtener los perfiles masa/carga fue un MALDI-TOF Microflex (Bruker, Billerica MA, EEUU) y las mediciones se hicieron en modo lineal utilizando el programa Flex-Control del paquete software del propio equipo. Los pesos moleculares teóricos (PM_{TEO}) de péptidos y proteínas fueron calculados utilizando la herramienta informática disponible en <https://web.expasy.org/protparam/>. Estos sirvieron para la interpretación de los obtenidos por el análisis de masa.

Dicroísmo Circular

Las mediciones de los espectros de Dicroísmo Circular en el UV-lejano fueron realizados en un equipo espectropolarímetro Jasco J-810 (Nikota, Japón) equipado con un termostato Peltier, utilizando cubetas de cuarzo Spectrasyl (Hellma, EEUU) con una longitud de paso óptico de 0.1 o 0.2cm. Cada medición se realizó por quintuplicado y el espectro final fue obtenido al promediar estas cinco medidas. Las mismas fueron realizadas entre 195 y 260 nm, con los siguientes parámetros: sensibilidad estándar, velocidad de escaneo 100nm/min, tiempo de respuesta de 4 segundos, ancho de banda de 2 nm y data pitch de 0.1 nm. El espectro generado al medir el tampón solo (blanco) fue restado de todas las medidas. Todas las mediciones fueron realizadas en tampón 10mM fosfato de sodio pH 7.0 a 20 ± 1 °C.

La elipticidad molar residual media $[\theta]_{MRW}$ fue calculada para cada espectro aplicando a los datos crudos obtenidos, la siguiente ecuación:

$$[\theta]_{MRW} (\text{deg cm}^2 / \text{dmol}) = \left(\frac{\text{mdeg}}{10 \times L \times [M] \times N} \right) \quad (1)$$

donde L es la longitud de paso óptico (cm), $[M]$ es la concentración de la muestra (en molar) y N es el número de enlaces peptídicos de la muestra (número de aminoácidos totales – 1).

Análisis del contenido de α -hélice en AdE1A.WT:

Se realizaron titulaciones con cantidades crecientes de Trifluoretanol (TFE) (ICN, Biomedicals Inc.) desde 0 a 65% (v/v) a 20 ± 1 °C. Se trabajó a concentración inicial de 15 μM de la proteína y se armaron 14 tubos con agregados creciente de TFE (con incrementos de 5% entre tubo y tubo). Los espectros luego fueron recolectados y corregidos luego por dilución. Los datos obtenidos de las titulaciones con TFE no pudieron ser ajustadas a un modelo de transición coil-hélice, por lo que el porcentaje de residuos en conformación de α -hélice fue calculado a partir de los valores de elipticidad molar a 222nm, a partir de la siguiente relación empírica [251]:

$$[\theta]_{222 \text{ nm}} = -39500 \times \left(1 - \frac{2.57}{n} \right) \quad (2)$$

dónde n , es el número de residuos de la proteína, considerando que esta relación empírica define el $[\theta]_{222 \text{ nm}}$ para una proteína con el 100% en conformación α -hélice.

Efecto del pH en el contenido de α -hélice en AdE1A.WT:

La dependencia de la formación de α -hélice con el pH fue medido a 16 μM AdE1A.WT a 20 °C en 4 condiciones: a) tampón 20 mM Tris:Cl pH 7.5; b) tampón 20 mM Tris:Cl pH 7.5 con 50% TFE; c) tampón 10mM formiato de sodio pH 5.0; d) tampón 10mM formiato de sodio pH 5.0 con 50% TFE trabajando a concentración 10 μM de la proteína. Condiciones de pH aún menores fueron ensayadas (pH 3.0 and 4.0) pero la proteína experimentó agregación.

Evaluación del contenido de estructura Poliprolina tipo II en AdE1A.WT:

La propensión a estructura tipo Poliprolina tipo II (PII) fue evaluada en AdE1A.WT midiendo los espectros de CD en el UV-lejano a 5°C, 20°C y 80°C. Como la estabilización de estructura PII ocurre a bajas temperaturas, por lo que la inducción de la banda PII

característica a 218 nm se evidencia por la diferencia entre el espectro de CD a 5°C menos el espectro de CD a 80°C.

Desnaturalización térmica de AdE1A.WT:

El escaneo térmico para evaluar la estabilidad de AdE1A.WT, se realizó con una muestra a concentración $1 < 0 \mu\text{M}$ de la proteína entre 20°C y 80°C a una velocidad de escaneo de 3 °C/min y la elipticidad cruda a 200 nm fue seguida en función de la temperatura.

Cromatografía de Exclusión Molecular y Dispersión de Luz Estática y Dinámica

Cromatografía analítica de exclusión molecular:

La cromatografía analítica de exclusión molecular (*Size exclusion chromatography* o SEC) fue realizada en una columna Superdex 75 HR 10/30 (24 ml) (GE Healthcare). Esta columna fue calibrada con los siguientes estándares de proteínas globulares: BSA (66 kDa), MBP (45 kDa) y Lisozima (14.3 kDa). El volumen muerto (*Void volume* o V_0) y el volumen total de la columna ($V_0 + V_i$ o V_T) fue determinado con la carga de Blue Dextrano y Acetona, respectivamente. Todas las corridas fueron hechas inyectando unos 100 μl de proteína en concentración a un flujo de 0.4 ml/min en tampón 20 mM fosfato de sodio, 200 mM NaCl, 1 mM DTT a pH 7.0 a temperatura ambiente. Para cada una de las proteínas y complejos proteicos analizados por SEC, se calculó el coeficiente de partición (K_{av}) conociendo los volúmenes de elución particulares (V_e), a partir de la fórmula:

$$K_{av} = \frac{(V_e - V_0)}{(V_T - V_0)} \quad (3)$$

Los valores de los Pesos Moleculares aparentes (PM_{app}) de proteínas libres y complejos fue hallada al interpolar sus valores de K_{av} en la curva de calibración construida con el $-\log$ del peso molecular de los estándares vs sus valores K_{av} .

Dispersión de luz estática:

La dispersión de luz estática (*Static Light Scattering* o SLS) fue llevado a cabo acoplado a las mediciones de SEC, para determinar el peso molecular medio de los picos individuales de las proteínas y también para la determinación de las estequiometrías de interacción de los complejos de RbAB con AdE1A.WT y mutantes. En el SLS, la intensidad de la luz dispersada es medida en función del ángulo de dispersión para obtener la información requerida del analito

en solución, como por ejemplo el peso molecular, el cual determinado en forma independiente de su comportamiento hidrodinámico. Para ello se usó un instrumento detector Precision Detector PD200 (Precision Detectors Inc, China) acoplado en tándem a un sistema de cromatografía de alto rendimiento (HPLC) y un refractómetro diferencial LKB2142. Las señales de dispersión de luz a 90° (LS) y el índice de refracción (RI) de las proteínas y complejos que eluyeron, fueron analizados con el software Discovery32 suministrado por Precision Detectors.

Dispersión de luz dinámica

La distribución del radio hidrodinámico de AdE1A.WT fue medido utilizando dispersión de luz dinámica (*Dinamic Light Scattering* o DLS) en un espectrómetro Wyatt Dynapro (Wyatt Technologies, EEUU). Todas las medidas fueron realizadas a 2 mg/ml (165 μM) en tampón 20mM fosfato de sodio, 200 mM NaCl, 1 mM DTT pH 7.0 a temperatura ambiente. Las muestras fueron previamente filtradas con filtros de 0.22 μM (Millipore) y dispuestas en una placa de vidrio negra de 96 pocillos (In Vitro Scientific P96-1.5H-N) cubiertas por una tapa de vidrio de alto rendimiento (0.17+/-0.005mm) antes de realizar las mediciones.

Cálculos de los Radios hidrodinámicos:

Para el cálculo del radio hidrodinámico (R_H) experimental de las proteínas en su estado nativo se utilizaron las formulaciones empíricas desarrolladas por Uversky [119, 120, 252]. En su trabajo de revisión de datos de literatura preexistentes, recopiló los valores de PM_{TEO} y sus R_H conocidos para distintos grupos de proteínas. Con datos de 20 proteínas en su estado nativo -globular- (G), 21 proteínas desnaturalizadas por acción de Cloruro de guanidinio ($D^{Gdm.Cl}$), 7 proteínas desnaturalizadas por acción de Urea (D^U), 8 proteínas en estado conformacional tipo “*Molten-Globule*” (MG) y 7 proteínas en estado conformacional tipo “*Pre-Molten-Globule*” (PMG), logró encontrar que los logaritmos del radio hidrodinámico y el peso molecular teórico de las proteínas ($\log(R_H)$ vs $\log(PM_{TEO})$) ajustaban a relaciones lineales, con parámetros que dependían de las conformaciones en las que se encontraban dichas proteínas. Así, describió las siguientes fórmulas empíricas:

$$\log R_H^G = -0.204 + 0.357 \log PM \quad (4)$$

$$\log R_H^{D^{Gdm.Cl}} = -0.723 + 0.543 \log PM \quad (5)$$

$$\log R_H^{D^U} = -0.649 + 0.521 \log PM \quad (6)$$

$$\log R_H^{\text{MG}} = -0.053 + 0.334 \log \text{PM} \quad (7)$$

$$\log R_H^{\text{PMG}} = -0.210 + 0.392 \log \text{PM} \quad (8)$$

Estas fórmulas empíricas fueron validadas por ensayos de Cromatografía de Exclusión Molecular propios, en los que trabajó con proteínas conocidas que se encontraban las mismas conformaciones/condiciones que los sets de proteínas estudiados. En primer lugar, logró relacionar los V_e de las proteínas en sus diferentes estados con sus respectivos R_H , con la relación:

$$\frac{1000}{V_e} = (0.724 \pm 0.0003)R_H + (52.1 \pm 0.3) \quad (9)$$

Como esta relación tiene parámetros dependientes de la columna utilizada (en su caso una Superosa 12) su utilización es limitada. No obstante, al ser SEC es una técnica que no modifica los estados conformacionales de las proteínas, observó que sus R_H experimentales hallados por esta metodología (usando la ecuación 9) concordaron con los R_H reportados en la literatura, los cuáles siguen las relaciones antes expuestas (ecuaciones 4 a 8). Ahora bien y por tanto, para una proteína en estudio de la que no se conoce su estado conformacional, el PM_{ap} hallado por SEC puede ser utilizado para calcular su R_H experimental utilizando la ecuación 4, que la supone en su estado nativo.

Para los cálculos de los R_H teóricos predichos se utilizaron las formulaciones de Wilkins [117, 118] y de Marsh y Forman-Kay [117]. Ambos están basados en Ley de escalamiento teórica para la predicción de dimensiones de homopolímeros, que sigue la forma

$$R_H = R_0 N^{\nu} \quad (10)$$

en la que R_0 es una constante y N es el número de unidades monoméricas que conforman el polímero. El exponente (ν), llamado exponente de escalamiento o de Flory es predicho que asuma valores 3/5 o 1/3 según el homopolímero esté totalmente solvatado o sus cadenas totalmente colapsadas, respectivamente. Como lo demostró Flory, un homopolímero con una cadena totalmente solvatada es observado en un buen solvente en el cual las interacciones polímero-solvente están favorecidas sobre las interacciones polímero-polímero, mientras que una cadena colapsada se observa en presencia de un solvente que favorece las interacciones polímero-polímero sobre las interacciones polímero-solvente. De esta forma, el exponente

puede llegar a valer 1/2 en “solventes Flory” en las que las interacciones polímero-solvente y polímero-polímero son equivalentes [253, 254]

En su trabajo, Wilkins encontró valores R_0 y ν , que hacen a la fórmula de la Ley de escalamiento aplicable para proteínas globulares (G) y desnaturalizadas (D), teniendo en cuenta la longitud de sus cadenas polipeptídicas. Para ello trabajó con una serie de 7 proteínas globulares y 7 proteínas bajo condiciones altamente desnaturalizantes a las que le determinó su R_H por PFG-RMN, para luego encontrar que estos valores ajustaban a las siguientes relaciones empíricas:

$$R_H^G = 4.75 \times N^{0.29} \quad (11)$$

$$R_H^D = 2.21 \times N^{0.57} \quad (12)$$

Estas relaciones luego fueron validadas por comparación con datos de literatura y otras predicciones teóricas. Ambas fueron usadas para predecir el R_H de las tres variantes de AdE1A.

En el trabajo de Marsh y Forman-Kay Wilkins, se realizó una búsqueda en la literatura de R_H conocidos que conformaron tres sets de proteínas: 20 globulares (G), 27 químicamente desnaturalizadas (D) y 32 intrínsecamente desordenadas (IDP). A cada grupo ajustaron la Ley de escalamiento y encontraron valores R_0 y ν , permitiéndoles definir las siguientes fórmulas:

$$R_H^G = 4.92 \times N^{0.285} \quad (13)$$

$$R_H^D = 2.33 \times N^{0.549} \quad (14)$$

$$R_H^{IDP} = 2.49 \times N^{0.509} \quad (15)$$

En la validación de estas relaciones, encontraron un buen acuerdo con los datos experimentales para las proteínas globulares y para las desnaturalizadas (RMSD 0.56 Å y 1.20 Å respectivamente). No obstante para las IDP, el acuerdo fue peor (RMSD 4.13 Å). Además habían observado que en varios casos aún cuando se espera que las IDP sean más compactas que las proteínas químicamente desnaturalizadas, las IDPs estaban más expandidas. Estas cuestiones los llevó a hacer un análisis en mayor profundidad sobre el R_H y grado de compactación de IDPs, que los condujo al desarrollo de la fórmula que tiene en cuenta

componentes de la secuencia los cuáles influyen directamente en los R_H de IDPs. Esta ecuación presenta la siguiente forma:

$$R_H^{IDP} = (AP_{pro} + B)(C|Q| + D)S_{his} \times 2.49 \times N^{0.509} \quad (16)$$

donde N es el número total de residuos en la secuencia proteica, P_{pro} es la fracción de prolinas, Q es la carga neta, S_{his} es un factor si la proteína tiene tag de histidinas, A - D son constantes obtenidas de ajustes. Cuando validaron los R_H obtenidos con esta fórmula, el RMSD fue de 2.37 Å, mucho mejor que el RMSD 4.13 Å, que se veía para la fórmula básica de escalamiento para IDPs. Las cuatro formulaciones fueron utilizadas en este trabajo para predecir el R_H de las tres variantes de AdE1A.

Espectroscopía de Fluorescencia

Las medidas de Espectroscopía de Fluorescencia se llevaron a cabo en un espectropolarímetro Jasco FP-6200 (Nikota, Japon) ensamblado en geometría L y acoplado a un termostato Peltier. Las longitudes de onda de excitación y de emisión fueron de 495 y 520 nm respectivamente, con un ancho de banda de 4nm. Todas las mediciones fueron hechas en tampón 20 mM fosfato de sodio, pH 7.0, 200 mM NaCl, 2 mM DTT y 0.1% Tween-20 a 20 ± 1 °C. Para cada punto de la titulación, se hicieron tres medidas de anisotropía (r) (y fluorescencia (F), sólo en el caso puntual de la interacción RbAB con AdE1A.WT). Estas 3 mediciones luego fueron promediadas. Para las mediciones de anisotropía fue usado un Factor G por cada sonda distinta evaluada, el cual fue calculado y configurado directamente por el equipo.

El parámetro medido en estos experimentos: anisotropía (r), queda definido como la razón entre el componente polarizado y la intensidad total:

$$r = \frac{(I_z - I_y)}{(I_x + I_y + I_z)} \quad (17)$$

cuando la excitación está polarizada a lo largo del eje z, la emisión del fluoróforo es simétrica en torno a este eje y estadísticamente resulta que $I_x = I_y$, dónde $I_y = I_{vh}$ e $I_z = I_{vv}$, tal que:

$$r = \frac{(I_{vv} - I_{vh})}{(I_{vv} + 2I_{vh})} \quad (18)$$

Para corregir el sesgo del sistema de detección de la polarización, se introduce un factor experimental de corrección G, por lo que r queda definido entonces como:

$$r = \frac{(I_{vv} - G \times I_{vh})}{(I_{vv} + 2G \times I_{vh})} \quad (19)$$

Titulaciones directas:

Para realizar las titulaciones directas, se tituló una concentración fija de péptido o proteína marcadas con FITC contenidos en una cubeta de cuarzo, con cantidades crecientes de una solución concentrada de RbAB hasta que se alcanzó la saturación de la señal. La máxima dilución permitida fue del 20% del volumen inicial. Para asegurarse de que las mediciones fueran realizadas en un estado estacionario, se dejó que las mezclas alcancen el equilibrio por al menos dos minutos. Los fenómenos de agregación y oxidación de los péptidos y proteínas conteniendo cisteínas, se previnieron controlando cuidadosamente condiciones experimentales con el agregado de un agente reductor de alta pureza (DTT). Se realizaron múltiples medidas independientes para asegurar la reproducibilidad de los resultados. El número de repeticiones para cada interacción testeada fue variable (tres como mínimo) y se presentan la Tabla A2.1 del Anexo.

Las titulaciones directas se realizaron en dos regímenes de concentración. Las titulaciones hechas a concentraciones al menos 10 veces superior a las constantes de disociación (K_D), permitieron una estimación de la estequiometría de las interacciones. En este régimen, las señales de anisotropía y fluorescencia aumentan linealmente con las concentraciones crecientes de agente titulante hasta alcanzar la saturación. La estequiometría de los complejos fue determinada con la intersección de las dos rectas obtenidas de la extrapolación del ajuste lineal de los primeros y últimos valores de las señales obtenidas en la titulación. Todas las mediciones saturaron a una relación molar 1 : 1 titulado : titulante, indicando una estequiometrías (n) igual a 1. Las titulaciones hechas a concentraciones sub-estequiométricas (en el rango de las K_D) permitieron una estimación de la constante de disociación al equilibrio (K_D) para todos los complejos ensayados.

Ajuste de los datos obtenidos en las titulaciones

Como todos los complejos resultaron tener una estequiometría 1 : 1, el ajuste de los datos obtenidos se hizo a un modelo de asociación bimolecular simple, donde las especies A y X interactúan:



y donde la constante de asociación (K_A) del complejo AX puede definirse como:

$$K_A = \frac{[AX]}{[A] \times [X]} \quad (21)$$

y por lo tanto la constante de disociación (K_D) es K_A^{-1} , entonces:

$$K_D = \frac{[A] \times [X]}{[AX]} \quad (21)$$

Para el ajuste al modelo bimolecular simple, se consideró que la señal de anisotropía total medida Y, es la suma de las contribuciones de las señales de anisotropía de la especie libre Y_F (que es la especie A –péptido o proteína- derivatizada con FITC) y la del complejo Y_B (la especie derivatizada con FITC acomplejado con RbAB), definida por la relación:

$$Y = Y_F \times f_F + Y_B \times f_B + Y_L \quad (22)$$

donde f_F y f_B son las fracciones de cada una de las especies tal que $f_F + f_B = 1$ e Y_L es un término lineal que depende de la concentración de RbAB: $[X]_0$, por lo que $Y_L = C \times [X]_0$. Si:

$$f_F = \frac{[A]}{[A_0]} = \frac{[A_0] - [AX]}{[A_0]} \quad (23)$$

$$f_B = \frac{[AX]}{[A_0]} \quad (24)$$

donde $[A_0]$ es la concentración total fija de especie derivatizada con FITC, entonces $[A]$, $[X]$ y $[AX]$ en el equilibrio pueden calcularse como:

$$[A] = [A_0] - [AX] \quad (25)$$

$$[X] = [X_0] - [AX] \quad (26)$$

de modo que

$$[AX] = \frac{([A_0] - [AX]) \times ([X_0] - [AX])}{K_D} \quad (27)$$

la cual es una ecuación cuadrática que puede resolverse:

$$[AX] = \frac{([A_0] + [X_0] + K_D) + \sqrt{([A_0] + [X_0] + K_D)^2 - (4 \times [A_0] + [X_0])}}{2} \quad (28)$$

Y si se reemplaza [A] y [AX] en la ecuación (11) y reorganizando se obtiene:

$$Y = Y_F + \frac{(Y_B - Y_F)}{[A_0]} \times \frac{([A_0] + [X_0] + K_D) + \sqrt{([A_0] + [X_0] + K_D)^2 - (4 \times [A_0] + [X_0])}}{2} + C \times [X_0] \quad (29)$$

donde K_D es la constante de disociación (en Molar) y constituye un parámetro que se desea estimar. $[A_0]$ es conocido y representa la concentración fija de la especie derivatizada con FITC y se considera como constante en esta ecuación ya que se toma como despreciable su dilución a medida que avanza la titulación. Los valores Y_B e Y_F pueden ser estimados a partir de los valores de anisotropía final e inicial de las curvas de titulación. El parámetro $C \times [X]_0$, toma en cuenta fenómenos de "bleaching" o también de agregación muy pequeños que pueden tener lugar a concentraciones proteicas altas. De este modo, la ecuación (29) queda expresada sólo como una función de la variable $[X_0]$, es decir que se mide Y a distintas concentraciones del titulante $-[RbAB]$ en cada punto de la curva- y luego se estiman Y_B e Y_F y K_D .

La ecuación: (29), fue implementada y los datos fueron ajustados usando el software Profit 7.0 (Quantumsoft, Zurich, Switzerland). Todos los ajustes se hicieron usando el algoritmo Levenberg Marquardt, lo que permitió obtener valores para cada parámetro y su correspondiente desvío estándar. Esta ecuación quedó expresada entonces como variable dependiente de $[RbAB]$, así:

$$Y = a_1 + a_2 \times \frac{([X_0] + a_3 + a_4) + \sqrt{([X_0] + a_3 + a_4)^2 - (4 \times a_3 \times [X_0])}}{2} + a_5 \times [X_0] \quad (30)$$

tal que en el ajuste:

$$[X_0] = [RbAB]$$

$$a_1 = Y_F$$

$$a_2 = (Y_B - Y_F) / [A_0]$$

$$a_3 = [A_0]$$

$$a_4 = K_D$$

$$a_5 = C$$

En primer lugar, cada curva de titulación fue individualmente ajustada punto a punto, obteniendo un valor de $a_4 = K_D$. Los parámetros a_1 y a_2 (Y_F e Y_B) fueron ajustados (usando datos crudos) o fijados a valores $Y_B - Y_F = 1$ e $Y_F = 0$, tras normalización de las señales. El parámetro $C \times [X]_0$ fue ajustado o fijado a cero según fuese necesario para lograr el ajuste de mejor calidad. Las curvas de titulación fueron hechas al menos a tres concentraciones diferentes de la especie derivatizada con FITC, para asegurar que el modelo describiese un comportamiento similar para el sistema independientemente de la concentración de dichas sondas.

Un único valor de K_D (K_D global) fue determinado para cada interacción luego de hacer un ajuste global de los datos recolectados habiendo trabajado a distintas concentraciones de una misma sonda $[A_0]$, y aplicando reiterativamente la ecuación (30), donde a_3 tiene la forma:

$$a_3 = [A_0] \times z \quad (31)$$

y z es un factor que multiplica la concentración más baja de sonda que fue probada $[A_0]$, de modo que $[A_0] \times z$ representa cada concentración más alta de sonda testeada. El número de veces que se aplicó en este caso la ecuación 30, es igual a la cantidad de concentraciones diferentes a la que fue testeada la misma sonda. Para estos casos, $[A_0]$ fue dejado como parámetro libre a fin de lograr el ajuste de mejor calidad posible y se utilizaron únicamente valores de señales normalizados, tal que siempre $Y_B - Y_F = 1$ e $Y_F = 0$. Para cada interacción, K_D individuales y global presentaron muy buen acuerdo (Ver Tabla A2.1 del Anexo).

La normalización de las señales de anisotropía (r) se hizo para que los cambios totales de esa señal en las titulaciones llevadas a cabo a diferentes concentraciones de una misma sonda, represente el mismo rango de variación y sea constante, tal que $Y_B - Y_F = 1$ e $Y_F = 0$. El valor de la señal normalizada (Y_N) fue determinado así:

$$Y_N = \frac{Y_x - A}{B - A} \quad (32)$$

dónde Y_x es el valor medido experimental de la señal r , A corresponde a Y_F , que es valor inicial de anisotropía (de la sonda libre) y B tiene la forma $a^* [X_0] + b$, dónde a y b son los parámetros correspondientes a la recta del ajuste de regresión lineal aplicado a los últimos puntos de la

curva monótonica de titulación (en condiciones de concentraciones saturantes donde la señal prácticamente ya no cambia) y $[X_0]$ es [RbAB] punto a punto en toda la curva.

Experimentos de competición:

Los experimentos de competición se llevaron a cabo titulando un complejo preensamblado de RbAB con el péptido E2F2-FITC (en una relación molar 1:1) a $5 \mu\text{M}$ con cantidades crecientes de los siguientes péptidos y proteínas sin marcar E2F2, AdE1ACR1, AdE1ACR2, AdE1.WT y AdE1A- ΔL . Es importante destacar que $5 \mu\text{M}$ es una concentración 5000 veces superior a la K_D de esta interacción, por lo que la misma está muy estabilizada.

La disrupción del complejo [RbAB:E2F2-FITC] fue seguida con el decrecimiento de la señal de anisotropía hasta alcanzar un valor de anisotropía correspondiente a E2F2-FITC libre. Los valores de IC_{50} fueron estimados directamente de las curvas en el punto en el que la concentración del agente desplazante produjo un descenso de la señal a un 50% del valor máximo de anisotropía. Como era esperado, los valores de IC_{50} fueron alrededor de 20 a 60 veces más altos que los valores obtenidos de K_D para cada interacción testeada. Estos cambios relativos obtenidos en los experimentos de competición, sumados al excelente acuerdo entre los valores de K_D obtenidos en los experimentos de Fluorescencia y los de calorimetría de titulación isotérmica (ITC por sus siglas en inglés) confirmaron que la presencia de la molécula de FITC en el extremo N-terminal de los péptidos y proteínas marcadas no modifican sustancialmente sus propiedades de interacción a RbAB.

Calorimetría de Titulación Isotérmica

Los experimentos de ITC se llevaron en colaboración con el laboratorio del Dr. Gary W. Daughdrill, del Centro para el Descubrimiento de Drogas e Innovación (CDDI) de *University of South Florida* (USF) en la ciudad de Tampa, Florida, EEUU, como parte de un proyecto conjunto y de una beca Fulbright –Ministerio de Educación de la Nación, de la que fui beneficiario en el año 2017. El laboratorio del Dr. Daughdrill cuenta con un equipo MicroCal VP-ITC (Malvern Panalytical) utilizado para realizar la mayoría de los experimentos calorimétricos planteados. Otra serie de experimentos se realizaron en las instalaciones del Laboratorio Europeo de Biología Molecular, en Heidelberg, Alemania, con un equipo MicroCal PEAQ-ITC (Malvern Panalytical). Todos los experimentos de ITC fueron realizados en dos equipos: El primero de ellos, un equipo MicroCal VP-ITC (Malvern Panalytical) ubicado en el laboratorio del Dr. Gary W. Daughdrill, del Centro para el Descubrimiento de Drogas e Innovación (CDDI) de *University of South Florida* (USF) en la ciudad de Tampa, Florida, EEUU.

El segundo, un equipo MicroCal PEAQ-ITC (Malvern Panalytical) ubicado en las instalaciones del Laboratorio Europeo de Biología Molecular, en Heidelberg, Alemania.

Los experimentos se realizaron en 20 mM fosfato de sodio, 200 mM NaCl, 5mM β -Me a pH 7.0 y 20.0 ± 0.1 °C, a menos que se indique lo contrario. Previo a las titulaciones las muestras utilizadas en las titulaciones fueron co-dializadas en el tampón previamente descrito por al menos 48hs a 4 ± 1 °C. Las muestras fueron de-gasificadas antes de hacer la carga en la celda y en la jeringa. para evitar injerencias de calores de mezcla y de posibles burbujas en la adquisición de los datos.

Titulaciones directas por ITC:

Las medidas en el MicroCal VP-ITC fueron realizadas aplicando 28 inyecciones de 10 μ l a un flujo de 0.5 μ l por segundo y las hechas en el MicroCal PEAQ-ITC, con 13 inyecciones de 3 μ l cada una. Los datos fueron analizados utilizando el paquete de análisis Origin, propio de cada equipo y tras descartar el primer punto, los datos fueron integrados usando un modelo de interacción a un único sitio, pudiendo determinar estequiometría (n), el cambio en la entalpía (ΔH) y constante de disociación al equilibrio (K_D) de los cuáles se reportaron también el cambio de la energía libre de interacción ($\Delta G = RT \ln K_D$) y el cambio en la entropía de los estados unidos y libres ($T\Delta S = \Delta H - \Delta G$). Para todos los sistemas evaluado, el valor de n fue cercano a 1, indicando una estequiometría 1:1 lo que validó el modelo de ajuste utilizado (Tabla 2.2, Capítulo II).

Evaluación del posible comportamiento alostérico en la interacción AdE1A-RbAB por ITC:

Una serie de experimentos de titulación de ITC fueron hechos para evaluar el comportamiento alostérico en *trans* entre el motivo E2F (AdE1A-CR1) y el motivo LxCxE (AdE1A-CR2) en la interacción de AdE1A con RbAB, cuando la región enlazante no se encuentra uniendo ambos motivos entre sí. En estos experimentos, un complejo formado por RbAB con AdE1A-CR2 (en una relación molar 1:1 a 30 μ M) fue titulado con cantidades crecientes de AdE1A-CR1 y en otro experimento con AdE1A- Δ L, para chequear posibles efectos alostéricos del motivo E2F sobre la interacción del motivo LxCxE con RbAB- LxCxE. Complementariamente, complejos de RbAB pre-saturados con AdE1A-CR1 y con AdE1A- Δ L (ambos a relación molar 1:1 y 30 μ M), fueron titulados con cantidades crecientes de AdE1A-CR2 para evaluar el mismo efecto alostérico pero desde el otro motivo.

Titulaciones de ITC dependientes de temperatura:

Las titulaciones de ITC a diferentes temperaturas de AdE1A-CR1 y Ad5E1A-ΔL con RbAB fueron hechas a 10.0, 15.0, 20.0 y 30.0 ± 0.1 °C y a 10.0, 20.0 y 30.0 ± 0.1 °C respectivamente. Estos experimentos permitieron el cálculo de la capacidad calórica de interacción a presión constante (ΔC_p). Esta capacidad calórica es la cantidad de calor necesaria para elevar en un grado °C la temperatura de una cantidad de sustancia (generalmente 1g de agua pura). A presión constante resulta que:

$$C_p = \frac{dQ_p}{dT} = \frac{dH}{dT} \quad (33)$$

Si la capacidad calórica es constante con la temperatura, la variación de la entalpía será:

$$\Delta H = nC_p \int_{T_1}^{T_2} dT = nC_p \Delta T \quad (34)$$

debido a esto, ΔC_p fue determinada para cada interacción de las pendiente obtenida del análisis de regresión lineal del gráfico ΔH vs temperatura.

Cálculos de parámetros estructurales derivados de ITC:

Parámetros estructurales derivados de los datos de calorimetría fueron estimados por aplicación de ecuaciones semi-empíricas aplicadas a interacciones proteína-ligando [192]. Los cambios en las áreas superficiales polares y no polares accesibles al solvente (en inglés polar, *non-polar Accessible Surface Areas* - ΔASA_p , ΔASA_{np} -) fueron calculadas resolviendo las siguientes ecuaciones de forma secuencial:

$$\Delta H_{int(T^\circ C)} = \Delta H_{int(T_H)} + \Delta C_p (T - T_H) \quad (35)$$

donde $\Delta H_{int(T^\circ C)}$ es el valor experimental de entalpía a la temperatura de trabajo (T) y $\Delta H_{int(T_H)}$ es la entalpía a la temperatura de convergencia entálpica, a la cual la contribución apolar se asume igual a cero que se tomó 295.15 K [196, 255]. Esta ecuación (35) proveyó el valor de $\Delta H_{int(T_H)}$, útil para el cálculo de las ecuaciones que siguen:

$$\Delta H_{int(T_H)} = \Delta h_{np} \Delta ASA_{np} + \Delta h_p \Delta ASA_p \quad (36)$$

$$\Delta C_p = \Delta c_{np} \Delta ASA_{np} + \Delta c_p \Delta ASA_p \quad (37)$$

donde Δh_{np} , Δh_p , Δc_{np} y Δc_p son parámetros empíricos y asumen los siguientes valores: $\Delta c_{np} = 0.45$ $\Delta c_p = -0.26$ en cal /K mol Å^2 y $\Delta h_{np} = -7.35$ y $\Delta h_p = 31.06$ en cal /mol Å^2 [192, 256, 257]. Como $\Delta H_{int(T)}$ y ΔCp ya se conocían, al igual que los parámetros Δh_{np} , Δh_p , Δc_{np} y Δc_p , estas ecuaciones (36) y (37) se resolvieron simultáneamente como sistema de dos ecuaciones con dos incógnitas (ΔASA_p , ΔASA_{np}).

Cálculos de la contribuciones entrópicas en la interacción del motivo E2F con RbAB y número de residuos involucrados:

Los cambios en la entropía asociada a la interacción proteína ligando, fue calculada como sigue:

$$\Delta S_{int(T)} = \Delta S_{solv} + \Delta S_{rt} + \Delta S_{config} \quad (38)$$

donde ΔS_{int} es la entropía a la temperatura experimental (T), ΔS_{rt} es la contribución entrópica rotacional-traslacional, que por consideraciones teóricas y empíricas asumen valores entre -4 y -10 cal/K mol, tomándose como referencia el valor de -8 cal/K mol [192], ΔS_{solv} es la entropía de solvatación o efecto hidrofóbico y ΔS_{config} refleja la contribución de los re-arreglos estructurales que se dan en a interacción. ΔS_{solv} fue calculada como:

$$\Delta S_{solv} = \Delta C_p \ln\left(\frac{T}{T_s^*}\right) \quad (39)$$

donde T es el valor de temperatura experimental (en Kelvin) y T_s^* es el valor de la temperatura de convergencia de entropía (385K) a la cual la contribución apolar se asume igual a cero. El número de residuos involucrados en las interacciones fueron estimados como:

$$X_{res} = \frac{\Delta S_{config}}{\Delta S_{residuo}} \quad (40)$$

donde $\Delta S_{residuo}$ es igual a -4.3 cal/K mol y representa el valor promedio estimado de contribución por residuo a la ΔS_{config} .

Resonancia Magnética Nuclear

Los experimentos de Resonancia Magnética Nuclear (RMN) fueron realizados en el laboratorio del Dr. Gary W. Daughdrill (CDDI- USF, Florida, EEUU). Estos experimentos se llevaron a cabo usando un espectrómetro Varian VNMRS 800 MHz (Agilent, California, EEUU) equipado con una crio-sonda de gradiente en el eje Z para experimentos con pulsos de triple resonancia. Una serie de experimentos bidimensionales ^1H - ^{15}N HSQC de elevada sensibilidad y tridimensionales HNCACB, HNCO y CBCA(CO)NH [258-260] fueron hechos para la asignación de las resonancias de las cadenas carbonadas en muestras de AdE1A.WT, AdE1A- ΔE y AdE1A- ΔL marcadas uniformemente con ^{13}C - ^{15}N a 700 μM , 975 μM y 850 μM respectivamente. Todas las mediciones se realizaron en 10 % D₂O, 20 mM fosfato de sodio, 200 mM NaCl, 2 mM DTT a pH 7.0 y a 25 °C. Los experimentos HSQC se adquirieron con ventanas espectrales de 9689.9 Hz y 1024 puntos complejos para la dimensión t₁ y con una ventana de 2106.4 Hz y 128 puntos complejos para la dimensión t₂. Los experimentos HNCACB se adquirieron usando ventanas espectrales de 9689.9, 14075.1, y 2106.4 Hz, con 1024, 128 y 32 puntos complejos para las dimensiones t₁, t₂ y t₃ respectivamente. Los experimentos HNCO usaron ventanas espectrales de 9689.9, 2010.4 y 2106.4 Hz con 1024, 64 y 32 puntos complejos para las dimensiones t₁, t₂ y t₃ respectivamente. Los CBCA(CO)NH usaron ventanas de 9689.9, 14072.6 y 2106.4 Hz con 1024, 128 y 32 puntos complejos para las dimensiones t₁, t₂ y t₃ respectivamente.

Los paquetes de software NMRPipe y NMRViewJ fueron utilizados para procesar y analizar todos los espectros de RMN [261].

Experimentos de asignación de núcleos:

La asignación logró completarse para más del 87% de $^{13}\text{C}_\alpha$, 83% de $^{13}\text{C}_\beta$, 71% de $^{13}\text{C}'$, 71% de ^{15}N de las amidas y 71% ^1H de las amidas para 114 residuos en AdE1A.WT y 113 residuos en las mutantes AdE1A- ΔL y AdE1A- ΔE . Las 18 Prolinas sobre un total 114 o 113 residuos de las 3 secuencias, no pudieron ser asignadas debido a limitaciones técnicas de los experimentos bi- y tridimensionales aplicados.

Determinación de los desplazamientos químicos secundarios:

La desplazamientos químicos residuo-específicos en *random-coil* fueron generados para las tres secuencias usando una biblioteca de desplazamientos químicos corregidos por residuos colindantes en Proteínas intrínsecamente Desordenadas (IDP por sus siglas en inglés) [215]. Los desplazamientos químicos secundarios ($\Delta\delta$), fueron calculados como la

substracción de los desplazamientos químicos obtenidos experimentalmente menos los desplazamientos químicos en “*random coil*”. Usando los desplazamientos químicos de ^1H de amidas, ^{15}N de amidas, $^{13}\text{C}_\alpha$, $^{13}\text{C}_\beta$, y $^{13}\text{C}'$ fueron obtenidas las Distribuciones de Población de Estructura Secundaria para cada una de las tres especies AdE1A.WT, AdE1A- ΔL y AdE1A- ΔE , haciendo uso del análisis $\delta 2d$ [219]. Las perturbaciones en los desplazamientos químicos en AdE1A.WT producido por las mutaciones, fueron calculados residuo por residuo a partir de la siguiente expresión [262]:

$$\sqrt{(\Delta^1\text{H})^2 + \left(\frac{\Delta^{15}\text{N}}{6.5}\right)^2} \quad (41)$$

Experimentos de relajación:

Experimentos de relajación fueron adquiridos en un espectrómetro Varian VNMRs 600 MHz (Agilent, California, EEUU) en muestras de AdE1A.WT, AdE1A- ΔE y AdE1A- ΔL marcadas uniformemente con ^{15}N a 420 μM , 690 μM y 680 μM respectivamente. El método de detección inversa 2D RMN se usó para medir relajación espín-red o relajación longitudinal (R1), relajación espín-espín o relajación transversal (R2), relajación rotacional (R1 ρ) y experimentos ^1H - ^{15}N NOE [222]. Todos los experimentos de relajación se adquirieron con ventanas espectrales de 7267.4 Hz y 512 puntos complejos para la dimensión t1 y ventanas espectrales de 1578.9 Hz y 128 puntos complejos para la dimensión t2. Los valores R1 se determinaron colectando espectros bidimensionales con los siguientes tiempos de retraso: 10, 50, 110, 310, 500, 650, 1000, 1500 y 1900 ms. Los valores R2 se adquirieron con los siguientes tiempos de retraso: 10, 30, 50, 90, 110, 150, 170, 190, 210 y 250 ms. Los valores R1 ρ fueron colectados con los tiempos de retraso: 10, 30, 90, 110, 150, 170, 190, 210, 250 y 350 ms. Una función decreciente monoexponencial fue usada para ajustar la intensidad de los picos obtenidos en los experimentos de relajación.

Experimentos TROSY:

Los experimentos bidimensionales ^1H - ^{15}N TROSY se realizaron sobre muestras de AdE1A.WT, AdE1A- ΔE y AdE1A- ΔL marcadas con ^{15}N en estado libre y cada una en complejo en relación molar 1:1 con RbAB a 525 μM (AdE1A.WT), 300 μM (AdE1A- ΔE) y 315 μM (AdE1A- ΔL). Para el caso de del complejo [RbAB:AdE1A.WT] se describe a continuación:

Se trabajó secuencialmente como sigue:

- a. En primer lugar, se preparó una muestra uniformemente marcada con ^{15}N de AdE1A.WT a concentración 525 μM , volumen final 600 μl y en el equipo Varian VNMRS 800 MHz se adquirió su espectro TROSY.
- b. Se comparó el espectro $^1\text{H}^{15}\text{N}$ -HSQC de AdE1A.WT a 700 μM previamente adquirido, con el nuevo espectro TROSY. Al corroborar una excelente superposición de los picos, se transfirieron las asignaciones hechas sobre el $^1\text{H}^{15}\text{N}$ -HSQC al espectro TROSY.
- c. Luego, una muestra concentrada y sin marcar de RbAB en idéntico tampón de trabajo, fue mezclado cuidadosamente con una segunda muestra de AdE1A.WT marcada uniformemente con ^{15}N y que como la usada para adquirir el primer TROSY, también estaba a concentración 525 μM . La mezcla se hizo en una relación molar 1:1. El producto final diluido fue concentrado nuevamente a 525 μM de complejo 1:1 en 600 μl . Aquí es importante decir que no se observó agregación alguna durante todo este proceso incluso trabajando a tan alta concentración.
- d. Se adquirió el espectro TROSY del complejo [RbAB:AdE1A.WT]. Esto también es importante porque se obtuvo un espectro TROSY de buena calidad e interpretable.
- e. Se superpusieron los espectros TROSY de la proteína libre y en complejo con RbAB.

Los residuos interactuantes fueron determinados (en conjunto con los datos de los experimentos cuantitativos/termodinámicos) a partir de los valores obtenidos de la relación:

$$\frac{I}{I_0} \quad (42)$$

donde I es la intensidad del pico en el estado unido e I_0 es la intensidad del pico en el estado libre o basal. Las asignaciones de los picos de las proteínas AdE1A en sus formas acomplejadas con RbAB no se realizaron, sino que se utilizaron las asignaciones obtenidas del análisis sobre las formas libres en solución. Debido a esto, aquellos picos que desplazaron mucho hacia otras zonas del espectro post-interacción con RbAB, no pudieron identificarse a qué residuos correspondían.

Experimento de titulación

El experimento de titulación implicó la adquisición de una serie de 14 espectros ^1H - ^{15}N TROSY trabajando con 2 muestras uniformemente marcadas con ^{15}N de AdE1A.WT libre y otra con AdE1A.WT en complejo 1:1 con RbAB, ambas a 525 μM . Los primeros espectros (espectro inicial y final de la titulación) se obtuvieron de estas muestras sin mezclar. Cada

espectro del medio fue adquirido después de hacer simultáneamente mezclas cruzadas de cantidades crecientes entre ambas muestras (Ver Tabla MM4). Para cada punto se extrajo igual volumen de ambas muestras y dicho volumen se mezcló con el resto de la muestra complementaria, manteniendo siempre un volumen final de 600 μ l en cada tubo, asegurando de que al final de cada mezcla siempre haya 525 μ M AdE1A.WT marcada uniformemente con 15 N. Este procedimiento permitió obtener no sólo un *continuum* de 14 puntos (picos) cubriendo desde una relación molar 0 a 1, sino también el régimen de intercambio en el que se encontraban cada uno de los residuos de AdE1A.WT en complejo con RbAB.

Tabla MM4. Puntos del experimento de Titulación de AdE1A.WT

Mezcla cruzada	Volumen extraído para mezclar (μ l)	Concentración de AdE1A libre (μ M)	Concentración del complejo Rb:AdE1A (μ M)	Punto de titulación	% de AdE1A unido a pRb tras la mezcla	Concentración de pRb tras la mezcla (μ M)
0	0	525.0	0.0	1	0.0	0.0
1	25	503.1	21.9	2	4.2	21.9
2	25	483.1	41.9	3	8.0	41.9
3	50	446.3	78.7	4	15.0	78.7
4	50	415.7	109.3	5	20.8	109.3
5	100	364.6	160.4	6	30.5	160.4
6	100	330.6	194.4	7	37.0	194.4
6	100	194.4	330.6	8	63.0	330.6
5	100	160.4	364.6	9	69.5	364.6
4	50	109.3	415.7	10	79.2	415.7
3	50	78.7	446.3	11	85.0	446.3
2	25	41.9	483.1	12	92.0	483.1
1	25	21.9	503.1	13	95.8	503.1
0	0	0.0	525.0	14	100.0	525.0

Herramientas computacionales y programas utilizados

Simulaciones de las interacciones entre AdE1A.WT y mutantes con RbAB:

Los complejos de AdE1A.WT, AdE1A- Δ E y AdE1A- Δ L unidas a RbAB se modelaron utilizando una versión extendida de un método recientemente propuesto para generar conjuntos conformacionales realistas para IDPs [263]. Este método utiliza información estructural local dependiente de la secuencia, codificada en una extensa base de datos de fragmentos de tres residuos. El mismo, crea conformaciones de forma incremental, residuo por residuo, muestreando valores de ángulos diédricos de la base de datos, evitando choques estéricos. Para modelar el complejo [RbAB:AdE1A.WT], los dos motivos de unión: E2F (40-

PPTLHELYDL-49) y LxCxE (121-DLTCHEAGF-129) en los extremos N- y C-terminal respectivamente, se consideraron estáticos, conservando las conformaciones extraídas de estructuras determinadas experimentalmente (PDBID: 2R7G y 1GUX, respectivamente). El “*linker*” entre estos dos motivos, fue considerado como una cadena polipeptídica larga que adapta su conformación manteniendo los dos motivos fijos a RbAB rígidamente. El muestreo conformacional considerando dichas restricciones sobre el “*linker*” se realizó utilizando un algoritmo inspirado en robótica [264], que se adaptó para utilizar valores de ángulo diédrico de la base de datos mencionada anteriormente. Para cada conformación factible del “*linker*”, se muestrearon conformaciones geoméricamente compatibles de las extremos cortos N- y C-terminales (primeros y últimos aminoácidos de la cadena) usando la estrategia básica explicada en [263]. Para cada ensamble se simularon unas 10000 estructuras.

Cálculo de radios hidrodinámicos para los modelos simulados de interacción entre AdE1A.WT y mutantes con RbAB:

Se utilizó el programa HYDROPRO (versión 10) [265, 266] para calcular los radios hidrodinámicos (R_H) de los 3 complejos simulados. HYDROPRO se ejecutó en 1000 modelos seleccionados por EOM para el complejo [RbAB:AdE1A.WT] y 1000 conformaciones seleccionadas al azar de los complejos entre RbAB y AdE1A mutantes. Los cálculos se realizaron a temperaturas de 20 y 25 °C con las correspondientes viscosidades del disolvente de 0.01 y 0.009 poises, respectivamente. Los valores de radio del elemento atómico (AER), peso molecular, volumen específico parcial y densidad del disolvente se ajustaron a 2.9 Å, 54590 Da, 0,702 cm³/g y 1 g/cm³ para ambas temperaturas.

Dispersión de Rayos X de ángulo pequeño (SAXS)

La serie de experimentos SAXS del complejo entre RbAB y AdE1A presentadas en esta tesis, se hicieron en colaboración multilateral. Mientras que las muestras fueron generadas y ensambladas por este laboratorio, la recolección de los datos SAXS fueron realizados en las instalaciones del Laboratorio Europeo de Biología Molecular (Hamburgo, Alemania) y el modelado y análisis de los datos con el principal aporte del grupo del Dr. Pau Bernadó del *Centre de Biochimie Structurale* de la *Université de Montpellier* y del Dr. Juan Cortes, del LAAS-CNRS de la *Université de Toulouse* (Francia).

Brevemente, SAXS es una técnica en la que una solución de partículas es irradiada por un haz de Rayos X y la intensidad de dispersión de estos rayos son captados por un detector. El patrón de dispersión de la partícula de interés al que se le sustrae el patrón de dispersión del

solvente, se relaciona con la forma y tamaño de la partícula. Como estas partículas se orientan de modo azaroso en la solución, el patrón de dispersión es isotrópico y por tanto se colectan en dos dimensiones y se promedian en forma radial. Estos patrones isotrópicos son representados como curvas de $I(s)$ (intensidad de señal) unidimensionales radialmente promediadas [267].

En este caso, a partir de muestras “stocks” de AdE1A.WT y RbAB en altas concentraciones y tampón 20 mM fosfato de sodio, 200 mM NaCl, 1mM DTT pH 7.0, se preparó una serie de muestras a ser analizadas. En los experimentos de SAXS la intensidad de la dispersión es proporcional a la concentración de la muestra, a altas concentraciones la relación señal/ruido aumenta aunque también hay contribución indeseada a la señal por un efecto atractivo inter-partículas, que hace que asuman distancias en el mismo orden de magnitud que las distancias intra-partículas. Esto último se soluciona midiendo a concentraciones menores y confluyendo luego las señales para obtener una única curva de dispersión [267]. Por tanto se requirieron medidas a diferentes concentraciones. La tabla con las muestras preparadas se presenta a continuación:

Tabla MM5. Muestras a diferentes concentraciones utilizadas en experimentos SAXS

	Contenido	Concentración		Volumen (μ l)
		(mg/ml)	(μ M)	
1	RbAB	4	95	100
2	RbAB	2	47.5	100
3	RbAB	1	23.7	200
4	[RbAB:AdE1A.WT]	5.46	100:100	100
5	[RbAB:AdE1A.WT]	2.73	50:50	200
6	[RbAB:AdE1A.WT]	1.37	25:25	200

Una vez preparadas las muestras, estas fueron transportadas al Laboratorio Europeo de Biología Molecular en Hamburgo en donde los experimentos SAXS se realizaron utilizando las líneas P12 de los anillos DORIS y PETRAIII de dichas instalaciones. Las muestras fueron dispuestas en tubos capilares de cuarzo (100 μ l), los cuales fueron irradiados por un haz colimado monocromático de Rayos X a 1.24 Å y la intensidad de la dispersión de los rayos captados por un detector RX ubicado a 3.0 m de la muestra. Los perfiles medidos cubrieron un rango de transferencia de momento de $0.0026 < s < 0.73 \text{ \AA}^{-1}$. Los datos SAXS fueron medidos para RbAB como para el complejo [RbAB:AdE1A.WT] a 10 °C. También fue medido el tampón 20 mM fosfato de sodio, 200 mM NaCl, 1 mM DTT a pH 7.0. Los patrones de dispersión

obtenidos de la solución tampón fueron medidos antes y después de las medidas de las muestras y para ambos (tampón y muestras) las medidas se realizaron múltiples veces con el fin de detectar y corregir los datos por el efecto del daño provocado por la radiación.

Las curvas finales de dispersión de señal obtenidas a cada concentración de muestras fueron derivadas luego de restarles los patrones de dispersión promedio del tampón. Ninguna muestra mostró agregación ni hubo señal de tal efecto en las curvas obtenidas y analizadas. Los perfiles finales de SAXS para ambas muestras analizadas fueron obtenidas luego de confluir las curvas obtenidas a cada concentración para corregir los efectos de las atracciones inter-partículas (mencionado anteriormente). Los datos crudos fueron tratados luego con protocolos estándar utilizando la serie de programas ATSAS [268].

La intensidad de dispersión $I(0)$ y el radio de giro R_g , fueron evaluados utilizando la aproximación de Guinier –análisis lineal- [269], asumiendo que a ángulos muy pequeños ($s < 1.3/R_g$) la intensidad puede ser bien representada como:

$$I(s) = I(0) \times e^{-(sR_g)^2/3} \quad (43)$$

La distribución de probabilidades $p(r)$, luego fue calculada por Transformada de Fourier indirecta utilizando el programa GNOM –análisis no lineal- [270], aplicando rangos de transferencia de momentos de $0.01 < s < 0.33 \text{ \AA}^{-1}$ y $0.013 < s < 0.27 \text{ \AA}^{-1}$ para RbAB y el complejo, respectivamente.

Para el caso de RbAB, la estructura cristalográfica de RbAB (PDBID: 3POM), fue ajustada a los curvas experimentales de SAXS, utilizando el programa FOXS [271, 272]. Para este caso se obtuvo un ajuste óptimo ($\chi^2=0.86$) luego de que se modelaron los segmentos perdidos en la estructura cristalográfica (regiones enlazantes y segmentos N- y C-terminal) con el programa AllosMod-FoXS [273]. El caso del complejo [RbAB:AdE1A.WT] es distinto porque no existe estructura de alta resolución disponible (cristalográfica o RMN) por lo que los datos SAXS para esta muestra fue analizado con el Método de Optimización de Ensamblajes (EOM, por sus siglas en inglés) [235, 274]. Este método selecciona un sub-ensamble de conformaciones que juntos describen mejor los perfiles de datos experimentales de SAXS a partir de miles de modelos de conformaciones posibles para una proteína (o complejo). En este caso, fueron modeladas y generadas 10000 estructuras posibles del complejo [RbAB:AdE1A.WT] con ambos motivos unidos a RbAB de manera estática y el “*linker*” con conformación flexible y otras 10000 estructuras en las que sólo un motivo (independientemente de cual) se encontraba fijo a RbAB. Todos estos modelados poseían perfiles de SAXS teóricos

asociados y luego fueron computados por CRY SOL [275]. De estas posibles estructuras, fueron seleccionadas con EOM 200 sub-ensambles diferentes de 20 o 50 conformaciones que pueden describir colectivamente los datos experimentales de SAXS y fueron analizados en términos de distribuciones de R_g . Si bien un ensamble de 20-50 estructuras es muy pequeño para explicar el comportamiento de un complejo en el que está asociado una proteína IDP, esta subpoblación es usualmente suficiente para describir las propiedades globales del sistema en términos de tamaño y forma [267].

Modelado WLC o de cadenas entrópicas

Para describir el comportamiento de la región enlazante “*linker*” de AdE1A se utilizó el modelo de polímero flexible o cadena entrópica (o *Worm Like Chain*, WLC por sus siglas en inglés) [111] que considera a una cadena polipeptídica como un cilindro continuo con cambios azarosos en la dirección de sus movimientos en solución pero con un radio de curvatura constante. Para este modelo la distribución de densidad de probabilidad de las distancias *fin-a-fin* sólo depende de dos factores: una longitud de persistencia (l_p), que es la distancia que le toma a los movimientos para perder todo tipo de correlación y la longitud de contorno (l_c), que no es otra cosa que la longitud total de la cadena polipeptídica. Para péptidos largos, l_p asume un valor de 3 Å [239] y l_c es el producto del número total de residuos por la contribución promedio de un aminoácido (3.8 Å). La distribución puede ser calculada como:

$$\rho(r) = 4\pi r^2 \left(\frac{3}{4\pi l_p l_c} \right)^{\frac{3}{2}} \exp\left(\frac{-3r^2}{4l_p l_c} \right) \zeta(r, l_p, l_c) \quad (44)$$

donde el término final es definido como:

$$\zeta(r, l_p, l_c) = \left(1 - \frac{5l_p}{4l_c} + \frac{2r^2}{l_c^2} - \frac{33r^2}{80l_p l_c^3} - \frac{69l_p^2}{160l_c^2} - \frac{329r^2 l_p}{120l_c^3} + \frac{6799r^4}{1600l_c^4} - \frac{3441r^6}{2800l_p l_c^5} + \frac{1089r^8}{12800l_p^2 l_c^6} \right) \quad (45)$$

La densidad de probabilidad de la distancia *fin-a-fin* cuando el *linker* es fijado a una distancia dada (r_0) (en este caso obtenida de la estructura de rayos X del complejo de ambos motivos con RbAB) entre los sitios de interacción, puede ser relacionada con un valor de concentración efectiva en el estado unido y que puede ser expresado así:

$$C_{eff} = \frac{\rho(r_0) 10^{27} \text{Å}^3 l^{-1}}{4\pi r^2 L_0} \quad (46)$$

donde L_0 es el número de Avogadro, (r_0) es la distancia que separa ambos sitios de interacción y C_{eff} es la concentración efectiva expresada en Molar [57, 239, 240].

ANEXO

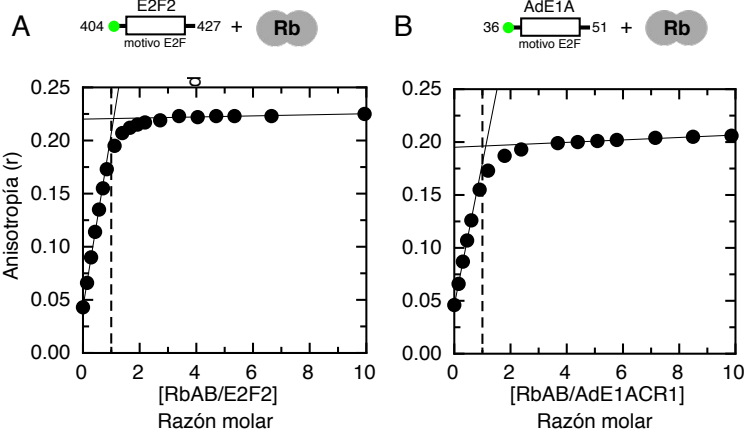
Tabla A2.1. Estequiometrías y Constantes de disociación de las interacciones entre péptidos y proteínas AdE1A y E2F2humana con RbAB

Sonda	n	(nM)	K_D (nM)	K_D (nM)	$r(0)$	$r(F)$	$F(0)$	$F(F)$
		Datos crudos		Datos normalizados	Señal de anisotropía		Señal de fluorescencia	
E2F2	1	1	1.18 ± 0.16	1.15 ± 0.05	0.062	0.163		
		5	0.91 ± 0.04	0.91 ± 0.02	0.049	0.209		
		10	0.71 ± 0.05	0.74 ± 0.03	0.043	0.219		
		20	1.39 ± 0.16	1.55 ± 0.05	0.043	0.215		
		30	1.06 ± 0.08	1.14 ± 0.05	0.043	0.227		
		Global		1.02 ± 0.03				
AdE1ACR1	1	100	109.3 ± 9.7	97.6 ± 3.6	0.046	0.198		
		200	100.2 ± 22.4	97.4 ± 5.6	0.049	0.206		
		300	101.3 ± 11.4	93.2 ± 4.2	0.049	0.206		
		500	105.5 ± 10.2	93.4 ± 4.4	0.050	0.211		
		1000	127.9 ± 12.2	75.1 ± 7.9	0.045	0.203		
		2000	109.5 ± 0.6	81.6 ± 4.9	0.046	0.206		
		Global		119.2 ± 2.1				
AdE1ACR2	1	100	113.4 ± 14.1	117.2 ± 6.5	0.028	0.076		
		200	93.1 ± 11.3	93.5 ± 2.9	0.036	0.091		
		300	93.7 ± 7.9	93.0 ± 2.6	0.033	0.087		
		500	97.9 ± 15.6	98.7 ± 3.7	0.032	0.088		
		1000	117.1 ± 28.3	111.9 ± 7.4	0.032	0.086		
		Global		104.6 ± 2.9				
AdE1ACR2 _{AC}	1	130	80.7 ± 12.8	80.5 ± 6.9	0.04	0.100		
		175	73.7 ± 5.2	73.7 ± 5.2	0.04	0.105		
		700	85.6 ± 7.4	82.8 ± 5.8	0.039	0.106		
		Global		72.4 ± 2.2				
AdE1ACR2 _{ACP}	1	30	20.1 ± 2.6	20.1 ± 2.6	0.05	0.099		
		50	20.2 ± 2.4	17.5 ± 1.3	0.04	0.093		
		100	19.5 ± 2.2	18.2 ± 2.2	0.039	0.092		
		Global		19.6 ± 0.5				
AdE1A.WT	1	0.5	0.025 ± 0.004	0.025 ± 0.003	0.075	0.191		
		0.5	0.026 ± 0.009	0.025 ± 0.008			111.6	94.3
		0.5	0.027 ± 0.006	0.027 ± 0.003	0.059	0.127		
		1	0.037 ± 0.009	0.034 ± 0.005	0.061	0.153		
		1	0.049 ± 0.019	0.050 ± 0.012			302.8	259.8
		2	0.046 ± 0.007	0.047 ± 0.005	0.065	0.197		
		2	0.065 ± 0.022	0.092 ± 0.025			451.3	374.3
		2	0.062 ± 0.008	0.062 ± 0.005	0.064	0.2		
		2	0.074 ± 0.036	0.090 ± 0.034			412.4	336.6
		2	0.065 ± 0.015	0.062 ± 0.009	0.062	0.175		
2	0.075 ± 0.010	0.074 ± 0.006			613.5	503.4		
Global		0.024 ± 0.004						
AdE1A-ΔE	1	50	42.6 ± 3.9	42.5 ± 3.6	0.05	0.105		
		200	53.9 ± 3.1	53.9 ± 3.1	0.05	0.106		
		800	60.2 ± 10.1	54.5 ± 10.3	0.048	0.105		
		Global		47.61 ± 1.7				
AdE1A-ΔL	1	200	197.7 ± 4.5	197.7 ± 4.5	0.058	0.220		
		400	207.5 ± 7.4	207.5 ± 7.4	0.056	0.232		
		800	225.2 ± 20.6	217.9 ± 10.9	0.023	0.219		
		Global		206.7 ± 3.1				

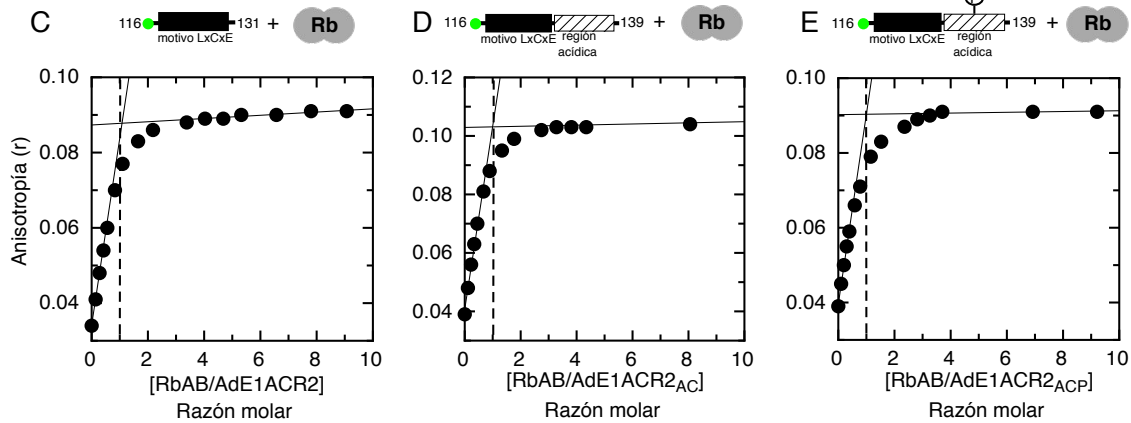
En la Tabla se muestran todas las mediciones realizadas para las interacciones ensayadas. Las condiciones de ensayo fueron tampón fosfato de sodio 20 mM, 200 mM NaCl, 2 mM DTT y 0.1% Tween-20, pH 7.0 y a 20 °C. Para todos los casos se presenta: la estequiometría hallada, las distintas concentraciones de especie marcadas con FITC a las que se realizaron las curvas de titulación, los valores de K_D hallados del ajuste de los datos crudos y normalizados entre 0 y 1 – que demuestran que la normalización no modifica sustancialmente los resultados y hay buena concordancia entre ellos-, el K_D del ajuste global -que es finalmente informado-. Se listan además los valores iniciales de anisotropía $r(0)$, que evidencian la anisotropía intrínseca de las especies derivatizadas las que se mantienen relativamente constantes e independientes de las concentraciones ensayadas y los valores de anisotropía final $r(F)$ que alcanzan valores estables para los complejos con RbAB. También se presentan los valores de fluorescencia iniciales y finales ($F(0)$ y $F(F)$) que sólo se midieron para AdE1A.WT, aunque representan rangos mas variables pero igualmente utilizables para hacer los correspondientes ajustes de K_D .

Figura A2.1. Medidas espectroscópicas de fluorescencia para determinar estequiometrías de interacciones con RbAB

Motivo E2F en E2F2 humano y AdE1A



Motivo LxCxE en AdE1A



AdE1A.WT y mutantes

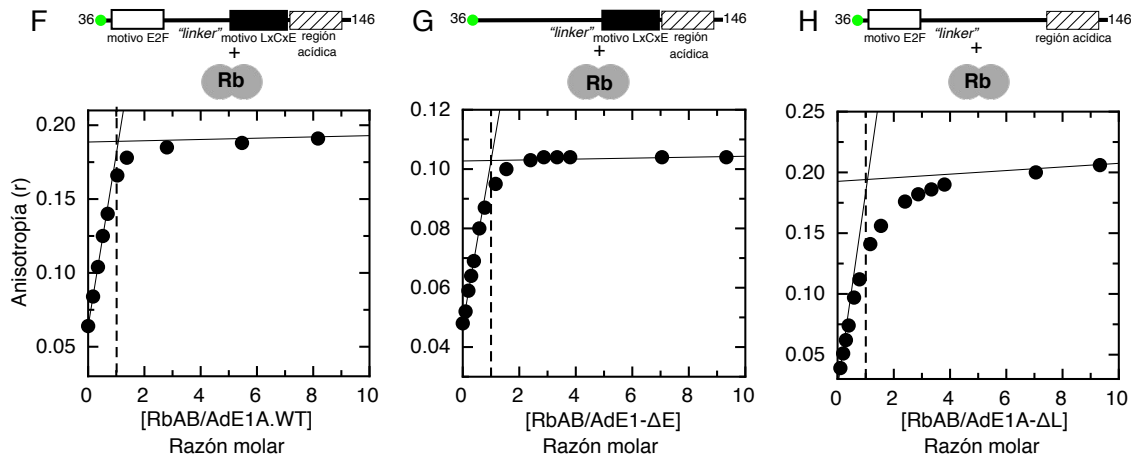
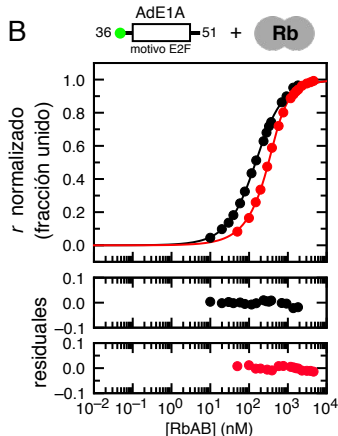
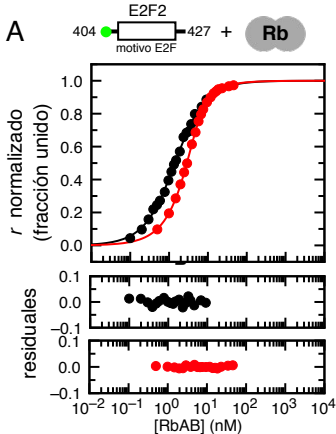


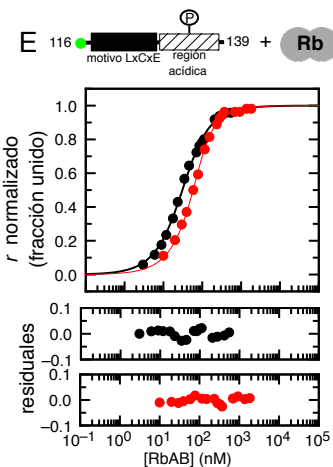
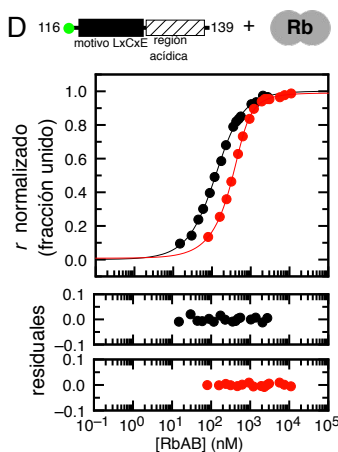
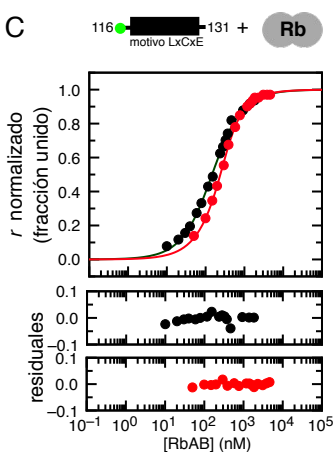
Figura A2.1. Medidas espectroscópicas de fluorescencia para determinar la estequiometría de interacción con RbAB. Las curvas de interacción en condiciones saturantes fueron obtenidas en tampón 20 mM fosfato de sodio, 200mM NaCl, 1 mM DTT, 0.1% Tween 20, pH 7.0 a 20°C. Las estequiometrías fueron determinadas de la extrapolación de dos ajustes lineales (línea rectas negras) al final y al inicio de las señales de anisotropía. La línea punteada indica razones molares 1:1 de titulante, evidenciando estequiometrías n=1. Medidas fueron hechas en las siguientes concentraciones de agente titulado (sonda) **A**) E2F2 30 nM, **B**) AdE1ACR1 2000 nM, **C**) AdE1ACR2 1000 nM, **D**) AdE1ACR2_{AC} 700 nM, **E**) AdE1ACR2_{ACP} 100 nM, **F**) AdE1A.WT 2 nM, **G**) AdE1A-ΔE 800 nM, **H**) AdE1A-ΔL 800 nM.

Figura A2.2. Medidas espectroscópicas representativas para la determinación de constantes de disociación de interacciones con RbAB

Motivo E2F en E2F2 humano y AdE1A



Motivo LxCxE en AdE1A



AdE1A.WT y mutantes

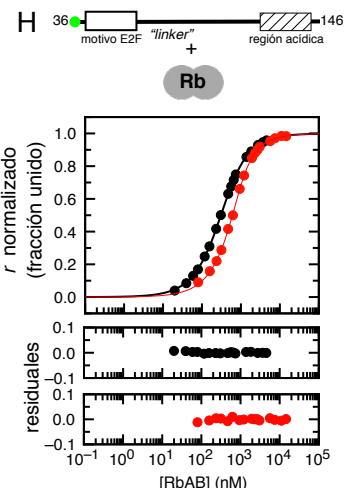
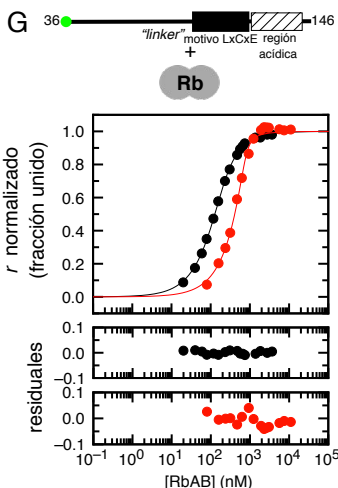
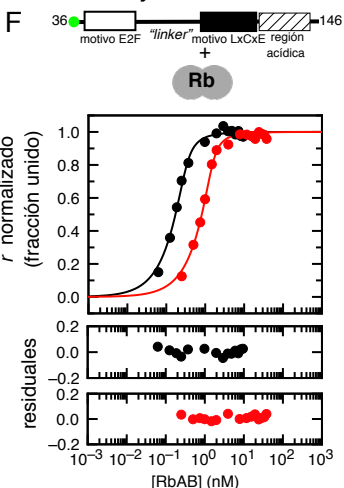


Figura A2.2: Medidas espectroscópicas representativas para la determinación de constantes de disociación de interacciones con RbAB. Curvas de titulación al equilibrio ajustado a un modelo de interacción 1:1. Intensidades de anisotropía de fluorescencia normalizada (entre 0 y 1) a dos concentraciones sub-estequiométricas distintas (círculos negros y rojos) ajustadas a curvas de unión 1:1 (líneas negras y rojas) para obtener K_D individuales y los residuales del ajuste. Estas curvas corresponden a una sonda a 2 concentraciones distintas para mostrar como el sistema se comporta de manera consistente, independientemente de las concentraciones. Las mismas muestran un buen ajuste a las distintas concentraciones ensayadas. A) E2F2 1 y 5 nM, B) AdE1ACR1 100 y 500 nM, C) AdE1ACR2 100 y 500 nM, D) AdE1ACR2_{AC} 130 y 700 nM, E) AdE1ACR2_{ACP} 30 y 100 nM, F) AdE1A.WT 0.5 y 2nM, G) AdE1A-ΔE 200 y 800 nM, H) AdE1A-ΔL 200 y 800 nM.

Figura A2.3. Medidas espectroscópicas para la interacción RbAB – E2F2humano

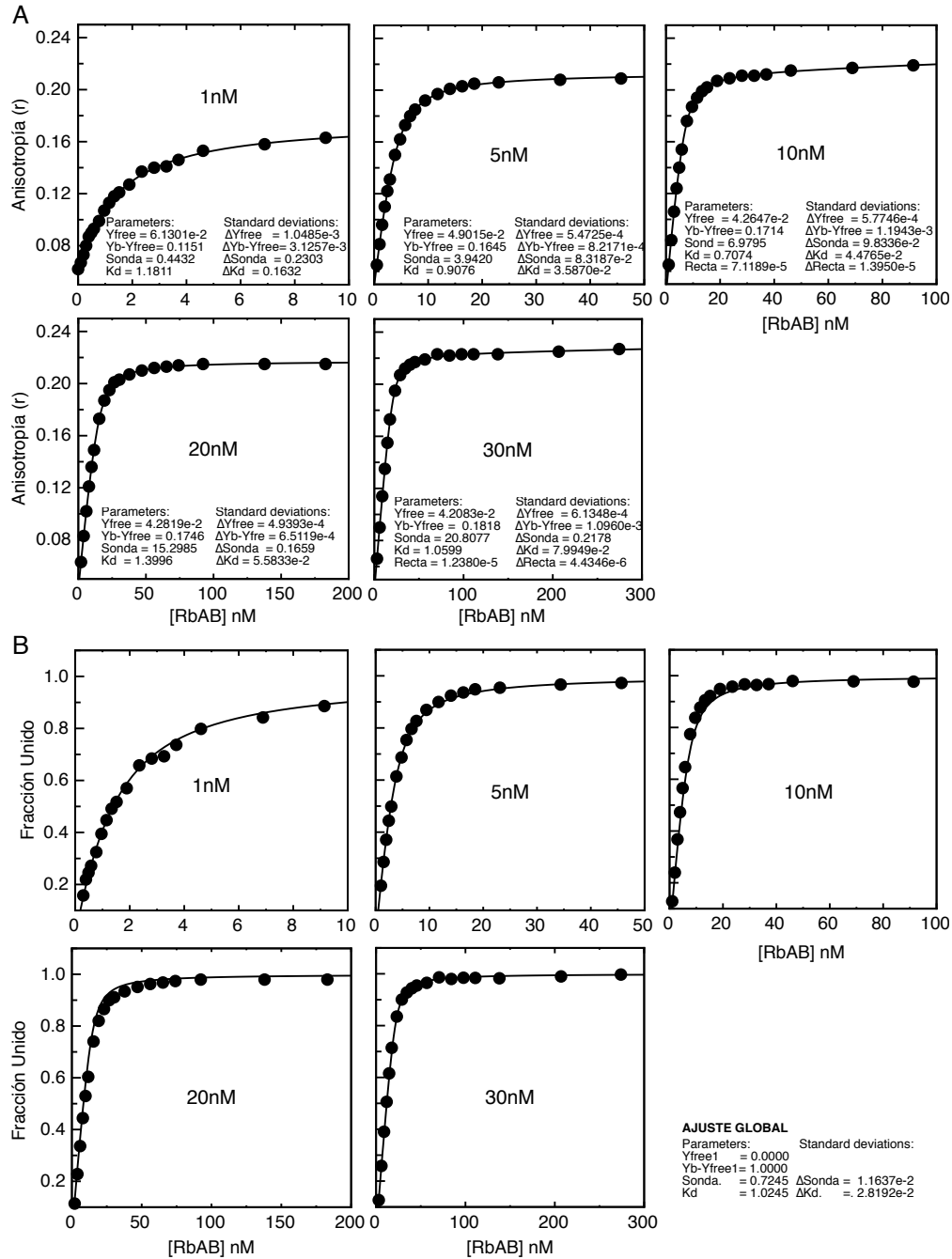


Figura A2.3: Medidas espectroscópicas para la interacción RbAB- E2F2humano. Curvas de titulación al equilibrio ajustado a un modelo de interacción 1:1. A) Señales de intensidad de anisotropía de fluorescencia (r) cruda a diferentes concentraciones de E2F2-FITC fueron ajustadas para obtener K_D individuales. Parámetros del ajuste están mostrados dentro de cada gráfico. B) Señales de intensidad de anisotropía de fluorescencia (r) normalizada entre 0 y 1, a diferentes concentraciones de E2F2-FITC que fueron usadas en conjunto para obtener K_D global. (Ver Tabla A2.1)

Figura A2.4. Medidas espectroscópicas para la interacción RbAB – AdE1ACR1

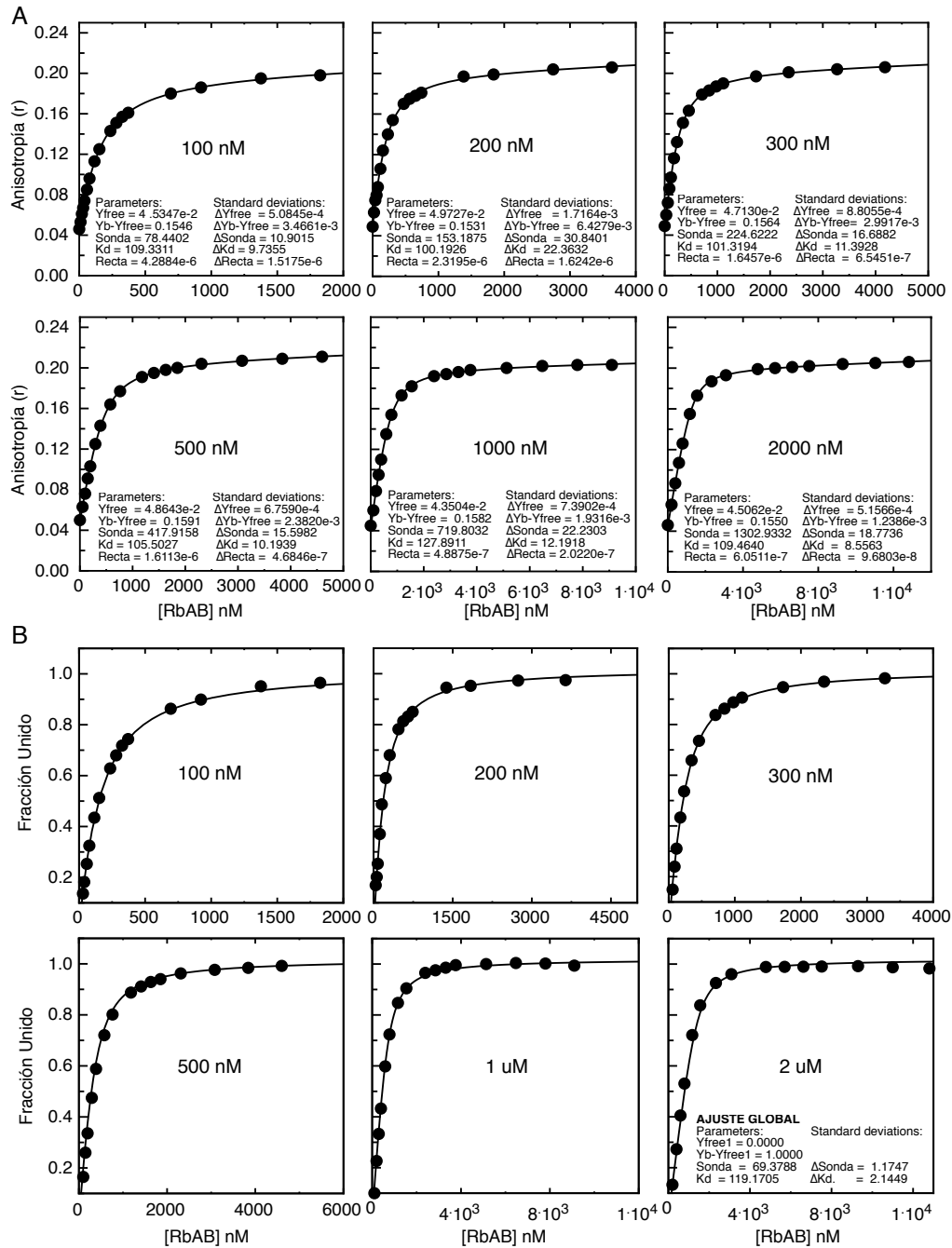


Figura A2.4: Medidas espectroscópicas para la interacción RbAB- AdE1ACR1. Curvas de titulación al equilibrio ajustado a un modelo de interacción 1:1. A) Señales de intensidad de anisotropía de fluorescencia (r) cruda a diferentes concentraciones de AdE1ACR1-FITC fueron ajustadas para obtener K_D individuales. Parámetros del ajuste están mostrados dentro de cada gráfico. B) Señales de intensidad de anisotropía de fluorescencia (r) normalizada entre 0 y 1, a diferentes concentraciones de AdE1ACR1-FITC que fueron usadas en conjunto para obtener K_D global. (Ver Tabla A2.1)

Figura A2.5. Medidas espectroscópicas para la interacción RbAB – AdE1ACR2

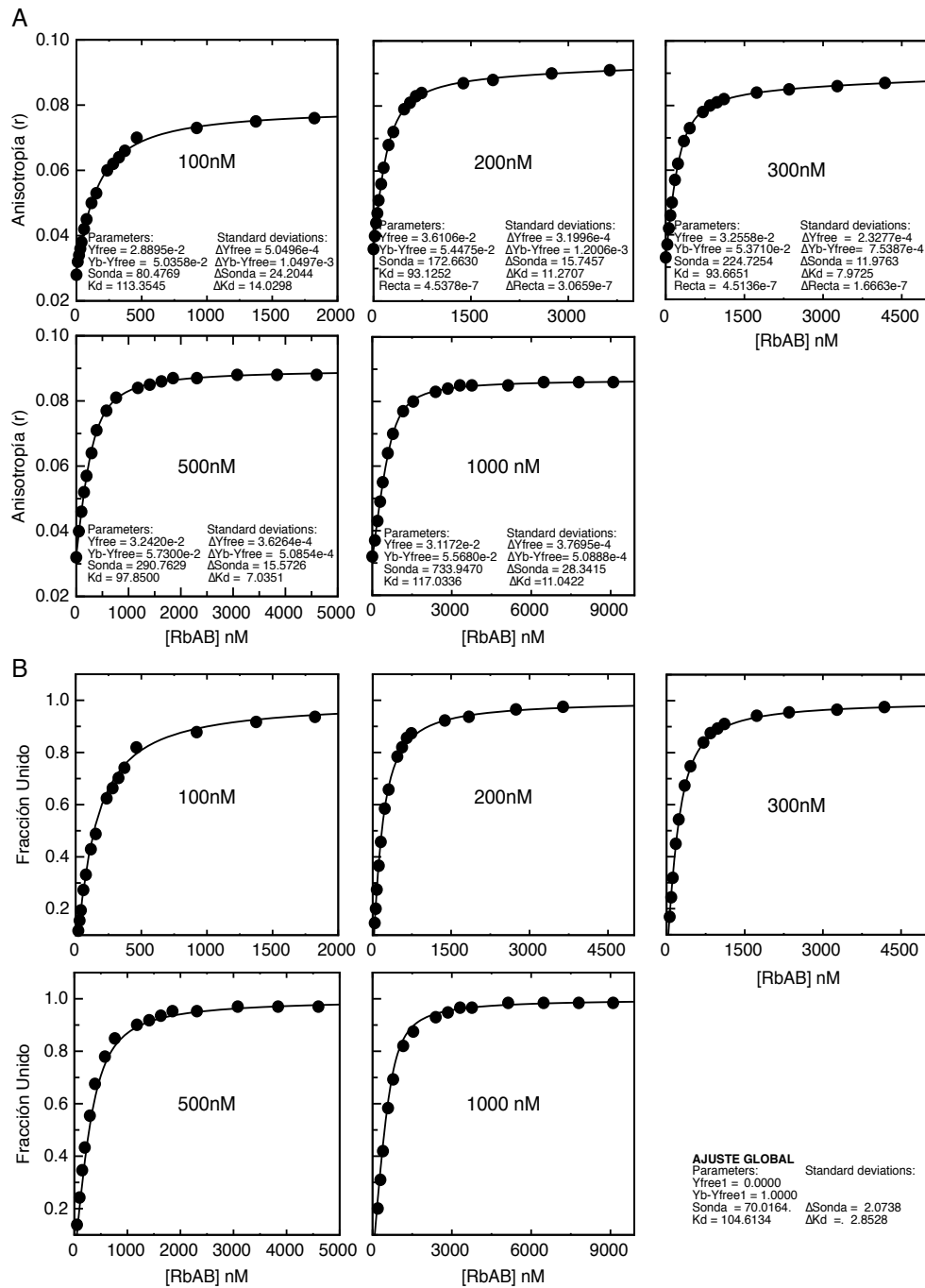


Figura A2.5: Medidas espectroscópicas para la interacción RbAB- AdE1ACR2. Curvas de titulación al equilibrio ajustado a un modelo de interacción 1:1. A) Señales de intensidad de anisotropía de fluorescencia (r) cruda a diferentes concentraciones de AdE1ACR2-FITC fueron ajustadas para obtener K_D individuales. Parámetros del ajuste están mostrados dentro de cada gráfico. B) Señales de intensidad de anisotropía de fluorescencia (r) normalizada entre 0 y 1, a diferentes concentraciones de AdE1ACR2-FITC que fueron usadas en conjunto para obtener K_D global. (Ver Tabla A2.1)

Figura A2.6. Medidas espectroscópicas para la interacciones de RbAB con AdE1ACR2_{AC} Y AdE1ACR2_{ACP}

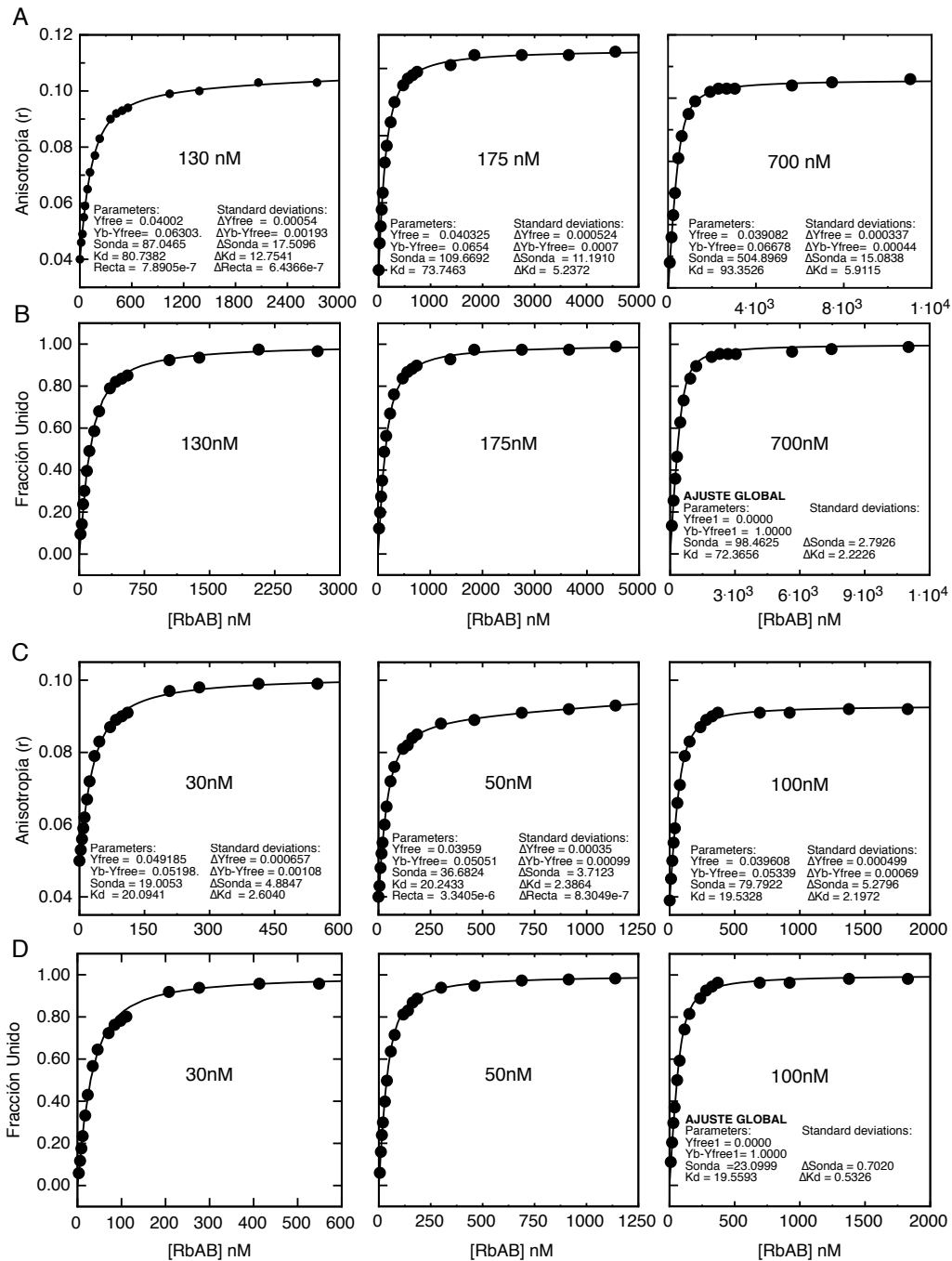


Figura A2.6: Medidas espectroscópicas para las interacciones de RbAB con AdE1ACR2_{AC} Y AdE1ACR2_{ACP}. Curvas de titulación al equilibrio ajustado a un modelo de interacción 1:1. Señales de intensidad de anisotropía de fluorescencia (r) cruda a diferentes concentraciones de AdE1ACR2_{AC}-FITC (A) y AdE1ACR2_{ACP}-FITC (C) fueron ajustadas para obtener K_D individuales. Parámetros del ajuste están mostrados dentro de cada gráfico. Señales de intensidad de anisotropía de fluorescencia (r) normalizada entre 0 y 1, a diferentes concentraciones de AdE1ACR2_{AC}-FITC (B) y AdE1ACR2_{ACP}-FITC (D) que fueron usadas en conjunto para obtener K_D global. (Ver Tabla A2.1)

Figura A2.7.A. Medidas espectroscópicas para la interacción RbAB – AdE1A.WT

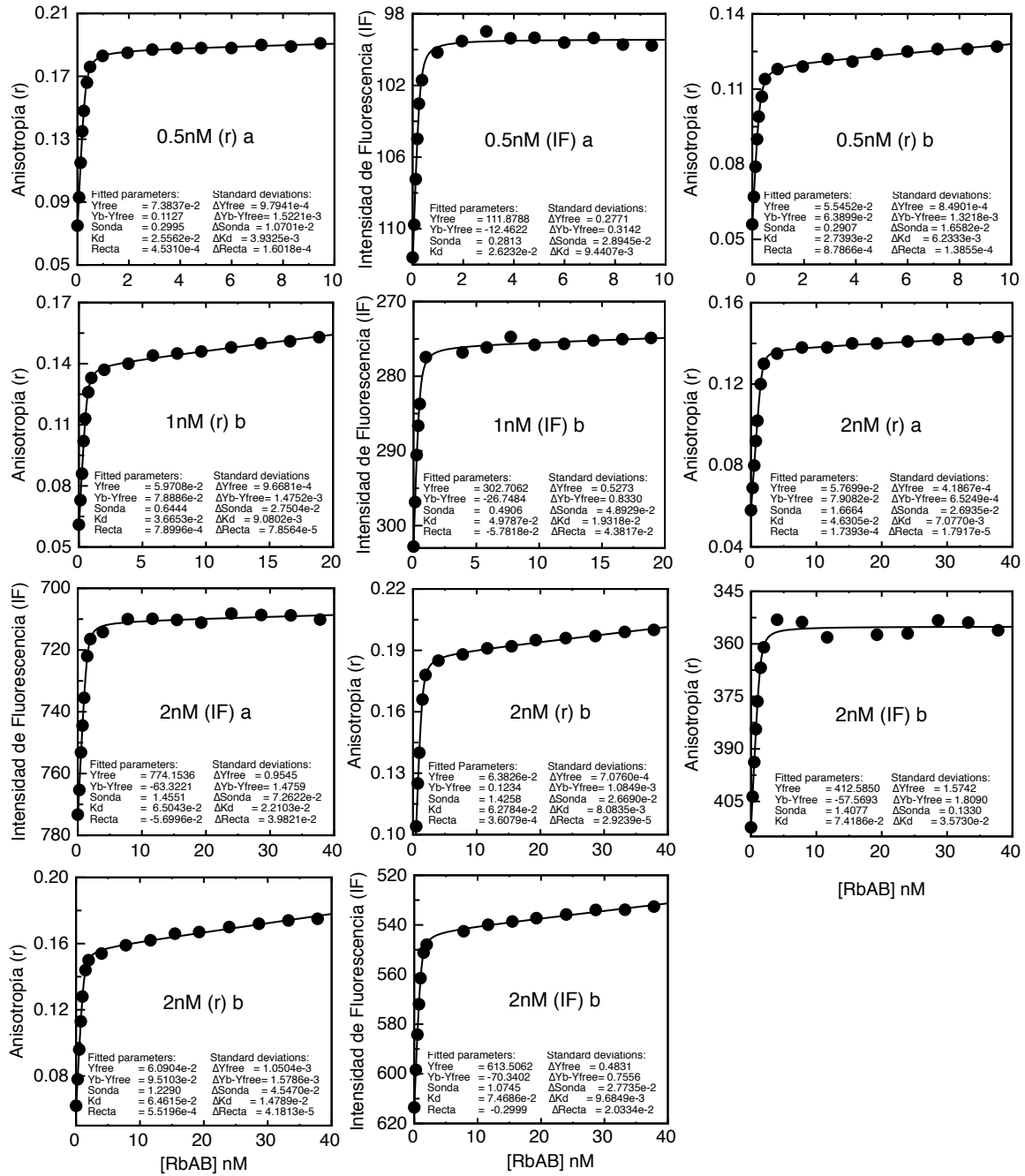


Figura A2.7.A.: Medidas espectroscópicas para la interacción RbAB- AdE1A.WT. Curvas de titulación al equilibrio ajustado a un modelo de interacción 1:1. Señales de intensidad de anisotropía de fluorescencia (r) cruda a diferentes concentraciones de AdE1A.WT-FITC fueron ajustadas para obtener K_D individuales. Parámetros del ajuste están mostrados dentro de cada gráfico.

Figura A2.7.B. Medidas espectroscópicas para la interacción RbAB – AdE1A.WT

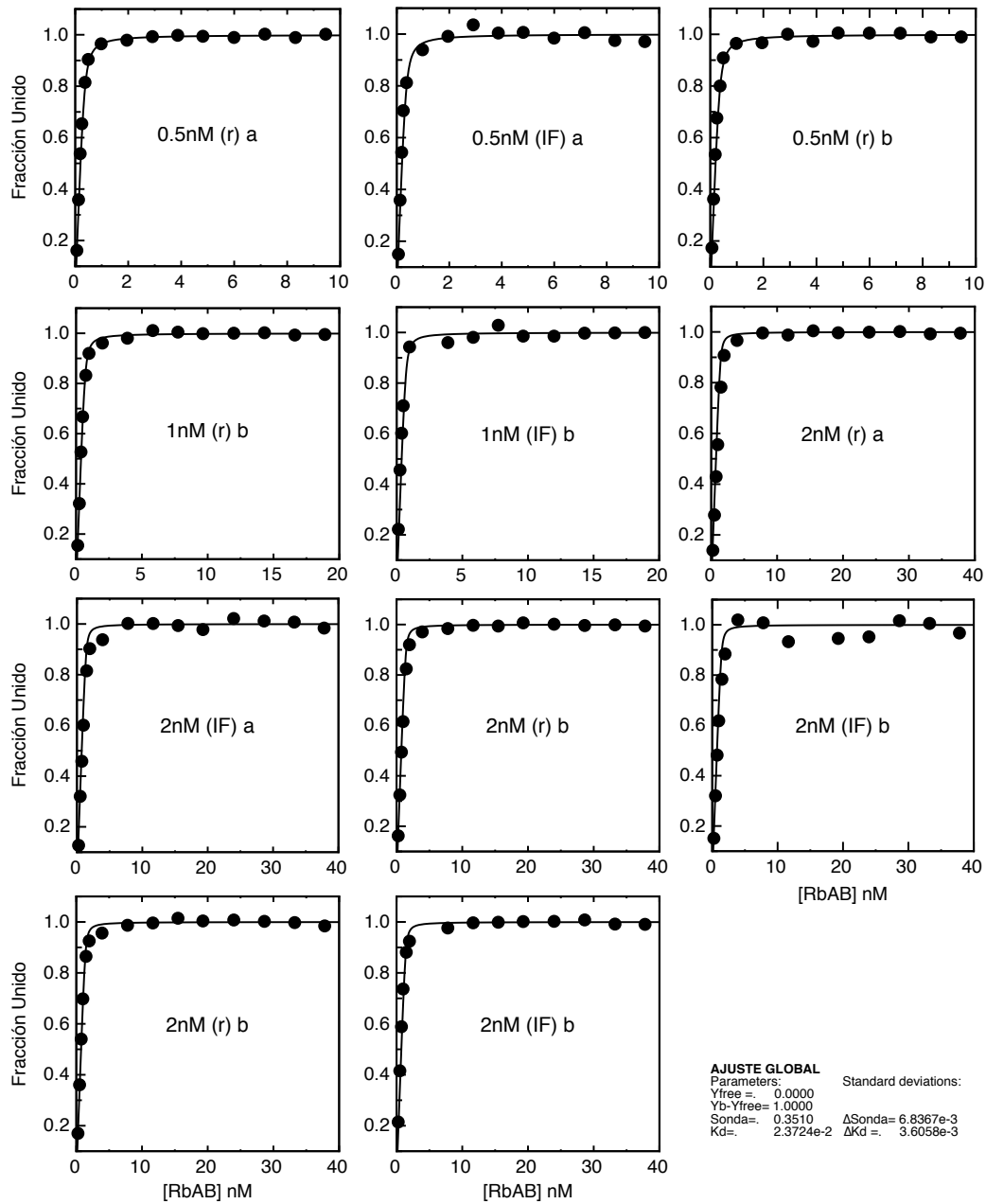


Figura A2.7.B.: Medidas espectroscópicas para la interacción RbAB- AdE1A.WT. Curvas de titulación al equilibrio ajustado a un modelo de interacción 1:1. Señales de intensidad de anisotropía de fluorescencia (*r*) normalizada entre 0 y 1, a diferentes concentraciones de AdE1A.WT-FITC que fueron usadas en conjunto para obtener K_D global. (Ver Tabla A2.1)

Figura A2.8. Medidas espectroscópicas para la interacciones de RbAB con AdE1A- Δ E y AdE1A- Δ L

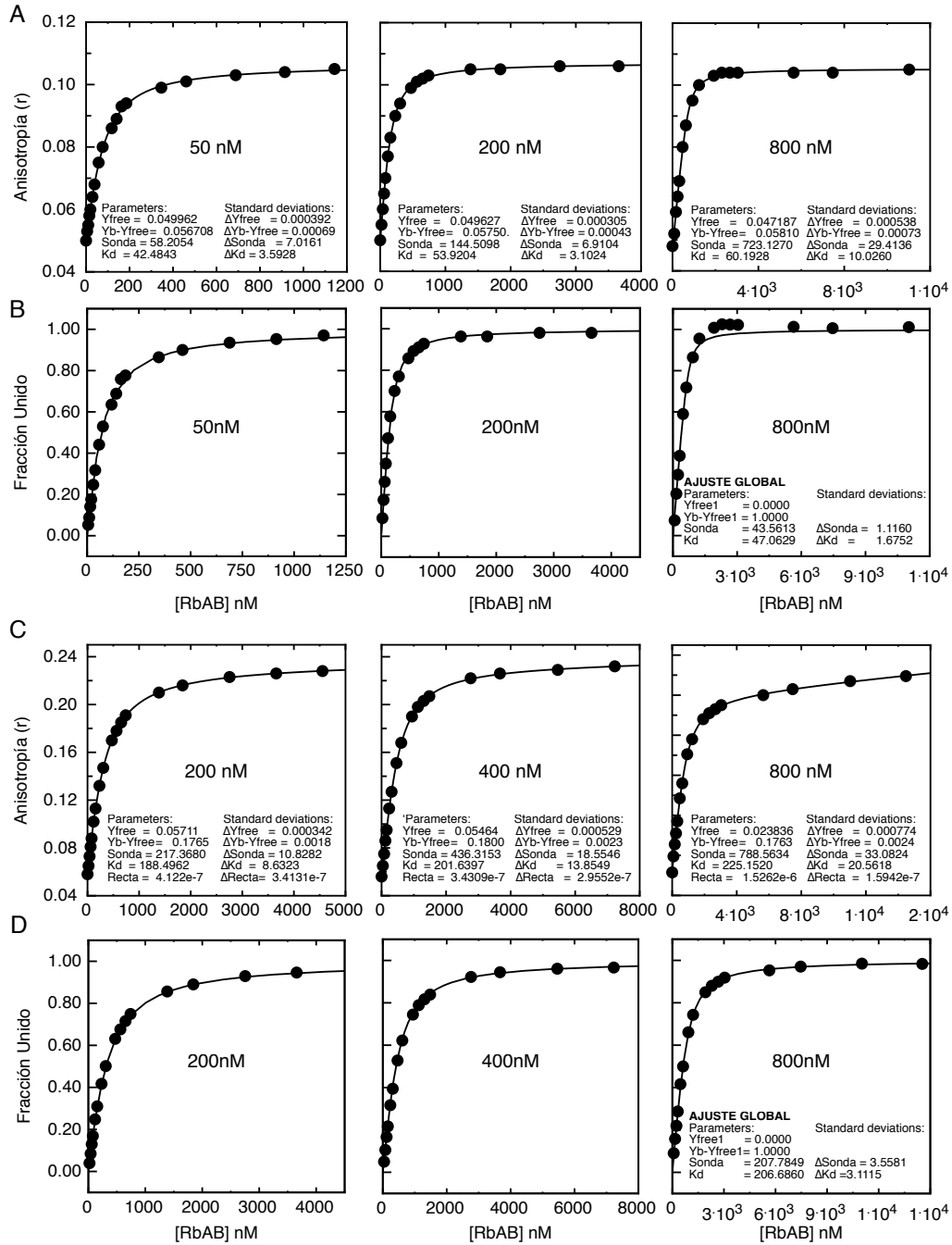


Figura A2.8: Medidas espectroscópicas para las interacciones de RbAB con AdE1A- Δ E y AdE1A- Δ L. Curvas de titulación al equilibrio ajustado a un modelo de interacción 1:1. Señales de intensidad de anisotropía de fluorescencia (r) cruda a diferentes concentraciones de AdE1A- Δ E -FITC (A) y AdE1A- Δ L FITC (C) fueron ajustadas para obtener K_D individuales. Parámetros del ajuste están mostrados dentro de cada gráfico. Señales de intensidad de anisotropía de fluorescencia (r) normalizada entre 0 y 1, a diferentes concentraciones de AdE1A- Δ E -FITC (B) y AdE1A- Δ L -FITC (D) que fueron usadas en conjunto para obtener K_D global. (Ver Tabla A2.1)

Figura A2.9. Medidas de ITC para la interacción RbAB – E2F2humano

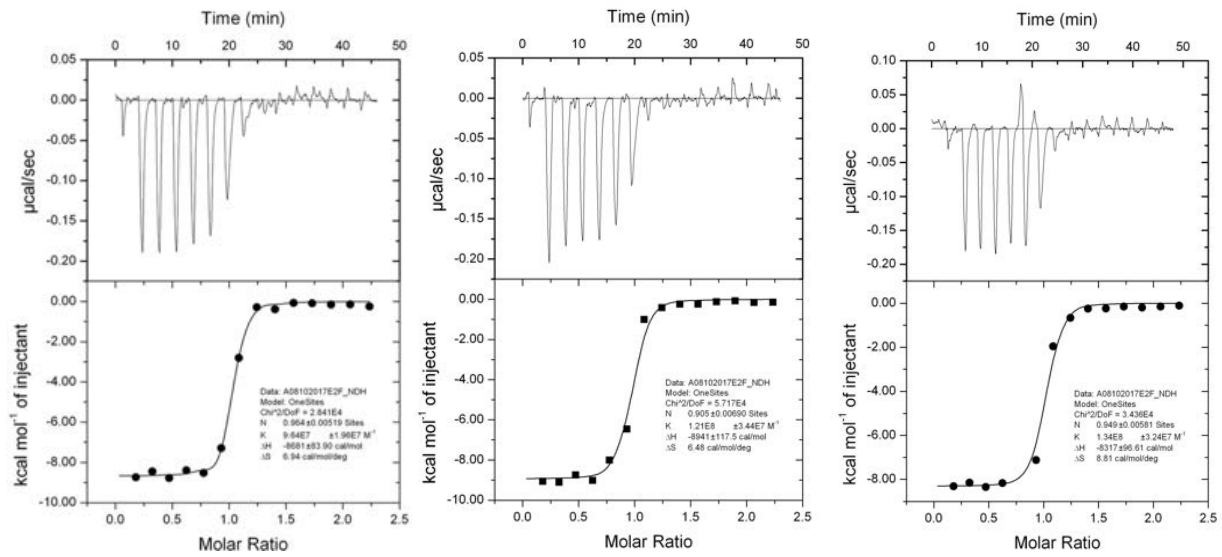


Figura A2.9: Medidas de ITC para la interacción RbAB – E2F2humano. Titulaciones calorimétricas hechas en fosfato de sodio 20mM, 200mM NaCl, 5 mM 2-mercapto etanol a pH 7.0. La celda fue cargada con RbAB aislado ($5\mu\text{M}$) y la jeringa con E2F2humano ($50\mu\text{M}$). Titulaciones hechas a 20°C .

Figura A2.10. Medidas de ITC para la interacción RbAB – AdE1ACR1

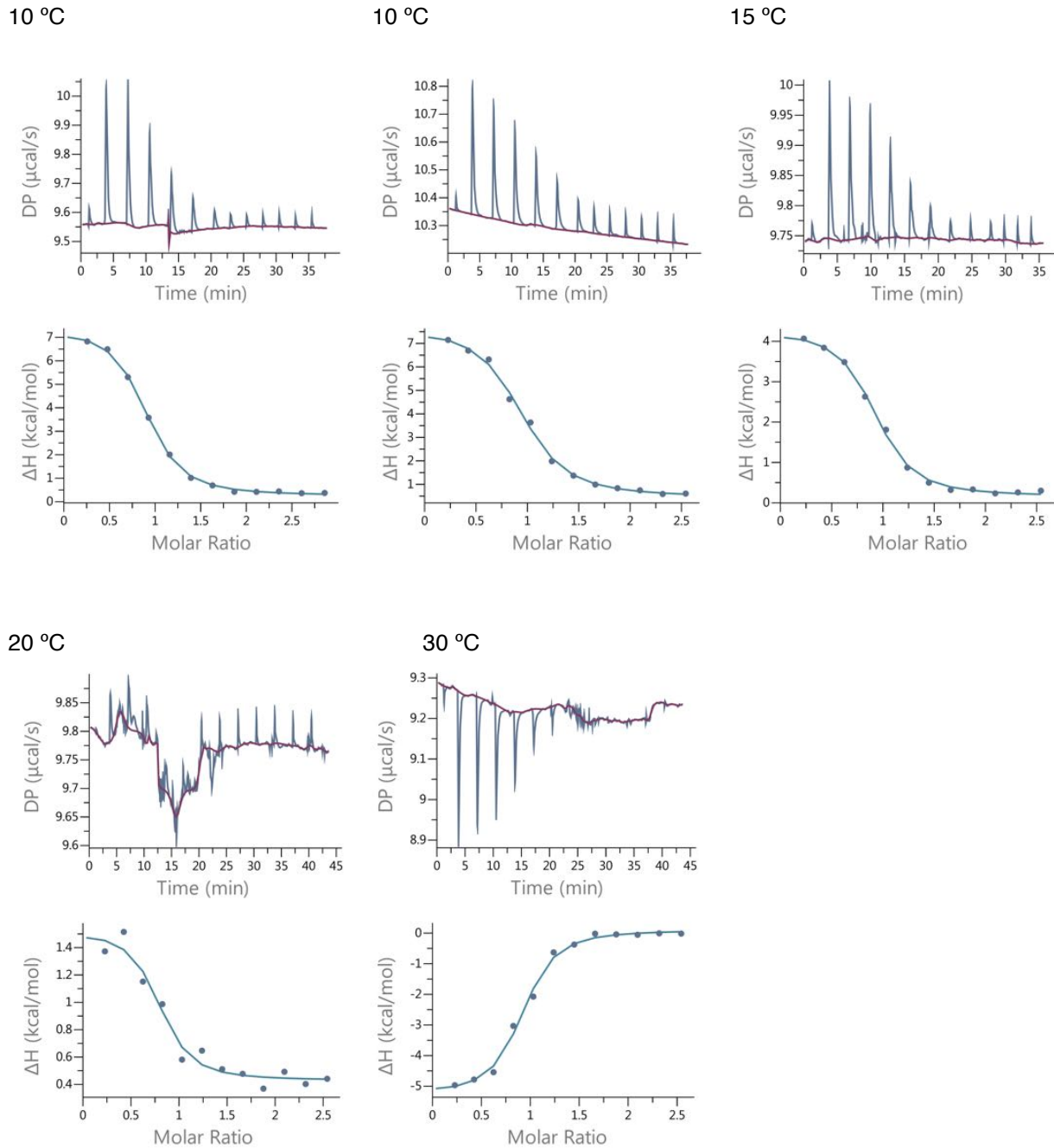


Figura A2.10: Medidas de ITC para la interacción RbAB – AdE1ACR1. Titulaciones calorimétricas hechas en fosfato de sodio 20mM, 200mM NaCl, 5 mM 2-mercapto etanol a pH 7.0. La celda fue cargada con RbAB aislado (30 μM) y la jeringa con AdE1ACR1 (300 μM). Titulaciones hechas a 10, 15, 20 y 30°C.

Figura A2.11. Medidas de ITC para la interacción RbAB – AdE1ACR2

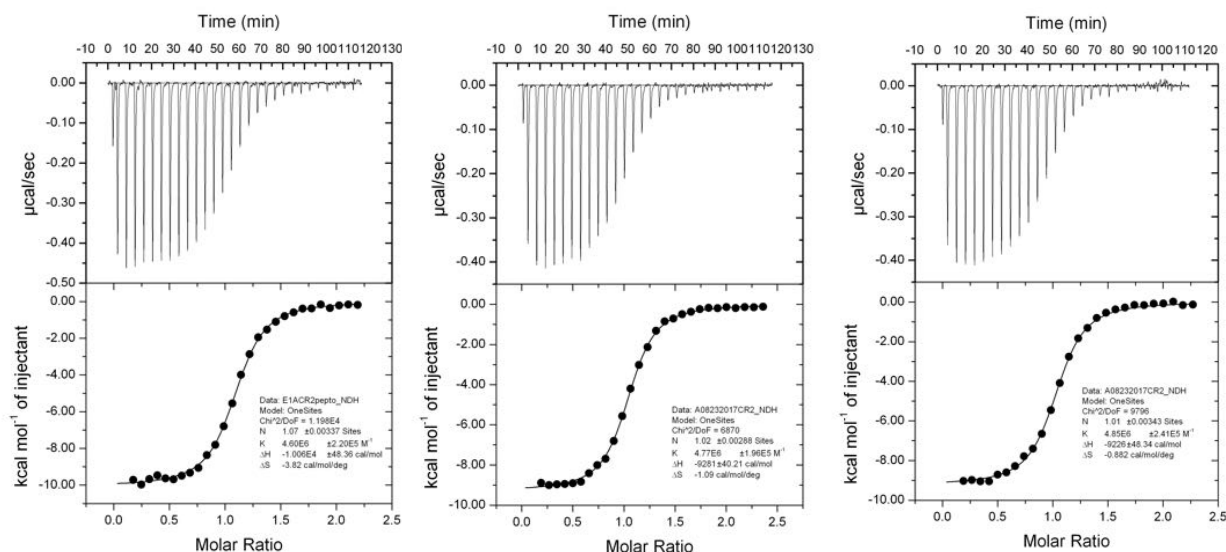


Figura A2.11: Medidas de ITC para la interacción RbAB – AdE1ACR2. Titulaciones calorimétricas hechas en fosfato de sodio 20mM, 200mM NaCl, 5 mM 2-mercapto etanol a pH 7.0. La celda fue cargada con RbAB aislado (15µM) y la jeringa con AdE1ACR2 (150µM). Titulaciones hechas a 20°C.

Figura A2.12. Medidas de ITC para la interacción RbAB – AdE1ACR2_{AC}

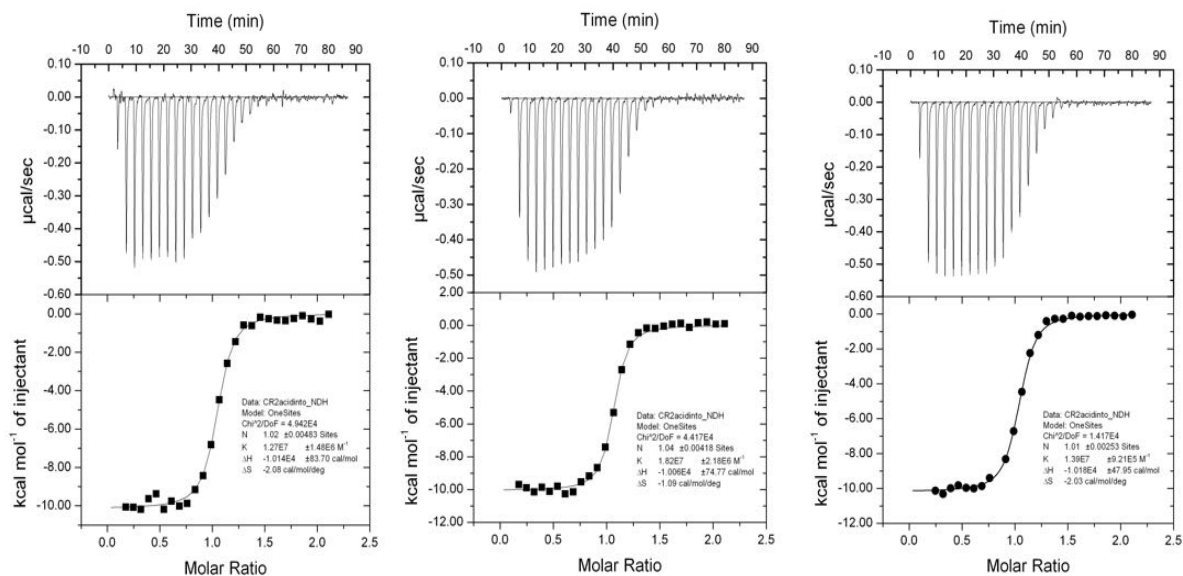


Figura A2.12: Medidas de ITC para la interacción RbAB – AdE1ACR2_{AC}. Titulaciones calorimétricas hechas en fosfato de sodio 20mM, 200mM NaCl, 5 mM 2-mercapto etanol a pH 7.0. La celda fue cargada con RbAB aislado (15µM) y la jeringa con AdE1ACR2_{AC} (150µM). Titulaciones hechas a 20°C.

Figura A2.13. Medidas de ITC para la interacción RbAB – AdE1ACR2_{ACP}

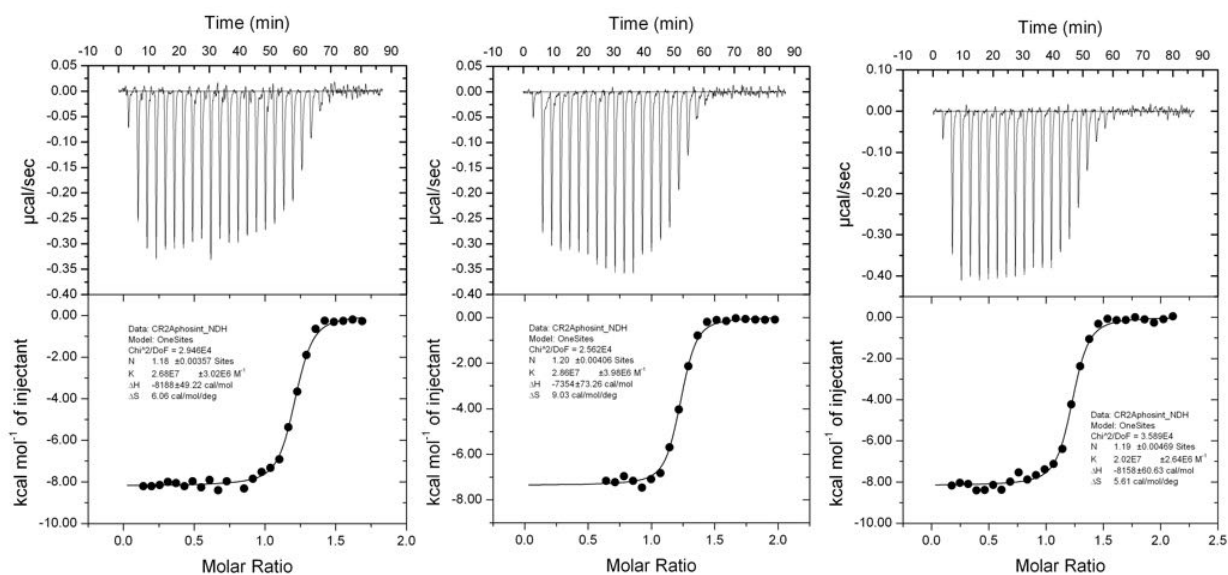


Figura A2.13: Medidas de ITC para la interacción RbAB – AdE1ACR2_{ACP}. Titulaciones calorimétricas hechas en fosfato de sodio 20mM, 200mM NaCl, 5 mM 2-mercapto etanol a pH 7.0. La celda fue cargada con RbAB aislado (15µM) y la jeringa con AdE1ACR2_{ACP} (150µM). Titulaciones hechas a 20°C.

Figura A2.14 Medidas de ITC para la interacción RbAB – AdE1A.WT

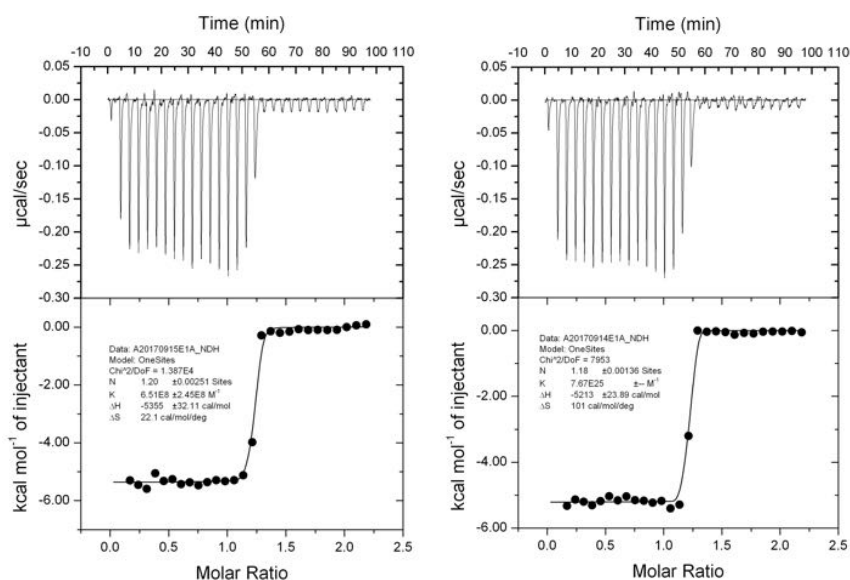


Figura A2.14: Medidas de ITC para la interacción RbAB – AdE1A.WT. Titulaciones calorimétricas hechas en fosfato de sodio 20mM, 200mM NaCl, 5 mM 2-mercapto etanol a pH 7.0. La celda fue cargada con RbAB aislado (15µM) y la jeringa con AdE1A.WT (150µM). Titulaciones hechas a 20°C.

Figura A2.15. Medidas de ITC para la interacción RbAB – AdE1A-ΔE

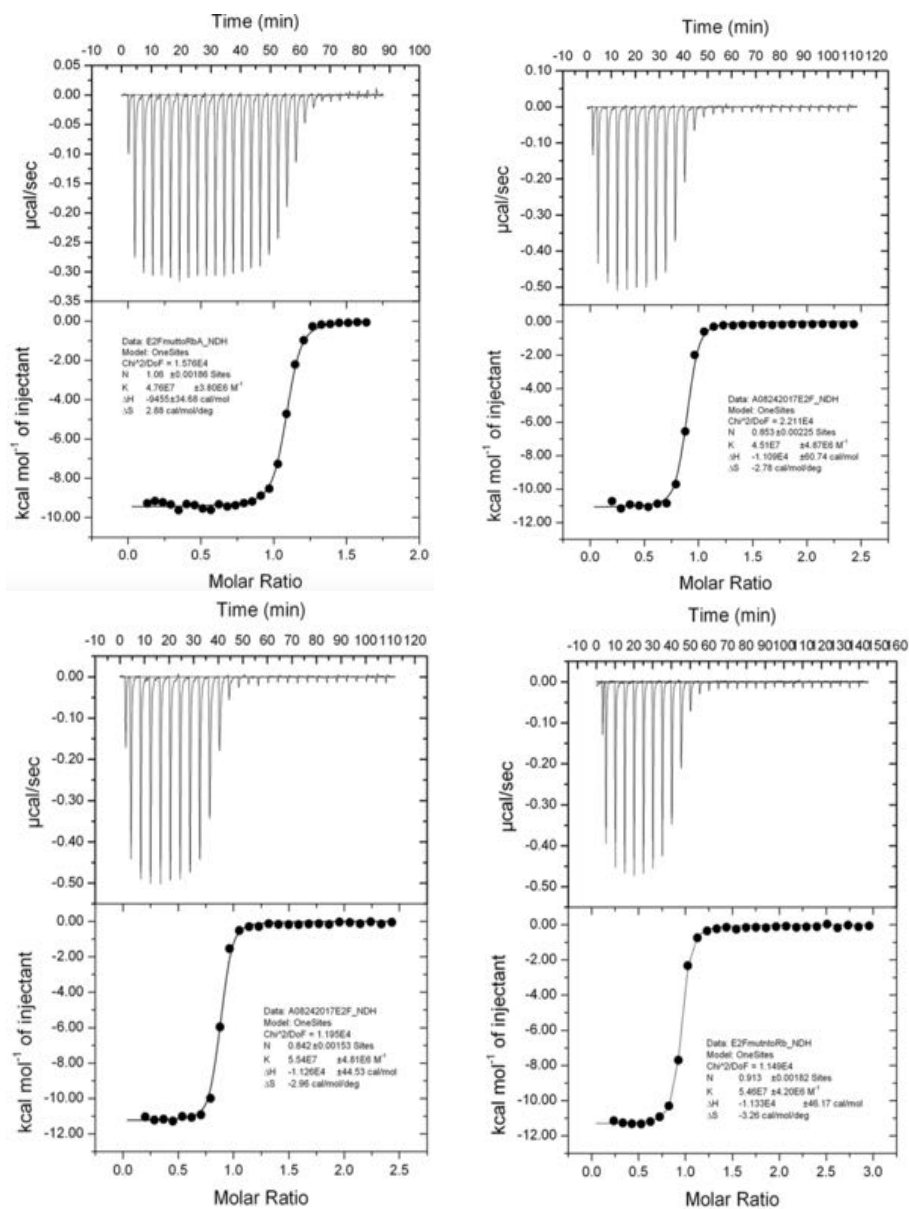


Figura A2.15: Medidas de ITC para la interacción RbAB – AdE1A-ΔE. Titulaciones calorimétricas hechas en fosfato de sodio 20mM, 200mM NaCl, 5 mM 2-mercapto etanol a pH 7.0. La celda fue cargada con RbAB aislado (15 μM) y la jeringa con AdE1A-ΔE (150 μM). Titulaciones hechas a 20°C.

Figura A2.16. Medidas de ITC para la interacción RbAB – AdE1A- Δ L

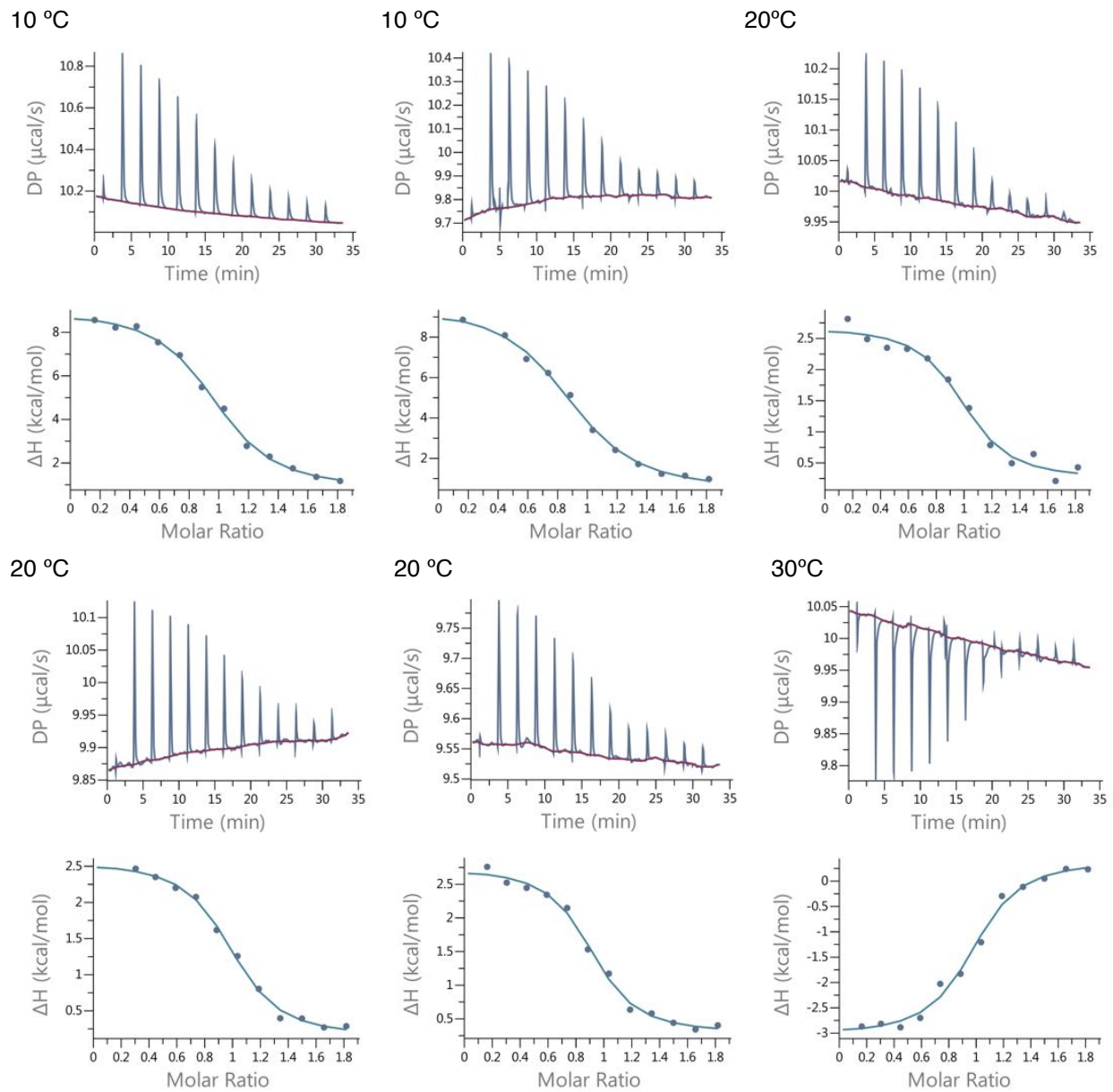


Figura A2.16: Medidas de ITC para la interacción RbAB – AdE1A- Δ L. Titulaciones calorimétricas hechas en fosfato de sodio 20mM, 200mM NaCl, 5 mM 2-mercapto etanol a pH 7.0. La celda fue cargada con RbAB aislado (30 μ M) y la jeringa con AdE1A- Δ L (300 μ M). Titulaciones hechas a 10, 20 y 30°C.

Figura A2.17. Medidas de ITC para la interacción [RbAB:AdE1ACR2] - AdE1ACR1

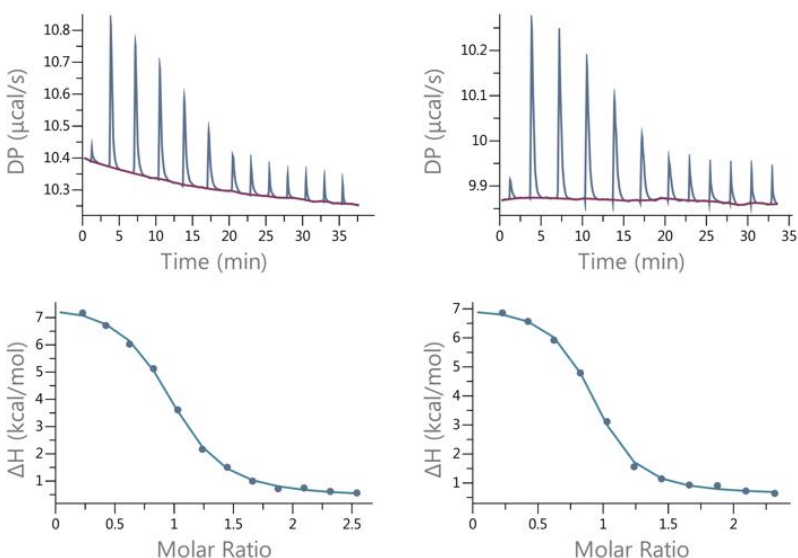


Figura A2.17: Medidas de ITC para la interacción [RbAB:AdE1ACR2] - AdE1ACR1. Titulaciones calorimétricas hechas en fosfato de sodio 20mM, 200mM NaCl, 5 mM 2-mercapto etanol a pH 7.0. La celda fue cargada con el complejo [RbAB:AdE1ACR2] (30µM) y la jeringa con AdE1ACR1 (300µM). Titulaciones hechas a 10°C.

Figura A2.18. Medidas de ITC para la interacción [RbAB:AdE1ACR2] - AdE1A-ΔL

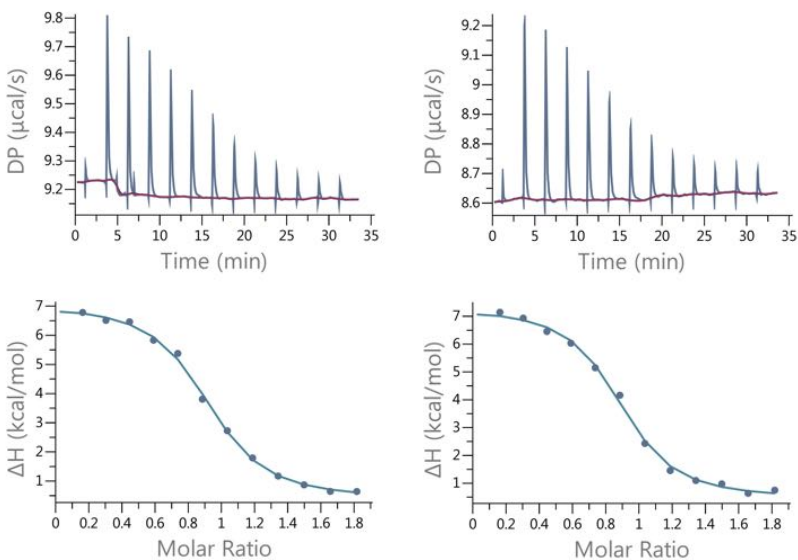


Figura A2.18: Medidas de ITC para la interacción [RbAB:AdE1ACR2] - AdE1A-ΔL. Titulaciones calorimétricas hechas en fosfato de sodio 20mM, 200mM NaCl, 5 mM 2-mercapto etanol a pH 7.0. La celda fue cargada con el complejo [RbAB:AdE1ACR2] (30µM) y la jeringa con AdE1A-ΔL (300µM). Titulaciones hechas a 10°C.

Figura A2.19. Medidas de ITC para la interacción [RbAB:AdE1ACR1] - AdE1ACR2

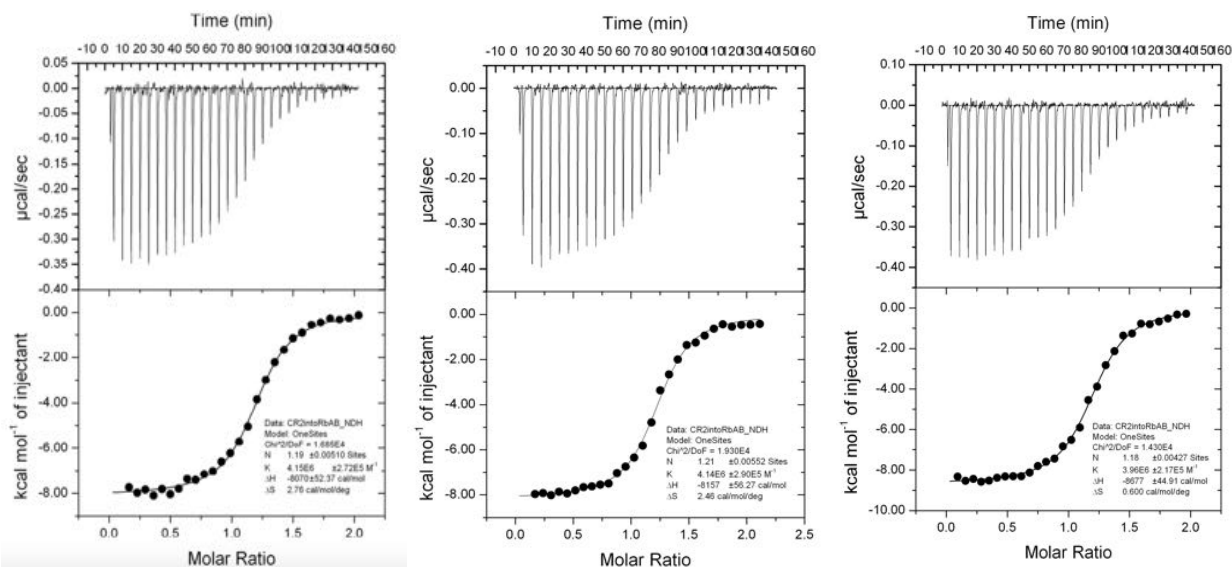


Figura A2.19: Medidas de ITC para la interacción [RbAB:AdE1ACR1] - AdE1ACR2. Titulaciones calorimétricas hechas en fosfato de sodio 20mM, 200mM NaCl, 5 mM 2-mercapto etanol a pH 7.0. La celda fue cargada con el complejo [RbAB:AdE1ACR1] (15µM) y la jeringa con AdE1ACR2 (150µM). Titulaciones hechas a 20°C.

Figura A2.20. Medidas de ITC para la interacción [RbAB:AdE1-ΔL] - AdE1ACR2

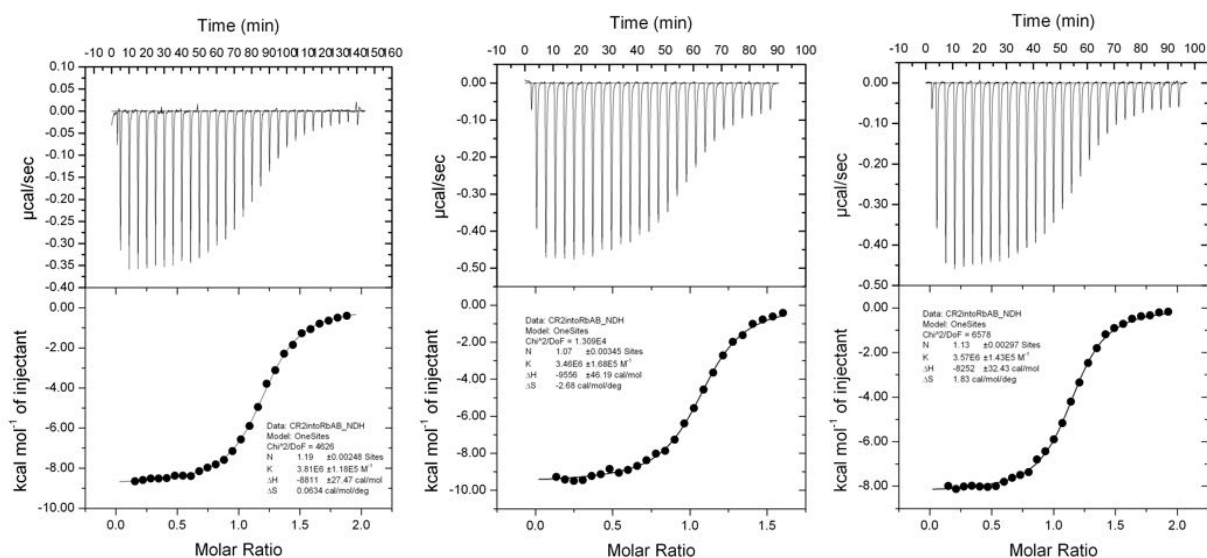


Figura A2.20: Medidas de ITC para la interacción [RbAB:AdE1-ΔL] - AdE1ACR2. Titulaciones calorimétricas hechas en fosfato de sodio 20mM, 200mM NaCl, 5 mM 2-mercapto etanol a pH 7.0. La celda fue cargada con el complejo [RbAB:AdE1A-ΔL] (15µM) Y la jeringa con AdE1ACR2 (150µM). Titulaciones hechas a 20°C.

Tabla A3.1. Asignaciones de AdE1A.WT

#	AA	¹³ C _α	¹³ C _β	¹³ C'	¹⁵ N	¹ H
38	F	57.47370148	39.91873169	174.8260193	0	0
39	E	53.62131882	30.21705055	0	125.0963287	8.150650024
40	P	0	0	0	0	0
41	P	62.99200821	32.09835052	177.0219879	0	0
42	T	61.80664063	69.89299774	174.5193939	114.6051636	8.189869881
43	L	55.41648483	42.27530861	177.041748	124.4604568	8.24862957
44	H	55.99425507	30.15450001	174.9949646	119.6560364	8.28390789
45	E	56.84124947	30.15196514	176.2462769	121.8205414	8.351630211
46	L	55.29841614	42.39996338	177.0053253	122.2783127	8.1895504
47	Y	57.45692444	38.91101837	175.4439545	119.9560471	7.988220215
48	D	54.22191048	41.14070511	175.1119232	121.6376572	8.218729973
49	L	55.34621048	42.60874939	177.0961304	122.1169662	7.997169971
50	D	54.36949921	41.05059052	176.2873993	121.2270889	8.363209724
51	V	62.12386513	32.76477432	176.3101959	119.9560471	7.978789806
52	T	61.84964943	69.87884903	173.9026489	118.3804474	8.237569809
53	A	50.51531982	18.33469963	0	128.5323334	8.309439659
54	P	62.95766068	32.16529083	176.8289642	0	0
55	E	56.50700378	30.39158058	176.0883026	121.2789688	8.45586586
56	D	52.01332	41.25714111	0	123.0228729	8.435440063
57	P	63.47385025	32.18347168	176.9849548	0	0
58	N	53.43348503	39.04151917	175.391922	118.5427628	8.534790039
59	E	56.87279892	30.38083076	0	121.5763092	8.175660133
60	E	56.67969131	30.33616066	176.3710022	0	0
61	A	52.57976532	19.19147015	177.9182739	125.3719101	8.274820328
62	V	62.40439034	32.83436966	176.411499	119.3822784	8.086319923
63	S	58.35549927	63.835989	174.3047943	119.4096069	8.321000099
64	Q	55.65537071	29.70125008	175.9684753	122.6100006	8.298800468
65	I	60.79681015	39.06282997	176.0714722	121.9487534	7.989303112
66	F	55.44195175	39.09434128	0	125.5397263	8.425430298
67	P	63.36132813	32.18695831	176.7224426	0	0
68	D	54.73461151	41.04146004	176.6206207	120.1801605	8.375329971
69	S	59.02842522	63.68170929	174.9390411	115.5183563	8.198699951
70	V	63.0493009	32.7166748	176.3555298	122.0057144	8.003458977
71	M	55.70055008	32.69810486	176.3995209	123.0245438	8.252169609
72	L	55.29402924	42.48771095	177.0900879	123.2336884	8.136360168
73	A	52.6607399	19.15071583	177.9018707	124.6113129	8.114549637
74	V	62.61125565	32.77635002	176.4939575	119.3309021	8.017979622
75	Q	55.93798065	29.61995506	175.8764343	123.9077911	8.388239861
76	E	56.89274025	30.42590427	177.0187836	122.4783936	8.439860344
77	G	45.64166069	0	174.3319244	110.0407028	8.43268013
78	I	61.24258041	38.95266533	175.9290619	119.791748	7.862949848
79	D	54.38987541	41.03661537	176.2138214	123.8508072	8.422570229
80	L	55.48847961	42.46382141	177.4093018	122.7784271	8.130379677
81	L	55.37561417	42.15764046	177.3736572	121.50354	8.125579834
82	T	61.83649063	69.8285141	173.6454773	114.3048325	7.783929825
83	F	55.44215012	39.11143875	0	122.9358063	8.123399734
84	P	0	0	0	0	0
85	P	62.66296005	32.07341003	176.4081879	0	0
86	A	50.35681152	18.14567947	0	125.6589127	8.409790039
87	P	63.42837906	31.98933029	177.6435242	0	0
88	G	45.10085487	0	174.0309448	109.7939377	8.508856773
89	S	56.59640121	63.37731934	0	117.0972137	8.143580437
90	P	63.09379959	32.06245041	176.681427	0	0
91	E	54.37865829	29.86895943	0	122.9544373	8.414489746
92	P	0	0	0	0	0
93	P	63.37934113	32.17889023	177.0800476	0	0
94	H	56.46497536	30.00349522	175.1409149	120.9913864	8.627779961
95	L	54.99829483	42.46069527	177.1927185	123.661293	8.106220245
96	S	58.27777481	63.77594566	174.3279724	116.9348526	8.297189713
97	R	55.88782501	31.00198936	175.9673157	123.0048981	8.294980049
98	Q	53.70639038	28.86494064	0	123.1734009	8.455280304
99	P	63.17823029	32.1971283	176.8212433	0	0
100	E	56.27315945	30.08153439	176.3629761	121.2797165	8.542090416
101	Q	53.51683044	29.1015892	0	122.0867996	8.40225029
102	P	0	0	0	0	0
103	E	56.76692963	29.9695797	176.671875	0	0

104	Q	55.89204979	30.17581463	175.9416656	121.7789307	8.382200241
105	R	55.96484947	30.98435497	175.9058533	122.6579666	8.306710243
106	A	52.4169445	19.19102001	177.5550537	125.4971008	8.333530426
107	L	55.09207535	42.6355896	177.618576	121.8717728	8.231769562
108	G	44.56362152	0	0	109.5503387	8.13076973
109	P	63.10945892	32.29798889	177.2290649	0	0
110	V	62.22076607	32.89286995	176.184494	120.322197	8.242440224
111	S	58.01397514	63.7376442	174.0231323	119.5137863	8.330639839
112	M	53.30144882	32.64104843	0	123.702652	8.366829872
113	P	63.27806091	32.14242935	176.4578705	0	0
114	N	53.24057961	38.74513054	174.8954163	118.409523	8.476103783
115	L	55.2993145	42.43600464	176.9174652	123.0421066	8.159509659
116	V	59.81734085	32.58496857	0	122.9768829	8.133319855
117	P	63.13185883	32.22306061	176.719986	0	0
118	E	56.61384392	30.46060562	176.4286346	121.4701004	8.44604969
119	V	62.33395386	32.81748962	175.9090729	122.4415436	8.229780197
120	I	60.85807991	38.97042084	175.4528961	125.68647	8.230319977
121	D	53.83547974	41.32465935	176.3209534	125.1799088	8.4230299
122	L	55.40098	42.07870483	177.986908	124.8051376	8.462670326
123	T	62.87172127	69.56099319	174.7523041	114.1819992	8.312959671
124	C	58.54325104	27.80183983	0	121.0039978	8.131270409
125	H	56.1109581	30.59070969	175.003006	0	0
126	E	56.49405479	30.33754063	176.1519928	122.5337296	8.390509605
127	A	52.73580933	19.35484505	178.0116119	125.175087	8.361450195
128	G	44.90852928	0	173.4022675	107.6490173	8.196240425
129	F	55.60057068	38.99501038	0	120.7666397	8.029569626
130	P	0	0	0	0	0
131	P	62.92269135	32.17998886	177.0099487	0	0
132	S	58.0240097	64.08892059	174.456604	116.439827	8.461939812
133	D	54.42668533	41.34260559	175.9682159	122.8916168	8.450750351
134	D	54.55898666	41.28128052	176.3065643	120.5675278	8.282569885
135	E	56.55287933	30.50139999	0	121.1973038	8.316399574
136	D	0	0	0	0	0
137	E	56.55287933	30.7415905	176.5680237	0	0
138	E	56.7537117	30.43916988	177.0428009	122.4492264	8.4896698
139	G	45.20516968	0	174.0378723	110.2346573	8.398460388
140	E	56.49635506	30.7163105	176.3893738	120.865593	8.293410301
141	E	56.45645905	30.41569996	175.9757385	122.3013611	8.49450016
142	F	57.52707672	39.70749474	175.2180023	122.1405716	8.256250381
143	V	62.07068062	33.06620598	175.3856354	123.558609	8.042300224
144	L	54.952425	42.6533699	176.5311432	126.6923523	8.225020409
145	D	54.09724426	41.27142525	174.5745087	121.7958527	8.145910263
146	Y	58.96345901	39.42071915	0	124.6595078	7.477159977

Tabla A3.2. Asignaciones de AdE1A-ΔE

#	AA	¹³ C _α	¹³ C _β	¹³ C'	¹⁵ N	¹ H
38	F	57.46929169	39.91289902	174.8041077	0	0
39	E	53.60271835	30.23572922	0	125.2528534	8.158650398
40	P	0	0	0	0	0
41	P	62.98117065	32.08486938	177.1553497	0	0
42	T	61.79909134	69.96531677	174.371109	114.5603638	8.158140182
43	A	0	0	0	126.8257828	8.319419861
44	A	0	0	0	0	0
45						
46	A	0	0	0	0	0
47	A	52.55218124	19.43363953	177.4853058	0	0
48	D	54.28071594	41.01625061	176.1730957	119.2335281	8.21944046
49	L	55.2336998	42.59325981	177.0731964	122.0772629	8.027819633
50	D	54.30761337	41.04783058	176.2566071	121.5618668	8.351059914
51	V	62.11032486	32.78547096	176.331955	120.1317368	8.015740395
52	T	61.87459373	69.85126877	173.9038239	118.5279236	8.263469696
53	A	50.51673889	18.33945084	0	128.6332092	8.32365036
54	P	62.78733063	32.14582825	176.8413086	0	0
55	E	56.46336555	30.47450066	176.0699768	121.3602066	8.468930244
56	D	52.01327133	41.33512878	0	123.0714188	8.440620422
57	P	63.47092056	32.19620132	176.9754181	0	0
58	N	53.43559074	39.03841019	175.3814392	118.5946732	8.543120384
59	E	56.86904907	30.37977028	0	121.6255035	8.182530403
60	E	56.6640892	30.33448982	176.3524628	0	0
61	A	52.58033562	19.20326519	177.8988495	125.4254303	8.282010078
62	V	62.3465004	32.90987968	176.3969574	119.4400024	8.094579697
63	S	58.25461006	63.81022453	174.3119659	119.4721832	8.329039574
64	Q	55.64291382	29.65499496	175.1071167	122.6486282	8.303959846
65	I	60.79438019	38.84435463	176.0565948	121.9861832	8.000829697
66	F	55.44311905	39.33435059	0	125.5693207	8.432370186
67	P	63.3637085	32.20124817	176.7109222	0	0
68	D	54.73756027	41.04990959	176.603653	120.2326431	8.383500099
69	S	59.01981544	63.6776104	174.9582062	115.564682	8.205579758
70	V	63.03481483	32.69682884	176.3417053	122.0497894	8.011540413
71	M	55.72887039	32.5892086	176.3904266	123.0753403	8.259730339
72	L	55.41513062	42.49753952	177.0771637	123.2770004	8.14338398
73	A	52.67439842	19.15561485	177.8909302	124.6455612	8.121390343
74	V	62.61179543	32.79873466	176.4823608	119.3547363	8.022689819
75	Q	55.96455002	29.65111065	175.8301392	123.9403229	8.395620346
76	E	56.88542557	30.58279037	177.0011597	122.5084076	8.445650101
77	G	45.63652039	0	174.3184509	110.091507	8.438960075
78	I	61.23707008	38.96496582	175.9199677	119.8540115	7.870729923
79	D	54.34723473	41.07131004	176.1976776	123.9008102	8.42994976
80	L	55.57601929	42.22907066	177.3899536	122.8491135	8.137909889
81	L	55.37360382	42.15715408	177.3553925	121.5617599	8.133440018
82	T	61.82672882	69.80929184	173.62883	114.359642	7.792650223
83	F	55.62385178	39.09128571	0	122.9449234	8.128629684
84	P	0	0	0	0	0
85	P	62.66341019	32.06581879	176.3777008	0	0
86	A	50.35546112	18.15703011	0	125.7138596	8.41588974
87	P	63.34806061	32.01237106	177.6293945	0	0
88	G	45.10378456	0	174.0158539	109.8492737	8.512840271
89	S	56.59310913	63.39025116	0	117.1501617	8.150320053
90	P	63.10514832	32.08776855	176.6666412	0	0
91	E	54.36753082	29.86866951	0	123.0234833	8.420410156
92	P	0	0	0	0	0
93	P	63.37229919	32.19252014	177.0643616	0	0
94	H	56.30389595	30.03713512	175.0982361	121.0418472	8.634460449
95	L	55.12318993	42.43112564	177.1757813	123.7344971	8.123009682
96	S	58.31020927	63.74503517	174.3185883	117.003212	8.309029579
97	R	55.77186584	31.002985	175.9617157	123.0595016	8.304300308
98	Q	53.73160934	28.88372993	0	123.2471237	8.470560074
99	P	63.16844177	32.20756912	176.8067169	0	0
100	E	56.62903023	30.26192474	176.3467255	121.3283081	8.548800468
101	Q	53.51921082	29.10136986	0	122.1389236	8.408690453
102	P	0	0	0	0	0
103	E	56.5680542	29.98579979	176.6594391	0	0
104	Q	56.25016594	30.1192503	175.9259644	121.8669434	8.389069557

105	R	56.08732414	31.04111004	175.8921509	122.6825485	8.309729576
106	A	52.42346001	19.19136047	177.5370331	125.5470734	8.340800285
107	L	55.10978508	42.62466431	177.6024628	121.9176025	8.237890244
108	G	44.56447983	0	0	109.6041336	8.137999535
109	P	63.1069603	32.32389069	177.2162476	0	0
110	V	62.24822998	32.90154076	176.1955414	120.3668137	8.249079704
111	S	58.09144974	63.801754	174.0079498	119.5739365	8.334540367
112	M	53.29859161	32.63100052	0	123.7805481	8.369370461
113	P	63.29296112	32.12234116	176.4449158	0	0
114	N	53.26353836	38.75434875	174.8682251	118.4327164	8.487609863
115	L	55.39213943	42.10789108	176.4276428	123.0888367	8.167639732
116	V	59.81695175	32.58351898	0	123.0024033	8.138130188
117	P	63.1787796	32.23043823	176.702713	0	0
118	E	56.64141655	30.50744915	176.4093628	121.5294571	8.449959755
119	V	62.33493996	32.81381416	175.8981628	122.4728775	8.23526001
120	I	60.81568527	38.97169304	175.4343414	125.7307968	8.237330437
121	D	53.83934021	41.28087425	176.3072968	125.2036133	8.427410126
122	L	55.39991379	42.10027504	177.9735107	124.8423233	8.469539642
123	T	62.83122444	69.58661652	174.7386322	114.2003021	8.320199966
124	C	58.54476929	27.80023956	0	121.0419998	8.136719704
125	H	56.02743912	30.31918907	174.9584198	0	0
126	E	56.68025017	30.21575546	176.1500244	122.5772171	8.398900032
127	A	52.74131012	19.36147499	177.9957428	125.2241898	8.369250298
128	G	44.91098404	39.00368118	173.3905182	107.7102966	8.20596981
129	F	55.58625031	39.00368118	0	120.8190231	8.034950256
130	P	0	0	0	0	0
131	P	62.9221611	32.18830872	176.9963531	0	0
132	S	58.01912117	64.10274506	174.4477386	116.4870834	8.469180107
133	D	54.55338478	41.4237957	175.9548492	123	8.465600014
134	D	54.46821404	41.30541039	176.2946625	120.5620575	8.290679932
135	E	56.55828094	30.50519943	0	121.2353363	8.322509766
136	D	0	0	0	0	0
137	E	56.72805023	30.56038094	176.5527649	0	0
138	E	56.68836975	30.43780994	177.0257568	122.4287186	8.498570442
139	G	45.21812439	0	174.0312653	110.2804565	8.405520439
140	E	56.49922562	30.5646801	176.4220123	120.9147873	8.300680161
141	E	56.53917503	30.37330055	175.9634247	122.3269424	8.501190186
142	F	57.52628899	39.70334625	175.2043457	122.1641464	8.261300087
143	V	62.05426979	33.07002068	175.3757324	123.5848312	8.04570961
144	L	54.7892952	42.74082947	176.5161285	126.7265396	8.231049538
145	D	54.01309586	41.29047966	174.5589294	121.8478699	8.154000282
146	Y	58.96630859	39.4180603	0	124.717659	7.485330105

Tabla A3.3. Asignaciones de AdE1A-ΔL

#	AA	¹³ C _α	¹³ C _β	¹³ C'	¹⁵ N	¹ H
38	F	55.63185883	39.90929031	174.812439	0	0
39	E	53.61968994	30.22941971	0	125.1308365	8.15569973
40	P	0	0	0	0	0
41	P	63.00254059	32.09680176	177.0070038	0	0
42	T	61.79762459	69.92101288	174.5082397	114.5815506	8.187789917
43	L	55.65509415	42.2551899	177.0366821	124.4450836	8.247480392
44	H	56.26856613	30.2260828	174.9025116	119.5753784	8.30012989
45	E	56.70738602	30.3313446	176.24263	121.8323517	8.361889839
46	L	55.30360031	42.41535568	176.991684	122.2737808	8.191630363
47	Y	57.44504547	38.913311	175.4343567	119.9478531	7.987909794
48	D	54.2751503	41.14263916	175.0884705	121.6212463	8.218079567
49	L	55.36481667	42.60042953	177.0746765	122.0942078	7.996870041
50	D	54.3629055	41.0384655	176.2686615	121.2207031	8.364020348
51	V	61.99547958	32.77067947	176.2950592	119.8862076	7.978479862
52	T	61.86006927	69.84361267	173.8864441	118.359581	8.238030434
53	A	50.51662827	18.34075928	0	128.5229034	8.307849884
54	P	63.05110168	32.19303894	176.7867737	0	0
55	E	56.51204109	30.5535202	176.0420685	121.2975464	8.458931923
56	D	52.00497055	41.35725021	0	123.0323868	8.435310364
57	P	63.47409058	32.20298004	176.973175	0	0
58	N	53.43486023	39.04430962	175.3786469	118.5318069	8.535719872
59	E	56.86698151	30.38098907	0	121.5638733	8.17512989
60	E	56.66886902	30.32652092	176.3537292	0	0
61	A	52.58606911	19.203475	177.902298	125.3648071	8.274749756
62	V	62.3526001	32.85840034	176.3850708	119.3728714	8.086190224
63	S	58.33645058	63.81368065	174.2695313	119.4077682	8.322669983
64	Q	55.63206482	29.67445946	175.9514923	122.5646973	8.294630051
65	I	60.79140472	38.83160973	175.5537567	121.9817886	7.989299774
66	F	55.43896103	39.24948883	0	125.4847488	8.42343998
67	P	63.36011887	32.20069122	176.700882	0	0
68	D	54.73362541	41.05062485	176.6066437	120.1768875	8.376279831
69	S	59.02641869	63.67823982	174.9433746	115.5102921	8.199150085
70	V	63.05031013	32.56677437	176.3391571	121.9380569	8.002940178
71	M	55.44270515	32.72203445	176.4008484	123.0066376	8.251390457
72	L	55.48466492	42.48742485	177.0774231	123.2251968	8.136079788
73	A	52.67401886	19.15744495	177.8905487	124.587822	8.113089561
74	V	62.61029053	32.811177521	176.4822998	119.3002777	8.016309738
75	Q	55.94893456	29.65323448	175.8763885	123.8452988	8.384080887
76	E	56.89015007	30.49475002	177.0002441	122.4482269	8.438879967
77	G	45.63535881	0	174.3168488	110.0372772	8.432020187
78	I	61.22900009	38.96084976	175.9172516	119.8018417	7.863810062
79	D	54.40224457	41.04721832	176.1985016	123.8475571	8.423649788
80	L	55.48317528	42.29154396	177.387146	122.7393112	8.136819839
81	L	55.38341904	42.11235046	177.3552246	121.5028763	8.126429558
82	T	61.82750511	69.80801392	173.6295471	114.3013763	7.785550117
83	F	55.49335098	39.11912918	0	122.8411713	8.127165794
84	P	0	0	0	0	0
85	P	62.65803909	32.07205963	176.3966217	0	0
86	A	50.35491943	18.15559006	0	125.6742096	8.407406807
87	P	63.40568161	32.05194092	177.6274567	0	0
88	G	45.10692406	0	174.0156708	109.8001633	8.506349564
89	S	56.59148026	63.37393951	0	117.0967331	8.143719673
90	P	63.09896088	32.06592941	176.6628113	0	0
91	E	54.35665894	29.87994003	0	122.9449463	8.410679817
92	P	0	0	0	0	0
93	P	63.37282944	32.19483948	177.0633392	0	0
94	H	56.33912086	29.98248005	175.0360718	120.9918365	8.628259659
95	L	55.0396595	42.44821548	177.1793671	123.7175827	8.130800247
96	S	58.33333969	63.75475502	174.3104858	116.9744797	8.308919907
97	R	55.91214561	31.03125	175.9631958	123.0110397	8.301119804
98	Q	53.67847061	28.88726997	0	123.1747437	8.457550049
99	P	63.17181015	32.19018173	176.8046722	0	0
100	E	56.67779922	30.07582474	176.3453674	121.2785568	8.542949677
101	Q	53.51580811	29.10193062	0	122.0801086	8.403429985
102	P	0	0	0	0	0
103	E	56.69343948	30.19956017	176.6582489	0	0
104	Q	56.41558456	29.41907024	175.922348	121.7886276	8.383839607

105	R	56.00064087	31.00144577	175.891449	122.6521378	8.308859825
106	A	52.75541496	19.05611515	177.5374908	125.4988327	8.334039688
107	L	55.09457016	42.6287365	177.6011047	121.9336395	8.234649658
108	G	44.5642395	0	0	109.551712	8.131349564
109	P	63.10486984	32.19763947	177.2139282	0	0
110	V	62.28184509	32.87586975	176.1842041	120.3175964	8.242670059
111	S	58.04833984	63.75557899	174.0061035	119.5389404	8.330980301
112	M	53.28216171	32.62311172	0	123.7139664	8.367650032
113	P	63.25299072	32.14897156	176.4434052	0	0
114	N	53.24690437	38.74264908	174.8759003	118.3833771	8.480270386
115	L	55.2480011	42.50606155	176.8969421	123.0401611	8.158143044
116	V	59.82926941	32.57563019	0	122.9246674	8.120094299
117	P	0	0	0	0	0
118	E	56.73733521	30.45318985	176.4669037	0	0
119	V	62.34817123	32.81834984	176.0154266	122.3972626	8.216070175
120	I	60.94791412	38.82050514	175.8605804	125.6777878	8.258609772
121	D	54.04265976	41.3004303	176.1935883	125.1813736	8.389610291
122	A	53.28796005	19.01708031	0	126.0750198	8.350720406
123	A	0	0	0	0	0
124	A	0	0	0	0	0
125						
126	A	52.8678093	19.32533073	177.7024078	0	0
127	A	52.61015892	19.27017975	178.0172424	122.5010376	7.998760223
128	G	44.94784546	0	173.422821	107.2296524	8.049369812
129	F	55.60427856	38.87174988	0	120.6811066	8.02038002
130	P	0	0	0	0	0
131	P	62.91897964	32.20138931	177.0020599	0	0
132	S	58.0217514	64.06739807	174.4410858	116.4536514	8.469129562
133	D	54.44499969	41.3126297	175.9483185	122.95224	8.453410149
134	D	54.52046967	41.36087418	176.2902069	120.5538864	8.285400391
135	E	56.54841995	30.49972916	0	121.189743	8.318699837
136	D	0	0	0	0	0
137	E	56.57820892	30.45351982	176.3757935	0	0
138	E	56.71848106	30.49102497	177.0129852	122.3038101	8.494540215
139	G	45.21746445	0	174.0276337	110.2304611	8.398920059
140	E	56.46600533	30.53678417	176.5658569	120.8599472	8.293410301
141	E	56.62134933	30.51327515	175.9567871	122.4204025	8.486820221
142	F	57.52326393	39.70647049	175.2036896	122.1253586	8.255060196
143	V	62.0556736	33.07628059	175.3686218	123.5441971	8.041130066
144	L	54.69237518	42.23319626	176.5147247	126.6837769	8.224749565
145	D	54.09323502	41.29727936	174.5572052	121.7945023	8.146769524
146	Y	58.96376038	39.42000961	0	124.6629333	7.478960037

Tabla A3.4. Desplazamientos químicos secundarios de AdE1A.WT

#	AA	$^{13}\text{C}_\alpha\Delta\sigma$	$^{13}\text{C}_\beta\Delta\sigma$	$^{13}\text{C}'\Delta\sigma$	$^{15}\text{N}\Delta\sigma$	$^1\text{H}\Delta\sigma$
38	F	0	0	0	0	0
39	E	-0.438681183	0.357050552	0	1.016328735	-0.209349976
40	P	0	0	0	0	0
41	P	-0.087991791	0.088350525	-0.158012085	0	0
42	T	0.046640625	0.302997742	0.049393921	0.405163574	-0.060130119
43	L	0.256484833	0.095308609	-0.188251953	0.130456848	0.08862957
44	H	0.014255066	0.514500008	0.2649646	-0.013963623	-0.11609211
45	E	0.301249466	0.071965141	-0.023723145	-0.309458618	-0.008369789
46	L	0.018416138	0.289963379	-0.154674683	-0.441687317	0.0195504
47	Y	-0.293075562	0.221018372	0.123954468	-0.733952942	-0.141779785
48	D	0.321910477	0.090705109	-0.628076782	-0.902342834	0.138729973
49	L	-0.10378952	0.49874939	-0.143869629	-0.063033752	-0.172830029
50	D	0.169499207	0.060590515	0.077399292	0.237088928	0.093209724
51	V	-0.396134872	0.184774323	-0.189804077	-1.203952942	-0.181210194
52	T	0.039649429	0.21884903	-0.207351074	-0.149552612	-0.012430191
53	A	0.195319824	-0.265300369	0	0.872333374	-0.030560341
54	P	-0.312339325	0.165290833	-0.281035767	0	0
55	E	-0.072996216	0.241580582	-0.141697388	0.308968811	-0.13413414
56	D	-0.06668	0.717141113	0	0.522872925	0.075440063
57	P	0.07385025	0.30347168	0.234954834	0	0
58	N	0.253485031	0.341519165	0.111921997	-0.597237244	-0.105209961
59	E	0.00279892	0.290830765	0	0.206309204	-0.244339867
60	E	0.049691315	0.21616066	0.091002197	0	0
61	A	0.01976532	0.161470146	0.198273926	0.281910095	-0.005179672
62	V	0.044390335	0.204369659	0.051499023	-0.327721558	-0.033680077
63	S	0.025499268	0.215988998	-0.185205688	-0.950393066	-0.138999901
64	Q	-0.214629288	0.251250076	0.148475342	0.26000061	-0.121199532
65	I	-0.32318985	0.542829971	0.221472168	0.218753357	-0.060696888
66	F	0.021951752	-0.025658722	0	0.119726257	0.095430298
67	P	0.441328125	-0.023041687	0.702442627	0	0
68	D	0.534611511	0.021460037	0.160620728	-0.539839478	-0.134670029
69	S	0.598425217	0.05170929	0.309041138	-1.051643677	-0.041300049
70	V	0.609300901	0.396674805	-0.014470215	-0.554285583	-0.196541023
71	M	0.220550079	0.208104858	0.259520874	-1.085456238	-0.127830391
72	L	0.194029236	0.397710953	-0.009912109	-0.046311646	-0.103639832
73	A	0.190739899	0.150715828	0.261870728	-0.078687134	-0.095450363
74	V	0.201255646	0.236350021	0.20395752	-0.3490979	-0.092020378
75	Q	0.097980652	0.189955063	0.066434326	-0.722208862	-0.111760139
76	E	0.12274025	0.265904274	0.048783569	0.008393555	-0.090139656
77	G	0.36166069	0	0.291924438	0.20070282	0.02268013
78	I	0.092580414	0.312665329	-0.11093811	-0.348251953	-0.097050152
79	D	0.329875412	0.106615372	-0.116178589	-0.15919281	0.042570229
80	L	0.118479614	0.483821411	-0.070698242	0.618427124	0.050379677
81	L	0.155614166	-0.002359543	-0.216342773	-1.286459961	-0.104420166
82	T	0.076490631	0.158514099	-0.474522705	-0.075167542	-0.236070175
83	F	-0.207849884	-0.048561249	0	-0.954193726	-0.166600266
84	P	0	0	0	0	0
85	P	-0.347039948	0.133410034	-0.281812134	0	0
86	A	0.156811523	-0.384320526	0	0.108912659	-0.010209961
87	P	0.088379059	0.049330292	0.03352417	0	0
88	G	-0.009145126	0	-0.079055176	0.493937683	-0.091143227
89	S	0.666401215	0.037319336	0	0.087213745	-0.046419563
90	P	-0.166200409	0.082450409	-0.228572998	0	0
91	E	0.128658295	0.218959427	0	0.544437256	-0.155510254
92	P	0	0	0	0	0
93	P	0.429341125	0.218890228	0.260047607	0	0
94	H	0.664975357	0.483495216	0.440914917	1.891386414	0.107779961
95	L	-0.18170517	0.250695267	-0.117281494	0.08129303	-0.123779755
96	S	0.097774811	0.105945663	-0.172027588	0.2348526	0.017189713
97	R	-0.262174988	0.461989365	-0.172684326	-0.495101929	-0.155019951
98	Q	0.256390381	-0.185059357	0	-0.036599121	-0.024719696
99	P	-0.051769714	0.127128296	-0.068756714	0	0
100	E	-0.29684055	0.001534386	-0.167023926	0.029716492	-0.037909584
101	Q	-0.073169556	0.141589203	0	-0.433200378	0.00225029
102	P	0	0	0	0	0
103	E	0.196929626	-0.110420303	0.141875	0	0

104	Q	0.142049789	0.735814629	0.091665649	0.388930664	-0.007799759
105	R	-0.065150528	0.344354973	0.055853271	-0.322033386	-0.093289757
106	A	-0.023055496	0.181020012	-0.004946289	0.07710083	0.013530426
107	L	-0.327924652	0.5555896	-0.49142395	0.631772766	0.051769562
108	G	0.433621521	0	0	-1.139661255	-0.24923027
109	P	0.139458923	0.177988892	0.349064941	0	0
110	V	0.000766068	0.282869949	-0.065505981	-0.59780304	-0.087559776
111	S	-0.356024857	0.327644196	-0.746867676	-0.786213684	-0.079360161
112	M	-0.018551178	0.531048431	0	0.332651978	-0.043170128
113	P	0.138060913	0.222429352	-0.112129517	0	0
114	N	0.170579605	0.155130539	-0.27458374	-0.43047699	-0.093896217
115	L	-0.040685501	0.386004639	-0.40253479	0.512106628	-0.010490341
116	V	-0.192659149	0.484968567	0	0.666882935	0.023319855
117	P	-0.048141174	0.233060608	-0.120014038	0	0
118	E	0.153843918	0.340605621	0.008634644	0.180100403	-0.09395031
119	V	-0.056046143	0.167489624	-0.270927124	0.981543579	0.089780197
120	I	-0.37192009	0.470420837	-0.397103882	1.036470032	0.010319977
121	D	-0.224520264	0.394659348	-0.009046631	1.169908752	0.0430299
122	L	-0.029020004	-0.041295166	0.216907959	2.175137634	0.252670326
123	T	0.811721268	0.080993195	0.352304077	-1.458000793	-0.217040329
124	C	0.323251038	-0.248160172	0	-0.516002197	-0.358729591
125	H	0.220958099	0.800709686	0.573005981	0	0
126	E	-0.035945206	0.187540627	0.081992798	0.003729553	0.000509605
127	A	-0.034190674	0.374845047	-0.378388062	0.045086975	0.031450195
128	G	-0.241470718	0	-0.407732544	-0.010982666	-0.063759575
129	F	0.090570679	-0.234989624	0	-0.443360291	-0.050430374
130	P	0	0	0	0	0
131	P	-0.127308655	0.169988861	-0.01005127	0	0
132	S	-0.255990295	0.398920593	0.116604004	0.309826965	-0.038060188
133	D	0.036685333	0.322605591	-0.041784058	0.761616821	0.010750351
134	D	0.008986664	0.311280518	-0.233435669	-0.222472229	0.002569885
135	E	-0.287120667	0.381399994	0	0.117303772	-0.063600426
136	D	0	0	0	0	0
137	E	-0.317120667	0.6315905	-0.111976318	0	0
138	E	-0.0662883	0.329169884	-0.047199097	0.449226379	0.0096698
139	G	-0.134830322	0	-0.052127686	0.324657288	-0.071539612
140	E	-0.083644943	0.426310501	-0.230626221	0.195592957	0.003410301
141	E	-0.113540955	0.285699959	-0.224261475	0.791361084	0.15450016
142	F	-0.362923279	0.127494736	-0.451997681	0.820571594	0.026250381
143	V	0.070680618	0.386205978	-0.104364624	1.318609009	0.012300224
144	L	-0.287574997	0.513369904	-0.458856812	0.682352295	-0.114979591
145	D	-0.172755737	0.311425247	-1.445491333	1.145852661	-0.084089737
146	Y	0	0	0	0	0

Tabla A3.5. Desplazamientos químicos secundarios de AdE1A-ΔE

#	AA	¹³ C _α Δσ	¹³ C _β Δσ	¹³ C'Δσ	¹⁵ NΔσ	¹ HΔσ
38	F	0	0	0	0	0
39	E	-0.457281647	0.375729218	0	1.172853394	-0.201349602
40	P	0	0	0	0	0
41	P	-0.098829346	0.074869385	-0.024650269	0	0
42	T	0.039091339	0.305316772	0.101109009	-0.02963623	-0.121859818
43	A	0	0	0	0.245782776	0.009419861
44	A	0	0	0	0	0
45	0	0	0	0	0	0
46	A	0	0	0	0	0
47	A	-0.107818756	0.383639526	-0.144694214	0	0
48	D	-0.029284058	0.11625061	-0.266904297	0.063528137	-0.00055954
49	L	-0.216300201	0.483259811	-0.166803589	-0.102737122	-0.142180367
50	D	0.107613373	0.057830582	0.046607056	0.57186676	0.081059914
51	V	-0.40967514	0.205470963	-0.168045044	-1.028263245	-0.144259605
52	T	0.064593735	0.191268768	-0.206176147	-0.002076416	0.013469696
53	A	0.196738892	-0.260549164	0	0.973209229	-0.01634964
54	P	-0.482669373	0.145828247	-0.268691406	0	0
55	E	-0.116634445	0.324500656	-0.160023193	0.390206604	-0.121069756
56	D	-0.066728668	0.795128784	0	0.571418762	0.080620422
57	P	0.070920563	0.316201324	0.225418091	0	0
58	N	0.255590744	0.338410187	0.101439209	-0.545326843	-0.096879616
59	E	-0.000950928	0.289770279	0	0.25550354	-0.237469597
60	E	0.034089203	0.214489822	0.072462769	0	0
61	A	0.020335617	0.17326519	0.178849487	0.335430298	0.002010078
62	V	-0.013499603	0.279879684	0.036957397	-0.269997559	-0.025420303
63	S	-0.075389938	0.190224533	-0.178034058	-0.887816772	-0.130960426
64	Q	-0.227086182	0.204994965	-0.712883301	0.298628235	-0.116040154
65	I	-0.325619812	0.32435463	0.206594849	0.256183167	-0.049170303
66	F	0.023119049	0.214350586	0	0.149320679	0.102370186
67	P	0.443708496	-0.008751831	0.690922241	0	0
68	D	0.537560272	0.029909592	0.143652954	-0.487356873	-0.126499901
69	S	0.589815445	0.047610397	0.328206177	-1.005317993	-0.034420242
70	V	0.594814835	0.376828842	-0.028294678	-0.510210571	-0.188459587
71	M	0.248870392	0.099208603	0.250426636	-1.034659729	-0.120269661
72	L	0.315130615	0.40753952	-0.022836304	-0.002999573	-0.09661602
73	A	0.204398422	0.155614853	0.250930176	-0.044438782	-0.088609657
74	V	0.201795425	0.258734665	0.19236084	-0.325263672	-0.087310181
75	Q	0.124550018	0.221110649	0.02013916	-0.689677124	-0.104379654
76	E	0.115425568	0.422790375	0.031159668	0.038407593	-0.084349899
77	G	0.356520386	0	0.278450928	0.251506958	0.028960075
78	I	0.087070084	0.32496582	-0.120032349	-0.285988464	-0.089270077
79	D	0.287234726	0.141310043	-0.132322388	-0.109189758	0.04994976
80	L	0.206019287	0.249070663	-0.090046387	0.689113464	0.057909889
81	L	0.153603821	-0.002845917	-0.234607544	-1.228240051	-0.096559982
82	T	0.066728821	0.13929184	-0.49117	-0.020357971	-0.227349777
83	F	-0.026148224	-0.06871429	0	-0.945076599	-0.161370316
84	P	0	0	0	0	0
85	P	-0.346589813	0.125818787	-0.312299194	0	0
86	A	0.155461121	-0.372969894	0	0.163859558	-0.00411026
87	P	0.008060608	0.072371063	0.019394531	0	0
88	G	-0.006215439	0	-0.094146118	0.549273682	-0.087159729
89	S	0.663109131	0.05025116	0	0.140161743	-0.039679947
90	P	-0.154851685	0.107768555	-0.243358765	0	0
91	E	0.117530823	0.21866951	0	0.613483276	-0.149589844
92	P	0	0	0	0	0
93	P	0.422299194	0.232520142	0.244361572	0	0
94	H	0.50389595	0.517135124	0.398236084	1.941847229	0.114460449
95	L	-0.056810074	0.221125641	-0.1342187	0.15449707	-0.106990318
96	S	0.130209274	0.075035172	-0.181411743	0.303211975	0.029029579
97	R	-0.378134155	0.462985001	-0.178284302	-0.440498352	-0.145699692
98	Q	0.281609344	-0.166270065	0	0.037123718	-0.009439926
99	P	-0.061558228	0.137569122	-0.083283081	0	0
100	E	0.059030228	0.181924744	-0.183274536	0.078308105	-0.031199532
101	Q	-0.070789185	0.141369858	0	-0.381076355	0.008690453
102	P	0	0	0	0	0
103	E	-0.0019458	-0.094200211	0.129439087	0	0
104	Q	0.500165939	0.679250298	0.075964355	0.476943359	-0.000930443

105	R	0.057324142	0.401110039	0.042150879	-0.297451477	-0.090270424
106	A	-0.016539993	0.181360474	-0.022966919	0.127073364	0.020800285
107	L	-0.31021492	0.544664307	-0.507537231	0.677602539	0.057890244
108	G	0.434479828	0	0	-1.085866394	-0.242000465
109	P	0.136960297	0.203890686	0.336247559	0	0
110	V	0.02822998	0.291540756	-0.054458618	-0.55318634	-0.080920296
111	S	-0.278550262	0.391753998	-0.762050171	-0.726063538	-0.075459633
112	M	-0.021408386	0.521000519	0	0.410548096	-0.040629539
113	P	0.152961121	0.202341156	-0.125084229	0	0
114	N	0.193538361	0.164348755	-0.301774902	-0.40728363	-0.082390137
115	L	0.052139435	0.057891083	-0.892357178	0.55883667	-0.002360268
116	V	-0.193048248	0.483518982	0	0.692403259	0.028130188
117	P	-0.001220398	0.240438232	-0.137286987	0	0
118	E	0.18141655	0.38744915	-0.010637207	0.239457092	-0.090040245
119	V	-0.055060043	0.163814163	-0.281837158	1.012877502	0.09526001
120	I	-0.414314728	0.471693039	-0.415658569	1.080796814	0.017330437
121	D	-0.22065979	0.350874252	-0.022703247	1.193613281	0.047410126
122	L	-0.030086212	-0.01972496	0.203510742	2.212323303	0.259539642
123	T	0.771224442	0.106616516	0.338632202	-1.439697876	-0.209800034
124	C	0.324769287	-0.249760437	0	-0.478000183	-0.353280296
125	H	0.137439117	0.529189072	0.5284198	0	0
126	E	0.150250168	0.065755463	0.080024414	0.047217102	0.008900032
127	A	-0.02868988	0.381474991	-0.394257202	0.094189758	0.039250298
128	G	-0.239015961	0	-0.419481812	0.050296631	-0.05403019
129	F	0.076250305	-0.226318817	0	-0.390976868	-0.045049744
130	P	0	0	0	0	0
131	P	-0.127838898	0.178308716	-0.023646851	0	0
132	S	-0.26087883	0.412745056	0.107738647	0.357083435	-0.030819893
133	D	0.163384781	0.4037957	-0.055150757	0.87	0.025600014
134	D	-0.081785965	0.335410385	-0.245337524	-0.227942505	0.010679932
135	E	-0.281719055	0.385199432	0	0.155336304	-0.057490234
136	D	0	0	0	0	0
137	E	-0.141949768	0.450380936	-0.127235107	0	0
138	E	-0.131630249	0.327809944	-0.064243164	0.428718567	0.018570442
139	G	-0.12187561	0	-0.058734741	0.370456543	-0.064479561
140	E	-0.080774384	0.274680099	-0.197987671	0.244787292	0.010680161
141	E	-0.030824966	0.243300552	-0.236575317	0.816942444	0.161190186
142	F	-0.363711014	0.123346252	-0.465654297	0.844146423	0.031300087
143	V	0.054269791	0.390020676	-0.114267578	1.344831238	0.01570961
144	L	-0.450704803	0.600829468	-0.47387146	0.716539612	-0.108950462
145	D	-0.256904144	0.33047966	-1.461070557	1.197869873	-0.075999718
146	Y	0	0	0	0	0

Tabla A3.6. Desplazamientos químicos secundarios de AdE1A-ΔL

#	AA	$^{13}\text{C}_\alpha\Delta\sigma$	$^{13}\text{C}_\beta\Delta\sigma$	$^{13}\text{C}'\Delta\sigma$	$^{15}\text{N}\Delta\sigma$	$^1\text{H}\Delta\sigma$
38	F	0	0	0	0	0
39	E	-0.440310059	0.369419708	0	1.050836487	-0.20430027
40	P	0	0	0	0	0
41	P	-0.077459412	0.086801758	-0.172996216	0	0
42	T	0.037624588	0.331012878	0.038239746	0.381550598	-0.062210083
43	L	0.495094147	0.075189896	-0.193317871	0.115083618	0.087480392
44	H	0.288566132	0.586082802	0.172511597	-0.094621582	-0.09987011
45	E	0.167386017	0.251344604	-0.027369995	-0.297648315	0.001889839
46	L	0.023600311	0.305355682	-0.16831604	-0.446219177	0.021630363
47	Y	-0.304954529	0.223311005	0.114356689	-0.742146912	-0.142090206
48	D	0.375150299	0.09263916	-0.651529541	-0.918753662	0.138079567
49	L	-0.085183334	0.490429535	-0.165323486	-0.085792236	-0.173129959
50	D	0.162905502	0.0484655	0.058661499	0.230703125	0.094020348
51	V	-0.524520416	0.190679474	-0.204940796	-1.273792419	-0.181520138
52	T	0.050069275	0.183612671	-0.223555908	-0.170419006	-0.011969566
53	A	0.196628265	-0.259240723	0	0.862903442	-0.032150116
54	P	-0.218898315	0.19303894	-0.323226318	0	0
55	E	-0.067958908	0.403520203	-0.187931519	0.327546387	-0.131068077
56	D	-0.075029449	0.817250214	0	0.53238678	0.075310364
57	P	0.074090576	0.322980042	0.223175049	0	0
58	N	0.254860229	0.344309616	0.098646851	-0.608193054	-0.104280128
59	E	-0.003018494	0.290989075	0	0.193873291	-0.24487011
60	E	0.038869019	0.20652092	0.073729248	0	0
61	A	0.026069107	0.173474998	0.182297974	0.274807129	-0.005250244
62	V	-0.007399902	0.228400345	0.025070801	-0.337128601	-0.033809776
63	S	0.006450577	0.193680649	-0.22046875	-0.95223175	-0.137330017
64	Q	-0.237935181	0.224459457	0.13149231	0.214697266	-0.125369949
65	I	-0.328595276	0.311609726	-0.296243286	0.251788635	-0.060700226
66	F	0.018961029	0.129488831	0	0.06474884	0.09343998
67	P	0.440118866	-0.009308777	0.680881958	0	0
68	D	0.533625412	0.030624847	0.1466437	-0.543112488	-0.133720169
69	S	0.596418686	0.048239822	0.313374634	-1.059707947	-0.040849915
70	V	0.610310135	0.246774368	-0.030842896	-0.621943054	-0.197059822
71	M	-0.037294846	0.232034454	0.260848389	-1.103362427	-0.128609543
72	L	0.384664917	0.39742485	-0.022576904	-0.054803162	-0.103920212
73	A	0.20401886	0.157444954	0.250548706	-0.10217804	-0.096910439
74	V	0.200290527	0.27177521	0.192299805	-0.37972229	-0.093690262
75	Q	0.108934555	0.223234482	0.06638855	-0.784701233	-0.115919113
76	E	0.12015007	0.334750023	0.030244141	-0.021773071	-0.091120033
77	G	0.35535881	0	0.276848755	0.197277222	0.022020187
78	I	0.079000092	0.320849762	-0.122748413	-0.338158264	-0.096189938
79	D	0.342244568	0.117218323	-0.131498413	-0.162442932	0.043649788
80	L	0.113175278	0.311543961	-0.092854004	0.579311218	0.056819839
81	L	0.163419037	-0.047649536	-0.234775391	-1.287123718	-0.103570442
82	T	0.067505112	0.138013916	-0.490452881	-0.078623657	-0.234449883
83	F	-0.156649017	-0.040870819	0	-1.048828735	-0.162834206
84	P	0	0	0	0	0
85	P	-0.351960907	0.132059631	-0.293378296	0	0
86	A	0.154919434	-0.374409943	0	0.124209595	-0.012593193
87	P	0.06568161	0.111940918	0.017456665	0	0
88	G	-0.003075943	0	-0.094329224	0.500163269	-0.093650436
89	S	0.661480255	0.033939514	0	0.086733093	-0.046280327
90	P	-0.161039124	0.085929413	-0.247188721	0	0
91	E	0.106658936	0.229940033	0	0.534946289	-0.159320183
92	P	0	0	0	0	0
93	P	0.422829437	0.234839478	0.243339233	0	0
94	H	0.539120865	0.462480049	0.336071777	1.891836548	0.108259659
95	L	-0.1403405	0.238215485	-0.130632935	0.137582703	-0.099199753
96	S	0.153339691	0.08475502	-0.18951416	0.274479675	0.028919907
97	R	-0.237854385	0.49125	-0.176804199	-0.488960266	-0.148880196
98	Q	0.228470612	-0.162730026	0	-0.035256348	-0.022449951
99	P	-0.05818985	0.120181732	-0.085327759	0	0
100	E	0.107799225	-0.004175262	-0.184632568	0.028556824	-0.037050323
101	Q	-0.074191895	0.141930618	0	-0.439891357	0
102	P	0	0	0	0	0
103	E	0.123439484	0.119560165	0.128248901	0	0
104	Q	0.665584564	-0.020929756	0.072348022	0.398627625	-0.006160393

105	R	-0.029359131	0.36144577	0.041448975	-0.327862244	-0.091140175
106	A	0.315414963	0.04611515	-0.022509155	0.078832703	0.014039688
107	L	-0.32542984	0.548736496	-0.508895264	0.693639526	0.054649658
108	G	1.664239502	0	0	-1.138287964	-0.248650436
109	P	0.134869843	0.077639465	0.333928223	0	0
110	V	0.061845093	0.265869751	-0.065795898	-0.602403564	-0.087329941
111	S	-0.321660156	0.345578995	-0.763896484	-0.76105957	-0.079019699
112	M	-0.037838287	0.513111725	0	0.34396637	-0.042349968
113	P	0.112990723	0.228971558	-0.126594849	0	0
114	N	0.176904373	0.152649078	-0.294099731	-0.456622925	-0.089729614
115	L	-0.091998901	0.456061554	-0.423057861	0.510161133	-0.011856956
116	V	-0.180730591	0.475630188	0	0.614667358	0.010094299
117	P	0	0	0	0	0
118	E	0.277335205	0.33318985	0.046903687	0	0
119	V	-0.041828766	0.168349838	-0.164573364	0.9372626	0.076070175
120	I	-0.282085876	0.320505142	0.010580444	1.0277878	0.038609772
121	D	-0.01734024	0.300430298	0.063588257	0.7713736	-0.020389709
122	A	0.587960052	0.077080307	0	1.4950198	0.170720406
123	A	0	0	0	0	0
124	A	0	0	0	0	0
125		0	0	0	0	0
126	A	0.297809296	0.345330734	0.032407837	0	0
127	A	-0.14984108	0.300179749	-0.472757568	-0.768962402	-0.251239777
128	G	-0.202154541	0	-0.387178955	-0.430347595	-0.210630188
129	F	0.094278564	-0.358250122	0	-0.528893433	-0.05961998
130	P	0	0	0	0	0
131	P	-0.131020355	0.191389313	-0.017940063	0	0
132	S	-0.258248596	0.377398071	0.101085815	0.323651428	-0.030870438
133	D	0.054999695	0.2926297	-0.061681519	0.82223999	0.013410149
134	D	-0.029530334	0.390874176	-0.249793091	-0.236113586	0.005400391
135	E	-0.291580048	0.379729156	0	0.109743042	-0.061300163
136	D	0	0	0	0	0
137	E	-0.291791077	0.343519821	-0.304206543	0	0
138	E	-0.101518936	0.381024971	-0.077014771	0.30381012	0.014540215
139	G	-0.122535553	0	-0.062366333	0.320461121	-0.071079941
140	E	-0.113994675	0.246784172	-0.054143066	0.189947205	0.003410301
141	E	0.051349335	0.383275146	-0.243212891	0.910402527	0.146820221
142	F	-0.366736069	0.12647049	-0.466310425	0.805358582	0.025060196
143	V	0.055673599	0.396280594	-0.121378174	1.304197083	0.011130066
144	L	-0.547624817	0.093196259	-0.475275269	0.673776855	-0.115250435
145	D	-0.176764984	0.337279358	-1.4627948	1.144502258	-0.083230476
146	Y	0	0	0	0	0

Figura A3.1. Desplazamientos químicos secundarios por residuo en AdE1A-ΔE

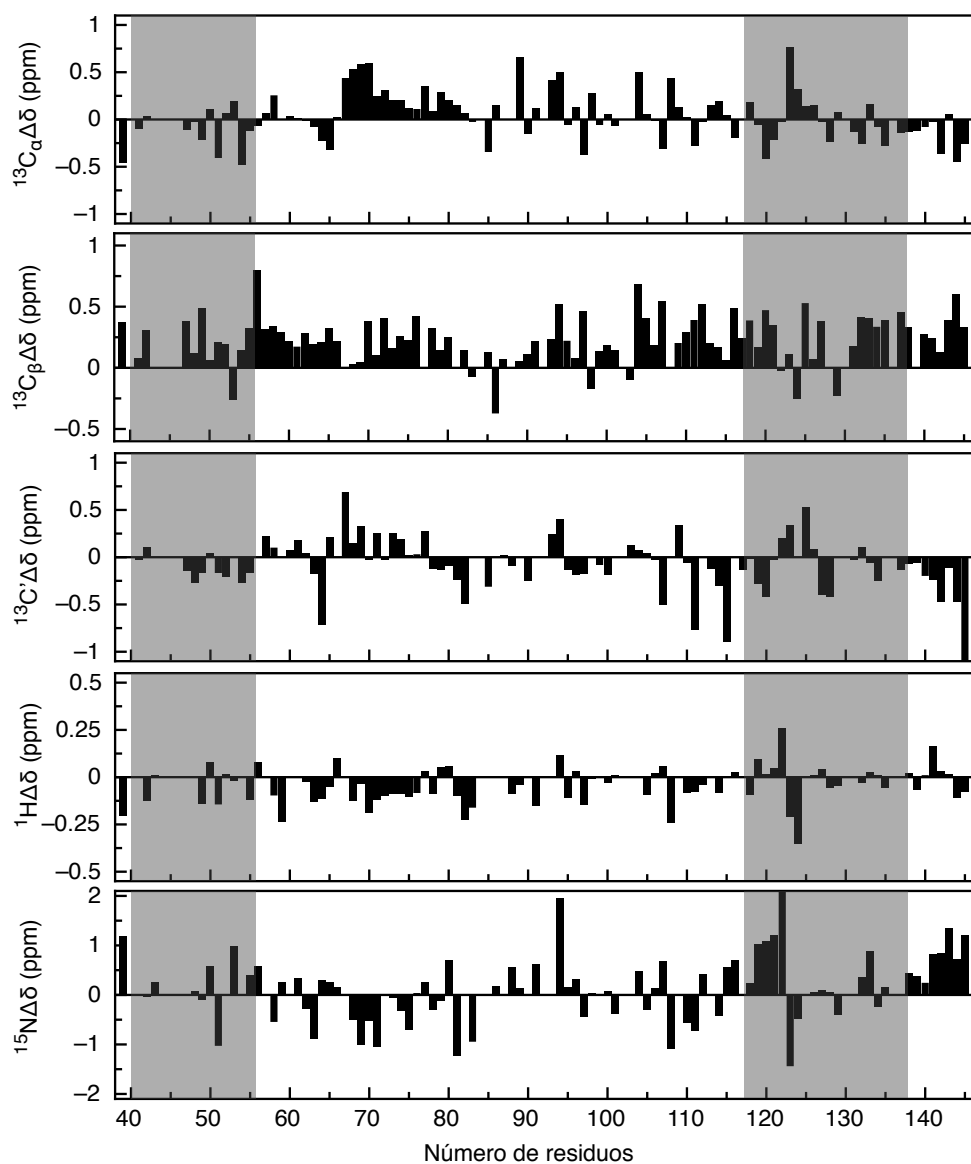


Figura A3.1: Cambio en los desplazamientos químicos de $^{13}\text{C}_\alpha$, $^{13}\text{C}_\beta$, $^{13}\text{C}'$, ^1H y ^{15}N por residuo para la proteína AdE1A-ΔE. Las regiones correspondientes a los motivos de interacciones y residuos flanqueantes están señalados en gris oscuro

Figura A3.2. Desplazamientos químicos secundarios por residuo en AdE1A- Δ L

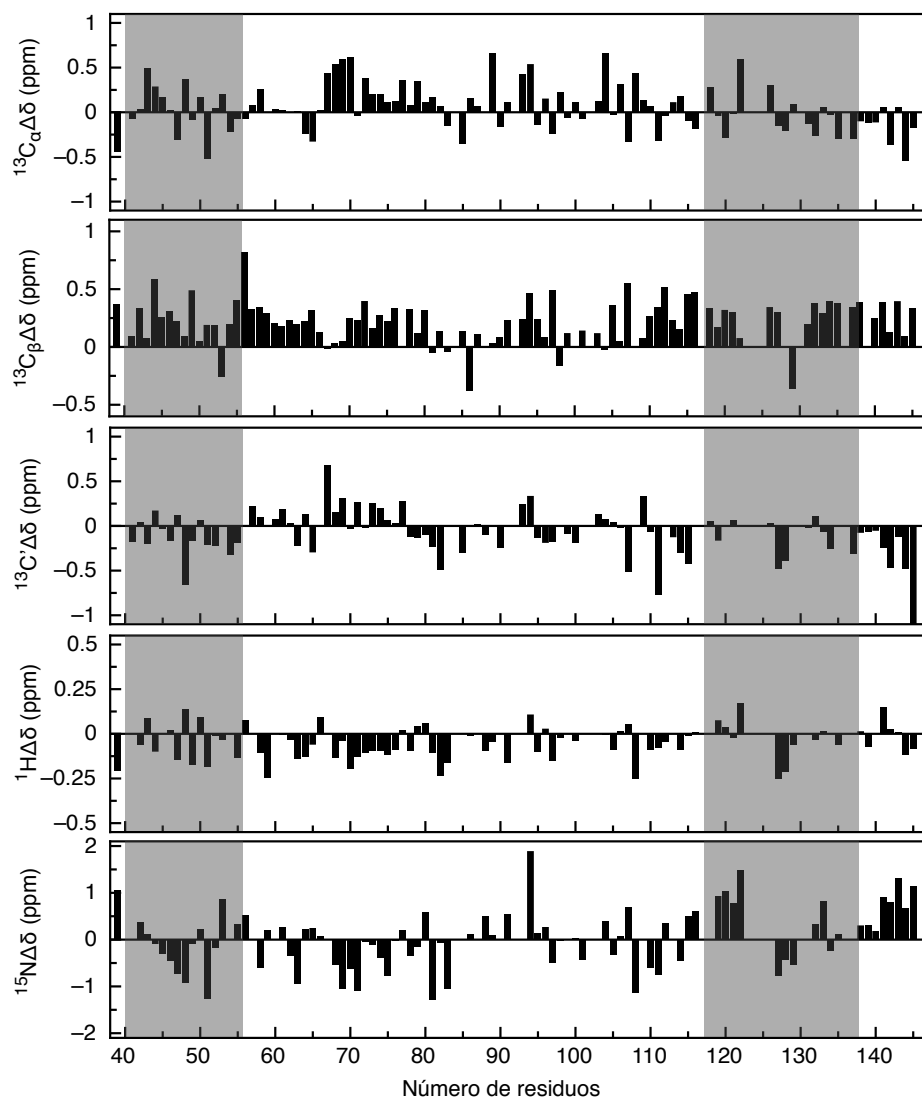


Figura A3.2: Cambio en los desplazamientos químicos de $^{13}\text{C}_\alpha$, $^{13}\text{C}_\beta$, $^{13}\text{C}'$, ^1H y ^{15}N por residuo para la proteína AdE1A- Δ L. Las regiones correspondientes a los motivos de interacciones y residuos flanqueantes están señalados en gris oscuro

Tabla A3.7. Probabilidad de estructura secundaria por residuo en AdE1A.WT

#	AA	α -Hélice	Lámina β	Coil	PPII
38	F				
39	E	0	0.105	0.767	0.129
40	P				
41	P	0.005	0.162	0.565	0.268
42	T	0.012	0.108	0.602	0.278
43	L	0.018	0.055	0.665	0.262
44	H	0.016	0.026	0.711	0.247
45	E	0.016	0.023	0.721	0.24
46	L	0.019	0.031	0.722	0.229
47	Y	0.021	0.039	0.72	0.22
48	D	0.025	0.041	0.71	0.224
49	L	0.034	0.033	0.714	0.218
50	D	0.028	0.045	0.702	0.225
51	V	0.017	0.064	0.723	0.196
52	T	0.006	0.102	0.699	0.192
53	A	0.002	0.131	0.685	0.182
54	P	0	0.168	0.623	0.209
55	E	0.002	0.151	0.636	0.212
56	D	0.006	0.14	0.668	0.187
57	P	0.01	0.152	0.658	0.18
58	N	0.008	0.099	0.735	0.158
59	E	0.005	0.04	0.817	0.137
60	E	0.005	0.011	0.84	0.145
61	A	0.006	0.023	0.775	0.196
62	V	0.005	0.05	0.7	0.245
63	S	0.005	0.08	0.639	0.276
64	Q	0.006	0.099	0.638	0.257
65	I	0.006	0.133	0.649	0.212
66	F	0.005	0.157	0.656	0.181
67	P	0.009	0.171	0.617	0.203
68	D	0.014	0.114	0.636	0.236
69	S	0.013	0.057	0.675	0.255
70	V	0.013	0.027	0.694	0.267
71	M	0.018	0.026	0.749	0.206
72	L	0.024	0.03	0.785	0.161
73	A	0.02	0.031	0.806	0.144
74	V	0.015	0.03	0.775	0.18
75	Q	0.017	0.029	0.699	0.255
76	E	0.03	0.031	0.651	0.288
77	G	0.041	0.036	0.65	0.273
78	I	0.035	0.037	0.701	0.227
79	D	0.032	0.032	0.754	0.183
80	L	0.038	0.039	0.763	0.16
81	L	0.038	0.061	0.766	0.134
82	T	0.027	0.094	0.753	0.125
83	F	0.012	0.093	0.769	0.125
84	P				
85	P	0	0.115	0.655	0.229
86	A	0.002	0.107	0.671	0.22
87	P	0.004	0.132	0.628	0.235
88	G	0.007	0.18	0.583	0.229
89	S	0.005	0.195	0.602	0.198
90	P	0.002	0.229	0.582	0.187
91	E	0.002	0.223	0.619	0.156
92	P				
93	P	0.009	0.229	0.574	0.187
94	H	0.01	0.189	0.581	0.22
95	L	0.008	0.114	0.637	0.241
96	S	0.005	0.067	0.668	0.26
97	R	0.003	0.072	0.676	0.249
98	Q	0.002	0.122	0.635	0.241
99	P	0.001	0.187	0.553	0.259
100	E	0	0.165	0.558	0.277
101	Q	0	0.105	0.673	0.222
102	P				
103	E	0.002	0.041	0.811	0.145

104	Q	0.006	0.057	0.733	0.204
105	R	0.012	0.066	0.688	0.234
106	A	0.014	0.099	0.631	0.256
107	L	0.013	0.166	0.575	0.246
108	G	0.008	0.2	0.535	0.258
109	P	0.005	0.177	0.541	0.277
110	V	0.004	0.117	0.596	0.282
111	S	0.002	0.12	0.619	0.258
112	M	0.001	0.152	0.624	0.222
113	P	0.005	0.18	0.586	0.229
114	N	0.012	0.126	0.638	0.224
115	L	0.017	0.105	0.665	0.212
116	V	0.01	0.105	0.677	0.208
117	P	0.004	0.142	0.605	0.248
118	E	0.001	0.125	0.603	0.271
119	V	0.001	0.116	0.614	0.269
120	I	0.004	0.097	0.634	0.265
121	D	0.011	0.082	0.644	0.262
122	L	0.019	0.051	0.686	0.244
123	T	0.017	0.034	0.766	0.183
124	C	0.008	0.022	0.845	0.124
125	H	0.004	0.023	0.822	0.15
126	E	0.009	0.047	0.7	0.244
127	A	0.018	0.075	0.627	0.28
128	G	0.023	0.112	0.629	0.235
129	F	0.016	0.12	0.71	0.154
130	P				
131	P	0.001	0.228	0.572	0.199
132	S	0.002	0.18	0.602	0.215
133	D	0.004	0.106	0.666	0.224
134	D	0.005	0.054	0.713	0.227
135	E	0.003	0.032	0.781	0.185
136	D				
137	E	0.007	0.027	0.774	0.192
138	E	0.014	0.036	0.679	0.271
139	G	0.016	0.038	0.631	0.316
140	E	0.014	0.047	0.615	0.324
141	E	0.014	0.062	0.613	0.311
142	F	0.017	0.075	0.612	0.296
143	V	0.01	0.081	0.612	0.297
144	L	0.005	0.083	0.623	0.289
145	D	0.001	0.111	0.605	0.282
146	Y				

Tabla A3.8. Probabilidad de estructura secundaria por residuo en AdE1A-ΔE

#	AA	α-Hélice	Lámina β	Coil	PPII
38	F				
39	E	0	0.109	0.772	0.119
40	P				
41	P	0.002	0.123	0.6	0.275
42	T	0.002	0.095	0.609	0.294
43	A				
44	A				
45					
46	A				
47	A	0.011	0.009	0.861	0.12
48	D	0.02	0.015	0.779	0.186
49	L	0.028	0.022	0.725	0.225
50	D	0.023	0.044	0.689	0.244
51	V	0.014	0.068	0.712	0.207
52	T	0.005	0.106	0.694	0.195
53	A	0.002	0.138	0.68	0.181
54	P	0	0.18	0.61	0.21
55	E	0.002	0.164	0.613	0.22
56	D	0.006	0.152	0.639	0.204
57	P	0.01	0.159	0.64	0.192
58	N	0.008	0.101	0.727	0.164
59	E	0.005	0.041	0.815	0.139
60	E	0.004	0.011	0.837	0.147
61	A	0.006	0.025	0.77	0.2
62	V	0.004	0.054	0.691	0.251
63	S	0.003	0.087	0.627	0.283
64	Q	0.004	0.109	0.623	0.264
65	I	0.004	0.14	0.639	0.217
66	F	0.005	0.165	0.646	0.185
67	P	0.009	0.176	0.61	0.206
68	D	0.016	0.116	0.654	0.215
69	S	0.018	0.054	0.74	0.189
70	V	0.022	0.021	0.8	0.157
71	M	0.026	0.022	0.816	0.136
72	L	0.03	0.029	0.807	0.134
73	A	0.022	0.031	0.805	0.142
74	V	0.016	0.031	0.773	0.18
75	Q	0.016	0.031	0.696	0.256
76	E	0.029	0.034	0.648	0.29
77	G	0.039	0.039	0.646	0.275
78	I	0.035	0.039	0.697	0.23
79	D	0.033	0.033	0.749	0.186
80	L	0.044	0.024	0.779	0.152
81	L	0.051	0.027	0.82	0.103
82	T	0.048	0.044	0.844	0.065
83	F	0.027	0.082	0.82	0.07
84	P				
85	P	0	0.126	0.648	0.226
86	A	0.002	0.112	0.666	0.221
87	P	0.004	0.14	0.62	0.236
88	G	0.007	0.187	0.576	0.23
89	S	0.004	0.201	0.595	0.199
90	P	0.002	0.234	0.577	0.187
91	E	0.001	0.228	0.614	0.156
92	P				
93	P	0.008	0.238	0.562	0.192
94	H	0.007	0.201	0.563	0.229
95	L	0.005	0.126	0.617	0.252
96	S	0.003	0.073	0.657	0.267
97	R	0.003	0.075	0.67	0.252
98	Q	0.001	0.12	0.647	0.232
99	P	0.001	0.176	0.595	0.228
100	E	0.001	0.141	0.639	0.219
101	Q	0.001	0.079	0.745	0.175
102	P				
103	E	0.003	0.029	0.824	0.145
104	Q	0.007	0.043	0.753	0.197

105	R	0.013	0.059	0.703	0.226
106	A	0.014	0.098	0.636	0.252
107	L	0.013	0.167	0.574	0.246
108	G	0.008	0.202	0.532	0.259
109	P	0.005	0.179	0.54	0.277
110	V	0.004	0.119	0.596	0.281
111	S	0.002	0.121	0.62	0.257
112	M	0.001	0.153	0.625	0.221
113	P	0.004	0.18	0.588	0.227
114	N	0.01	0.125	0.644	0.221
115	L	0.014	0.105	0.674	0.207
116	V	0.008	0.106	0.681	0.205
117	P	0.003	0.145	0.605	0.247
118	E	0.001	0.136	0.593	0.269
119	V	0.001	0.144	0.591	0.264
120	I	0.004	0.14	0.599	0.257
121	D	0.01	0.108	0.624	0.257
122	L	0.018	0.061	0.678	0.243
123	T	0.015	0.035	0.765	0.185
124	C	0.008	0.021	0.846	0.125
125	H	0.004	0.02	0.826	0.15
126	E	0.009	0.041	0.705	0.244
127	A	0.018	0.076	0.625	0.281
128	G	0.023	0.124	0.617	0.237
129	F	0.015	0.142	0.687	0.155
130	P				
131	P	0.001	0.236	0.563	0.2
132	S	0.003	0.185	0.595	0.218
133	D	0.005	0.109	0.663	0.223
134	D	0.006	0.051	0.724	0.219
135	E	0.005	0.021	0.809	0.165
136	D				
137	E	0.007	0.024	0.779	0.189
138	E	0.013	0.036	0.677	0.274
139	G	0.015	0.037	0.63	0.318
140	E	0.014	0.043	0.617	0.326
141	E	0.014	0.059	0.616	0.311
142	F	0.016	0.074	0.612	0.298
143	V	0.01	0.083	0.606	0.301
144	L	0.004	0.087	0.614	0.295
145	D	0.001	0.117	0.595	0.287
146	Y				

Tabla A3.9. Probabilidad de estructura secundaria por residuo en AdE1A- Δ L

#	AA	α -Hélice	Lámina β	Coil	PPII
38	F				
39	E	0	0.107	0.764	0.128
40	P				
41	P	0.003	0.161	0.589	0.247
42	T	0.005	0.104	0.679	0.212
43	L	0.007	0.044	0.821	0.128
44	H	0.005	0.01	0.915	0.07
45	E	0.006	0.007	0.922	0.065
46	L	0.007	0.019	0.872	0.102
47	Y	0.014	0.034	0.788	0.164
48	D	0.023	0.039	0.728	0.21
49	L	0.035	0.034	0.713	0.218
50	D	0.027	0.046	0.701	0.226
51	V	0.016	0.066	0.723	0.196
52	T	0.005	0.102	0.7	0.193
53	A	0.002	0.132	0.684	0.183
54	P	0	0.173	0.614	0.213
55	E	0.002	0.163	0.615	0.221
56	D	0.006	0.153	0.639	0.203
57	P	0.01	0.16	0.639	0.191
58	N	0.008	0.101	0.728	0.163
59	E	0.005	0.041	0.816	0.138
60	E	0.004	0.011	0.838	0.146
61	A	0.006	0.024	0.772	0.198
62	V	0.005	0.051	0.697	0.248
63	S	0.005	0.08	0.637	0.279
64	Q	0.005	0.097	0.636	0.262
65	I	0.004	0.131	0.647	0.219
66	F	0.03	0.153	0.651	0.166
67	P	0.083	0.158	0.613	0.147
68	D	0.136	0.088	0.657	0.119
69	S	0.092	0.037	0.741	0.13
70	V	0.049	0.015	0.798	0.138
71	M	0.027	0.023	0.811	0.14
72	L	0.033	0.027	0.807	0.133
73	A	0.027	0.038	0.808	0.128
74	V	0.023	0.056	0.773	0.148
75	Q	0.02	0.079	0.684	0.217
76	E	0.031	0.064	0.638	0.267
77	G	0.04	0.049	0.644	0.267
78	I	0.034	0.038	0.7	0.228
79	D	0.031	0.032	0.753	0.184
80	L	0.038	0.038	0.762	0.161
81	L	0.038	0.061	0.767	0.134
82	T	0.027	0.094	0.753	0.126
83	F	0.012	0.093	0.769	0.126
84	P				
85	P	0	0.116	0.654	0.23
86	A	0.002	0.11	0.668	0.221
87	P	0.004	0.136	0.623	0.236
88	G	0.007	0.183	0.58	0.23
89	S	0.005	0.197	0.599	0.199
90	P	0.002	0.231	0.58	0.187
91	E	0.001	0.226	0.616	0.156
92	P				
93	P	0.008	0.235	0.565	0.191
94	H	0.008	0.195	0.571	0.226
95	L	0.007	0.117	0.63	0.246
96	S	0.005	0.068	0.665	0.262
97	R	0.003	0.073	0.674	0.251
98	Q	0.002	0.118	0.648	0.232
99	P	0.001	0.173	0.596	0.23
100	E	0.001	0.139	0.636	0.224
101	Q	0.001	0.085	0.731	0.183
102	P				
103	E	0.003	0.016	0.844	0.137
104	Q	0.008	0.023	0.779	0.19

105	R	0.015	0.046	0.714	0.225
106	A	0.018	0.091	0.639	0.252
107	L	0.015	0.164	0.574	0.247
108	G	0.008	0.198	0.534	0.259
109	P	0.005	0.174	0.542	0.279
110	V	0.004	0.116	0.598	0.282
111	S	0.003	0.119	0.621	0.257
112	M	0.001	0.153	0.624	0.221
113	P	0.005	0.182	0.585	0.229
114	N	0.011	0.128	0.636	0.225
115	L	0.017	0.092	0.686	0.205
116	V	0.012	0.048	0.788	0.152
117	P				
118	E	0.001	0.053	0.805	0.141
119	V	0.002	0.086	0.712	0.2
120	I	0.005	0.084	0.69	0.221
121	D	0.012	0.058	0.72	0.211
122	A	0.024	0.028	0.808	0.14
123	A				
124	A				
125					
126	A	0.026	0.021	0.796	0.157
127	A	0.037	0.047	0.701	0.215
128	G	0.039	0.077	0.679	0.204
129	F	0.025	0.098	0.726	0.152
130	P				
131	P	0.001	0.231	0.571	0.197
132	S	0.003	0.182	0.602	0.214
133	D	0.005	0.105	0.671	0.219
134	D	0.006	0.049	0.728	0.217
135	E	0.005	0.021	0.809	0.165
136	D				
137	E	0.007	0.025	0.775	0.193
138	E	0.014	0.036	0.676	0.274
139	G	0.016	0.036	0.633	0.315
140	E	0.014	0.043	0.623	0.32
141	E	0.015	0.06	0.62	0.306
142	F	0.017	0.075	0.615	0.293
143	V	0.01	0.082	0.614	0.295
144	L	0.004	0.084	0.626	0.285
145	D	0.001	0.114	0.607	0.278
146	Y				

Tabla A3.10. Parámetros de relajación por residuo de AdE1A.WT

#	AA	R1	Error R1	R2	Error R2	R1ρ	Error R1ρ	NOE	Error NOE
38	F								
39	E	0.93	0.03	1.32	0.085	1.81	0.07	-0.826	0.087
40	P								
41	P								
42	T	1.078	0.023	1.61	0.075	2.25	0.11	-0.335	0.018
43	L	1.09	0.025	1.72	0.13	2.63	0.065	-0.180	0.014
44	H	1.17	0.035	2.34	0.1	2.75	0.125	-0.237	0.014
45	E	0.9	0.035	1.54	0.08	2.06	0.09	-0.194	0.007
46	L	1.036	0.0235	1.42	0.055	2.14	0.06	-0.180	0.030
47	Y	0.97	0.03	1.20	0.065	1.76	0.065	-0.322	0.003
48	D	1.11	0.025	1.58	0.06	2.13	0.06	-0.153	0.018
49	L	1	0.03	1.12	0.065	1.78	0.085	-0.236	0.011
50	D	1.07	0.025	1.50	0.055	2.09	0.075	-0.303	0.027
51	V	0.95	0.035	1.20	0.06	1.76	0.06	-0.235	0.001
52	T	1.006	0.0235	1.58	0.075	2.01	0.065	-0.501	0.013
53	A	0.97	0.025	1.52	0.055	1.8	0.075	-0.440	0.004
54	P								
55	E	0.95	0.025	1.35	0.085	2.1	0.085	-0.324	0.027
56	D	0.977	0.023	1.50	0.04	2.24	0.06	-0.329	0.008
57	P								
58	N	1.042	0.0205	1.96	0.05	2.62	0.085	-0.286	0.024
59	E	1.05	0.025	1.63	0.05	2.32	0.06	-0.153	0.013
60	E								
61	A	0.937	0.0215	1.36	0.075	1.93	0.08	-0.490	0.002
62	V	0.88	0.03	1.15	0.075	1.68	0.08	-0.409	0.018
63	S	0.878	0.0165	1.88	0.065	2.48	0.075	-0.732	0.076
64	Q	0.97	0.018	1.89	0.08	2.46	0.07	-0.625	0.044
65	I	0.97	0.03	1.25	0.05	1.85	0.08	-0.374	0.025
66	F	1	0.025	1.57	0.095	2.37	0.08	-0.297	0.034
67	P								
68	D	1.107	0.0205	2.13	0.12	2.67	0.075	-0.177	0.003
69	S	1.067	0.0165	2.20	0.065	2.84	0.06	-0.093	0.009
70	V	1	0.025	1.25	0.06	1.85	0.075	-0.289	0.016
71	M	1.1	0.0215	1.92	0.08	2.76	0.07	-0.291	0.046
72	L	0.94	0.025	1.35	0.055	2.22	0.045	-0.307	0.001
73	A	0.99	0.03	1.78	0.08	2.47	0.08	-0.264	0.075
74	V	0.9	0.035	1.34	0.08	1.84	0.095	-0.276	0.032
75	Q	1.015	0.014	1.84	0.09	2.59	0.05	-0.457	0.005
76	E	1.048	0.023	1.74	0.08	2.56	0.065	-0.247	0.008
77	G	0.99	0.021	1.63	0.105	1.72	0.065	-0.302	0.018
78	I	1.01	0.035	1.37	0.06	1.94	0.07	-0.119	0.016
79	D	1.166	0.0185	1.73	0.08	2.58	0.075	-0.245	0.005
80	L	0.99	0.03	1.46	0.065	2.21	0.075	-0.138	0.009
81	L	1.06	0.03	1.60	0.04	2.23	0.055	-0.085	0.035
82	T	1.02	0.025	1.65	0.055	2.27	0.075	-0.211	0.049
83	F	0.99	0.03	1.46	0.06	2.21	0.08	-0.189	0.014
84	P								
85	P								
86	A	1	0.0245	1.46	0.08	1.99	0.08	-0.245	0.052
87	P								
88	G	0.818	0.0195	1.68	0.085	2.01	0.11	-0.482	0.124
89	S	0.797	0.0235	1.95	0.095	2.51	0.09	-0.235	0.006
90	P								
91	E	0.97	0.025	1.50	0.045	2.24	0.065	-0.318	0.023
92	P								
93	P								
94	H	1.087	0.0215	1.63	0.07	2.37	0.09	-0.317	0.031
95	L	1.05	0.04	2.60	0.18	3.3	0.18	-0.268	0.033
96	S	1.07	0.065	2.48	0.13	2.04	0.14	-0.985	0.058
97	R	1.08	0.03	2.15	0.12	2.68	0.075	-0.356	0.103
98	Q	0.999	0.0235	1.53	0.05	2.29	0.045	-0.323	0.021
99	P								
100	E	1.051	0.023	1.70	0.085	2.53	0.09	-0.410	0.048
101	Q	0.89	0.03	1.00	0.05	1.6	0.05	-0.374	0.010
102	P								
103	E								

104	Q	0.93	0.025	1.16	0.07	1.67	0.075	-0.533	0.014
105	R	0.97	0.0165	1.89	0.07	2.46	0.07	-0.642	0.035
106	A	0.954	0.02	1.76	0.05	2.28	0.09	-0.765	0.060
107	L	1.016	0.011	1.57	0.065	2.21	0.07	-0.246	0.048
108	G	0.798	0.0225	1.22	0.06	1.47	0.09	-1.399	0.003
109	P								
110	V	0.82	0.03	1.30	0.075	1.79	0.08	-1.223	0.019
111	S	0.871	0.019	1.88	0.065	2.48	0.07	-0.905	0.087
112	M	0.96	0.02	1.90	0.055	2.79	0.05	-0.768	0.006
113	P								
114	N	0.93	0.03	2.37	0.14	2.5	0.095	-0.693	0.095
115	L	0.957	0.019	1.41	0.05	2.18	0.045	-0.496	0.007
116	V	0.99	0.035	1.46	0.06	2.21	0.075	-0.265	0.026
117	P								
118	E	0.95	0.025	1.35	0.08	2.1	0.08	-0.279	0.006
119	V	0.93	0.03	1.45	0.04	2.22	0.05	-0.222	0.039
120	I	0.96	0.025	1.64	0.055	2.34	0.07	-0.263	0.047
121	D	1.081	0.0235	2.18	0.075	3.11	0.07	-0.112	0.039
122	L	1.26	0.035	1.88	0.14	3.01	0.045	-0.150	0.028
123	T	1.26	0.035	2.07	0.125	2.94	0.16	-0.017	0.028
124	C	1.73	0.18	3.04	0.205	4.3	0.3	-0.661	0.151
125	H								
126	E	0.9	0.1	1.29	0.3	2.86	0.085	-0.726	0.068
127	A	0.91	0.0215	1.99	0.12	2.55	0.16	-0.554	0.041
128	G	0.828	0.0235	2.03	0.06	2.18	0.09	-0.621	0.003
129	F	0.9	0.025	1.27	0.09	2.15	0.075	-0.451	0.116
130	P								
131	P								
132	S	0.942	0.021	1.74	0.065	2.42	0.095	-0.659	0.093
133	D	0.971	0.022	1.52	0.03	2.18	0.055	-0.443	0.024
134	D	0.9	0.035	1.19	0.05	1.83	0.06	-0.477	0.009
135	E	0.91	0.035	0.95	0.05	1.55	0.065	-0.587	0.031
136	D								
137	E								
138	E					1.47	0.095		
139	G	0.872	0.022	0.98	0.07	1.24	0.06	-0.946	0.007
140	E	0.92	0.03	0.91	0.05	1.45	0.07	-0.644	0.003
141	E	0.934	0.0215	0.93	0.08	1.46	0.085	-0.720	0.018
142	F	0.936	0.0245	0.89	0.08	1.53	0.08	-0.954	0.000
143	V	0.83	0.035	0.64	0.06	1.11	0.085	-0.948	0.030
144	L	0.84	0.035	0.51	0.08	0.87	0.08	-1.322	0.038
145	D	0.704	0.024	0.56	0.08	1.01	0.075	-1.859	0.009
146	Y	0.52	0.035	0.11	0.04	0.43	0.065	-2.386	0.053

Tabla A3.11. Parámetros de relajación por residuo de AdE1A-ΔE

#	AA	R1	Error R1	R2	Error R2	R1ρ	Error R1ρ	NOE	Error NOE
38	F								
39	E	0.92	0.03	0.75	0.16	1.82	0.065	-1.063	0.027
40	P								
41	P								
42	T	0.968	0.0205	0.94	0.225	2.17	0.095	-1.009	0.062
43	A	0.89	0.025	1.14	0.22	2.14	0.08	-0.944	0.010
44	A								
45									
46	A								
47	A								
48	D	0.9	0.03	1	0.19	1.81	0.06	-0.881	0.018
49	L	0.96	0.035	1.12	0.135	1.99	0.03	-0.601	0.024
50	D	0.99	0.03	0.94	0.175	1.89	0.045	-0.670	0.026
51	V	0.92	0.03	0.86	0.165	1.74	0.055	-0.591	0.026
52	T	0.96	0.03	1.2	0.2	2.12	0.065	-0.741	0.019
53	A	0.93	0.03	1.3	0.215	2.06	0.085	-0.596	0.003
54	P								
55	E	0.97	0.035	1.14	0.175	2.22	0.09	-0.409	0.027
56	D	1.02	0.03	1.51	0.125	2.328	0.0195	-0.293	0.009
57	P								
58	N	1.108	0.021	1.96	0.185	2.68	0.065	-0.212	0.013
59	E	1.08	0.035	1.53	0.17	2.48	0.055	-0.181	0.013
60	E								
61	A	0.99	0.03	1.1	0.18	2.1	0.07	-0.469	0.017
62	V	0.92	0.035	1.07	0.195	2.01	0.06	-0.446	0.023
63	S	0.94	0.03	1.68	0.17	2.5	0.055	-0.781	0.010
64	Q	1.002	0.022	1.73	0.21	2.7	0.065	-0.571	0.002
65	I	0.99	0.035	1.42	0.13	2.192	0.018	-0.307	0.021
66	F	1.02	0.035	1.67	0.185	2.72	0.085	-0.294	0.005
67	P								
68	D	1.12	0.03	2.08	0.21	3	0.075	-0.139	0.009
69	S	1.076	0.021	2.09	0.16	2.93	0.06	-0.093	0.024
70	V	0.96	0.03	1.22	0.14	1.99	0.03	-0.392	0.029
71	M	1.063	0.0215	2.01	0.175	3.16	0.075	-0.326	0.016
72	L	1.05	0.03	1.39	0.195	2.66	0.075	-0.229	0.024
73	A	1.03	0.03	1.8	0.17	2.76	0.085	-0.242	0.001
74	V	0.96	0.04	1.46	0.155	2.29	0.07	-0.316	0.038
75	Q	1.065	0.024	1.93	0.14	2.81	0.065	-0.398	0.013
76	E	1.07	0.03	1.73	0.165	2.77	0.06	-0.223	0.006
77	G	1.043	0.02	1.95	0.18	2.62	0.075	-0.320	0.004
78	I	1.07	0.035	1.6	0.15	2.35	0.055	-0.067	0.007
79	D	1.16	0.025	1.91	0.185	2.99	0.065	-0.171	0.001
80	L	1.02	0.04	1.41	0.2	2.52	0.075	-0.129	0.022
81	L	1.09	0.035	1.75	0.14	2.49	0.03	-0.094	0.021
82	T	1.06	0.035	1.57	0.18	2.56	0.055	-0.144	0.001
83	F	1.02	0.035	1.41	0.21	2.52	0.07	-0.198	0.017
84	P								
85	P								
86	A	1.01	0.03	1.45	0.2	2.38	0.06	-0.287	0.005
87	P								
88	G	0.87	0.025	1.72	0.19	2.32	0.07	-0.636	0.016
89	S	0.884	0.0235	1.79	0.2	2.77	0.07	-0.279	0.021
90	P								
91	E	1.04	0.035	1.51	0.145	2.328	0.0175	-0.310	0.009
92	P								
93	P								
94	H	1.126	0.0235	1.71	0.195	2.67	0.07	-0.265	0.011
95	L	1.09	0.045	2.51	0.195	3.87	0.115	-0.338	0.003
96	S	0.92	0.03	2.09	0.22	2.94	0.09	-1.028	0.129
97	R	1.02	0.03	1.93	0.21	3.17	0.065	-0.323	0.016
98	Q	1.02	0.03	1.41	0.15	2.51	0.045	-0.306	0.004
99	P								
100	E	1.06	0.03	1.76	0.215	2.73	0.075	-0.308	0.036
101	Q	1	0.03	0.93	0.135	1.759	0.011	-0.317	0.007
102	P								
103	E								

104	Q	0.96	0.03	0.89	0.18	1.92	0.07	-0.584	0.004
105	R	1.002	0.022	1.73	0.18	2.7	0.07	-0.551	0.003
106	A	0.988	0.0245	1.59	0.165	2.47	0.05	-0.626	0.017
107	L	0.9	0.03	1.14	0.195	2.18	0.06	-0.878	0.053
108	G	0.82	0.025	0.88	0.175	1.61	0.06	-1.026	0.019
109	P								
110	V	0.86	0.03	1.06	0.18	1.98	0.055	-1.057	0.027
111	S	0.94	0.03	1.68	0.16	2.5	0.05	-0.952	0.001
112	M	1.08	0.035	1.92	0.12	2.698	0.0175	-0.599	0.009
113	P								
114	N	0.963	0.023	2.11	0.2	2.87	0.075	-0.615	0.027
115	L	0.99	0.03	1.38	0.115	2.312	0.005	-0.421	0.028
116	V	1.02	0.035	1.41	0.195	2.52	0.07	-0.262	0.013
117	P								
118	E	0.97	0.035	1.14	0.195	2.22	0.09	-0.306	0.010
119	V	0.96	0.045	1.28	0.19	2.43	0.075	-0.209	0.014
120	I	0.94	0.045	1.47	0.185	2.62	0.07	-0.246	0.019
121	D	1.11	0.03	1.56	0.22	2.8	0.075	-0.227	0.001
122	L	1.06	0.03	1.57	0.175	2.7	0.095	-0.189	0.023
123	T	1.12	0.025	1.71	0.175	2.98	0.065	-0.142	0.013
124	C	1.06	0.04	2.3	0.18	3.18	0.06	-0.227	0.004
125	H								
126	E	0.98	0.03	1.69	0.16	2.73	0.065	-0.597	0.053
127	A	0.929	0.024	1.4	0.24	2.49	0.095	-0.797	0.014
128	G	0.829	0.0245	1.42	0.195	1.87	0.065	-0.973	0.028
129	F	0.93	0.03	1.14	0.19	2.15	0.065	-0.554	0.007
130	P								
131	P								
132	S	0.961	0.023	1.47	0.2	2.44	0.08	-0.624	0.024
133	D	0.99	0.025	1.2	0.195	2.41	0.045	-0.521	0.016
134	D	0.98	0.035	0.98	0.18	1.84	0.045	-0.502	0.005
135	E	0.98	0.03	0.93	0.16	1.79	0.04	-0.573	0.011
136	D								
137	E								
138	E								
139	G	0.92	0.03	0.76	0.16	1.58	0.075	-0.859	0.025
140	E	0.93	0.03	0.77	0.16	1.69	0.05	-0.649	0.014
141	E								
142	F	0.96	0.03	0.68	0.175	1.83	0.07	-0.837	0.042
143	V	0.88	0.03	0.41	0.15	1.49	0.065	-1.029	0.036
144	L	0.87	0.03	0.28	0.155	1.17	0.055	-1.172	0.025
145	D	0.75	0.035	0.27	0.155	1.3	0.06	-1.988	0.016
146	Y	0.58	0.03	0.05	0.09	0.77	0.035	-1.677	0.037

Tabla A3.12. Parámetros de relajación por residuo de AdE1A- Δ L

#	AA	R1	Error R1	R2	Error R2	R1 ρ	Error R1 ρ	NOE	Error NOE
38	F								
39	E	0.87	0.04	1.35	0.025	1.37	0.08	-1.043	0.057
40	P								
41	P								
42	T	0.96	0.04	1.91	0.1	1.75	0.07	-0.259	0.020
43	L	1.08	0.035	2.1	0.085	1.89	0.185	-0.166	0.004
44	H	0.956	0.022	2.99	0.085	2.31	0.135	-0.338	0.037
45	E	0.86	0.05	1.8	0.04	1.3	0.25	-0.342	0.031
46	L	0.94	0.055	1.89	0.075	1.7	0.045	-0.245	0.004
47	Y	0.9	0.055	1.41	0.06	1.67	0.18	-0.331	0.004
48	D	1.02	0.055	1.74	0.075	1.41	0.05	-0.150	0.006
49	L	0.95	0.055	2.12	0.065	1.5	0.2	-0.260	0.004
50	D	0.97	0.065	0.88	0.035	1.52	0.095	-0.291	0.004
51	V	0.87	0.055	1.41	0.07	1.48	0.05	-0.212	0.017
52	T	0.93	0.045	1.81	0.06	1.58	0.085	-0.448	0.015
53	A	0.91	0.045	1.83	0.05	1.22	0.17	-0.459	0.007
54	P								
55	E	0.9	0.06	1.75	0.055	1.41	0.185	-0.491	0.014
56	D	0.91	0.05	1.76	0.03	1.86	0.05	-0.374	0.003
57	P								
58	N	1.01	0.055	2.47	0.045	2.15	0.07	-0.226	0.001
59	E	0.98	0.05	2.05	0.035	1.84	0.09	-0.068	0.000
60	E								
61	A	0.88	0.05	1.45	0.085	1.38	0.15	-0.519	0.014
62	V	0.81	0.06	1.52	0.035	1.31	0.1	-0.464	0.009
63	S	0.85	0.03	2.39	0.04	1.81	0.11	-0.836	0.004
64	Q	0.93	0.035	2.16	0.045	1.96	0.135	-0.675	0.008
65	I	0.95	0.055	1.85	0.05	1.5	0.22	-0.332	0.004
66	F	0.97	0.045	1.81	0.045	1.53	0.25	-0.288	0.012
67	P								
68	D	1.06	0.03	2.6	0.06	2.22	0.14	-0.156	0.015
69	S	1.03	0.04	2.29	0.065	2.06	0.165	-0.104	0.009
70	V	0.95	0.055	1.85	0.045	1.85	0.045	-0.358	0.005
71	M	1.05	0.035	2.26	0.05	2.34	0.105	-0.273	0.002
72	L	0.87	0.05	1.68	0.05	1.73	0.095	-0.264	0.008
73	A	0.95	0.045	2.27	0.065	1.99	0.12	-0.270	0.008
74	V	0.85	0.045	1.64	0.06	1.47	0.13	-0.311	0.010
75	Q	0.99	0.035	2.13	0.045	1.99	0.2	-0.634	0.018
76	E	0.96	0.045	2.13	0.055	1.91	0.125	-0.271	0.012
77	G	0.98	0.035	2.36	0.075	1.58	0.075	-0.313	0.013
78	I	0.96	0.055	1.7	0.04	1.61	0.11	-0.095	0.007
79	D	1.06	0.06	1.81	0.03	1.89	0.21	-0.295	0.008
80	L	0.94	0.065	1.78	0.065	1.79	0.075	-0.185	0.015
81	L	0.97	0.06	2.1	0.04	1.94	0.065	-0.075	0.020
82	T	0.94	0.05	2	0.055	1.87	0.055	-0.189	0.003
83	F	0.91	0.06	1.62	0.035	1.66	0.08	-0.323	0.000
84	P								
85	P								
86	A	0.9	0.05	1.74	0.06	1.54	0.19	-0.309	0.006
87	P								
88	G	0.82	0.04	2.31	0.06	1.74	0.08	-0.565	0.023
89	S	0.82	0.04	2.25	0.03	1.95	0.065	-0.303	0.004
90	P								
91	E	0.91	0.05	2.28	0.095	1.86	0.14	-0.440	0.039
92	P								
93	P								
94	H	1.03	0.05	2.06	0.06	2.02	0.05	-0.347	0.016
95	L	0.97	0.055	2.94	0.065	2.5	0.35	-0.279	0.013
96	S	0.9	0.035	2.63	0.04	1.99	0.205	-1.162	0.096
97	R	0.92	0.05	2.33	0.055	2.17	0.235	-0.421	0.016
98	Q	0.94	0.05	1.55	0.03	1.75	0.18	-0.311	0.000
99	P								
100	E	0.98	0.05	2.09	0.06	1.86	0.155	-0.356	0.006
101	Q	0.91	0.045	0.85	0.035	1.58	0.215	-0.398	0.006
102	P								

103	E								
104	Q	0.88	0.055	1.32	0.04	1.25	0.06	-0.568	0.020
105	R	0.93	0.045	2.16	0.05	1.96	0.135	-0.734	0.014
106	A	0.88	0.035	1.93	0.03	1.78	0.11	-0.727	0.035
107	L	0.79	0.06	1.4	0.04	1.58	0.125	-0.909	0.019
108	G	0.74	0.035	1.24	0.065	1	0.095	-1.246	0.014
109	P								
110	V	0.79	0.04	1.26	0.065	1.21	0.15	-1.100	0.017
111	S	0.85	0.03	2.04	0.035	2.4	0.085	-1.167	0.111
112	M	0.915	0.023	2.39	0.05	2.36	0.035	-0.894	0.018
113	P								
114	N	0.9	0.03	2.49	0.11	2.17	0.1	-0.850	0.041
115	L	0.96	0.06	1.38	0.025	1.96	0.035	-0.506	0.000
116	V	0.91	0.06	1.62	0.04	1.66	0.085	-0.222	0.009
117	P								
118	E								
119	V	0.93	0.045	1.4	0.045	1.69	0.045	-0.319	0.004
120	I	0.86	0.05	1.58	0.035	1.5	0.11	-0.587	0.011
121	D	0.96	0.04	1.76	0.05	1.69	0.15	-0.647	0.014
122	A	0.91	0.05	1.61	0.045	1.55	0.07	-0.649	0.007
123	A								
124									
125	A								
126	A								
127	A	0.72	0.045	1.05	0.045	1.43	0.17	-1.285	0.002
128	G	0.69	0.035	1.69	0.045	1.18	0.085	-1.208	0.029
129	F	0.81	0.045	1.28	0.03	1.29	0.095	-0.814	0.001
130	P								
131	P								
132	S	0.86	0.045	1.6	0.065	1.47	0.09	-0.952	0.013
133	D	0.92	0.05	1.54	0.04	1.84	0.075	-0.600	0.003
134	D	0.87	0.055	1.28	0.045	1.21	0.06	-0.566	0.010
135	E	0.88	0.055	1.15	0.035	1.09	0.065	-0.599	0.015
136	D								
137	E								
138	E								
139	G	0.83	0.045	0.99	0.04	0.9	0.12	-0.978	0.011
140	E	0.85	0.05	1.27	0.04	1	0.08	-0.554	0.008
141	E	0.88	0.05	1.08	0.04	1.1	0.4	-0.703	0.012
142	F	0.89	0.045	0.9	0.025	1.18	0.045	-0.883	0.013
143	V	0.8	0.05	0.83	0.04	0.83	0.065	-1.198	0.017
144	L	0.77	0.055	0.65	0.04	0.48	0.145	-1.170	0.013
145	D	0.68	0.05	0.65	0.03	0.66	0.05	-1.573	0.018
146	Y	0.5	0.05	0.27	0.035	0.6	0.4	-2.383	0.074

Figura A3.3. Parámetros de relajación por residuo para AdE1A-ΔE

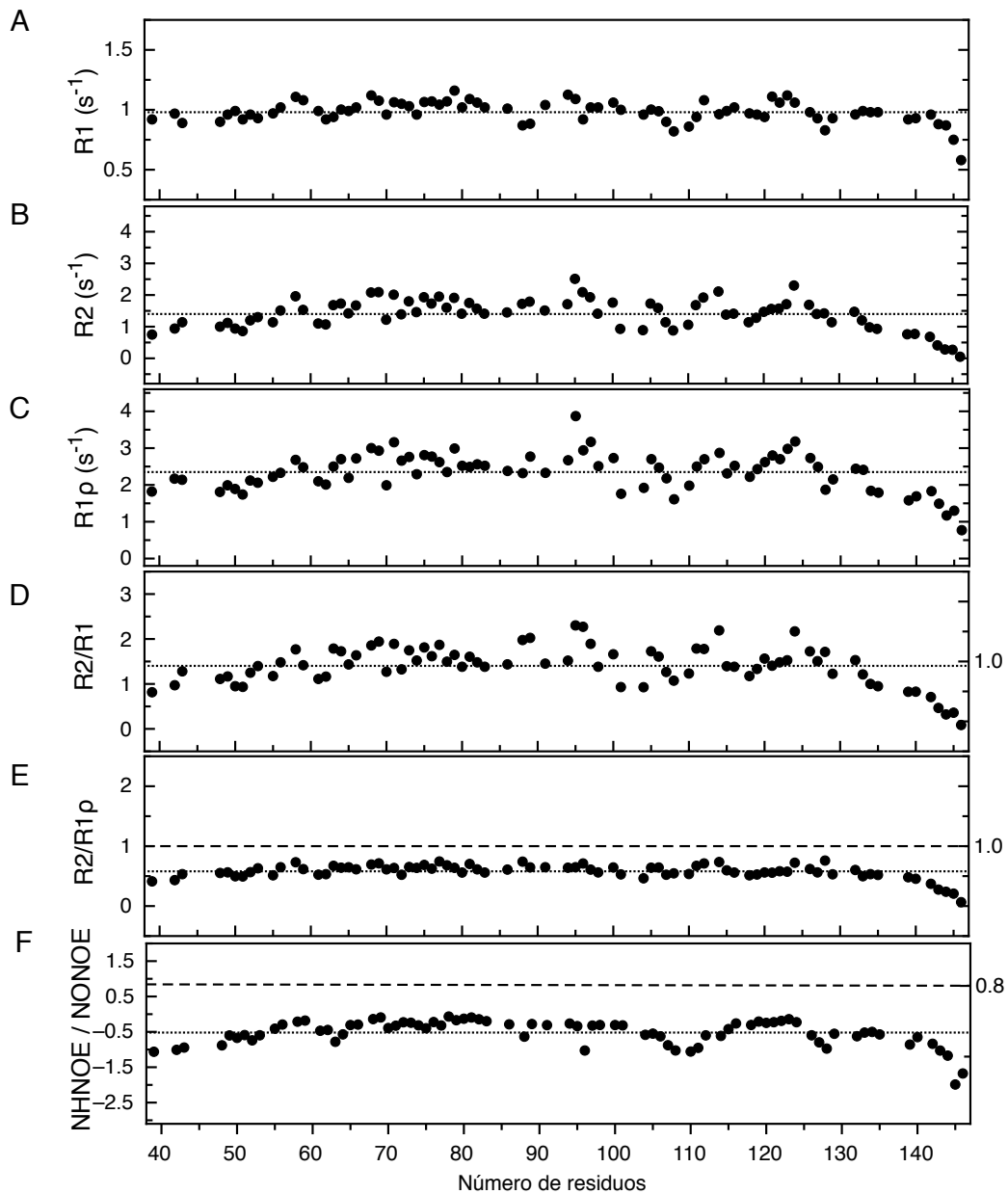


Figura A3.3. Parámetros de relajación por residuo para AdE1A-ΔE. A) Gráfico de relajación espín-red (R_1) B) Gráfico de relajación espín-espín (R_2) C) Gráfico de relajación rotacional ($R_{1\rho}$) D) Gráfico de razón R_2/R_1 . E) Gráfico de razón $R_2/R_{1\rho}$. Los valores por residuo se comparan con respecto a la unidad (línea de rayas). F) Gráfico de $NHNOE/NONOE$. El valor 0.8 (línea de rayas) es el máximo esperado para conformaciones plegadas o estructuradas. Los valores promedio de cada uno de los parámetros (presentados en la Tabla 3.4 Capítulo III) están señalados con la línea de puntos.

Figura A3.4. Parámetros de relajación por residuo para AdE1A- Δ L.

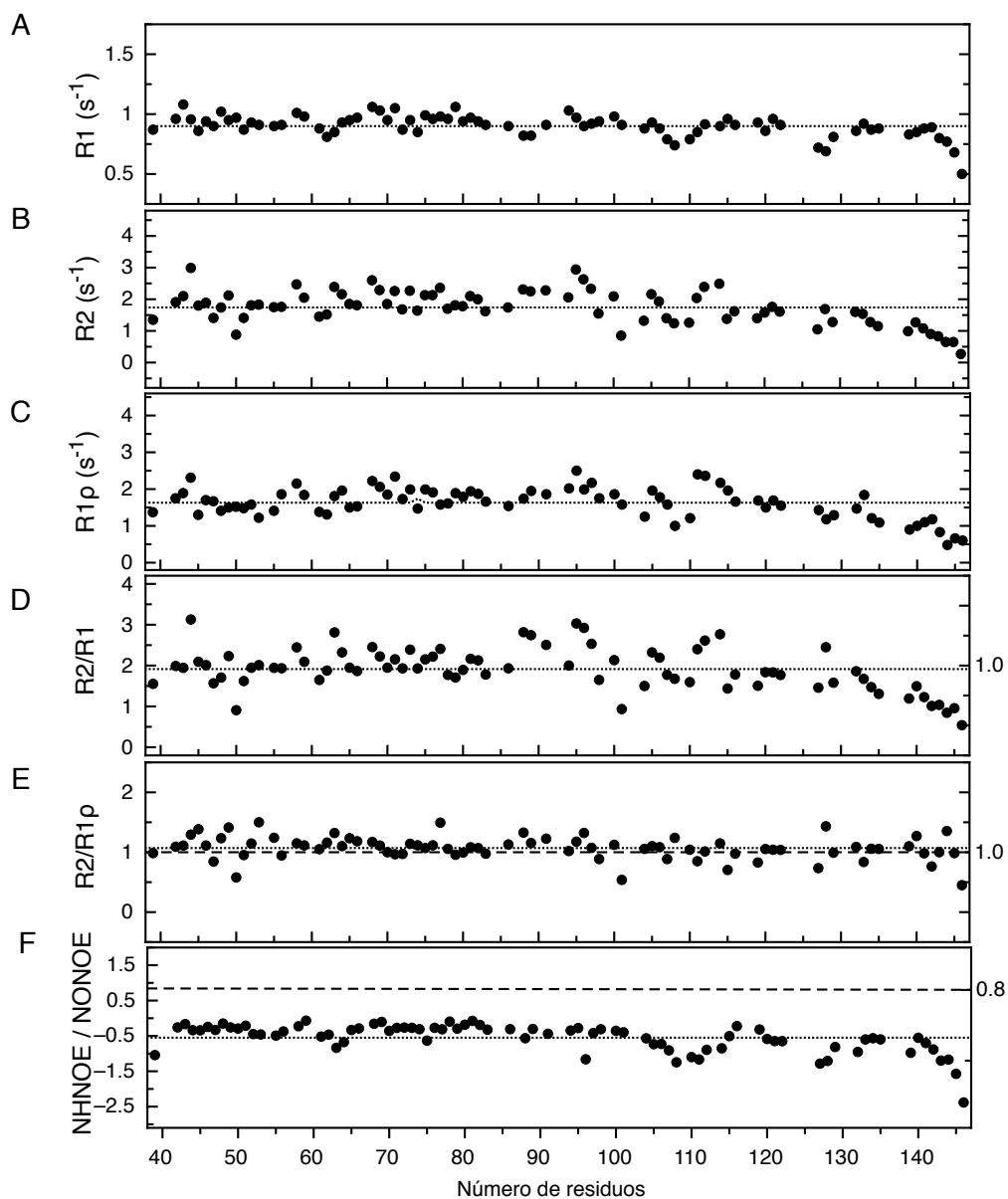


Figura A3.4. Parámetros de relajación por residuo para AdE1A- Δ L. A) Gráfico de relajación espín-red (R_1) B) Gráfico de relajación espín-espín (R_2) C) Gráfico de relajación rotacional ($R_{1\rho}$) D) Gráfico de razón R_2/R_1 . E) Gráfico de razón $R_2/R_{1\rho}$. Los valores por residuo se comparan con respecto a la unidad (línea de rayas). F) Gráfico de $NHNOE/NONOE$. El valor 0.8 (línea de rayas) es el máximo esperado para conformaciones plegadas o estructuradas. Los valores promedio de cada uno de los parámetros (presentados en la Tabla 3.4 Capítulo 3) están señalados con la línea de puntos.

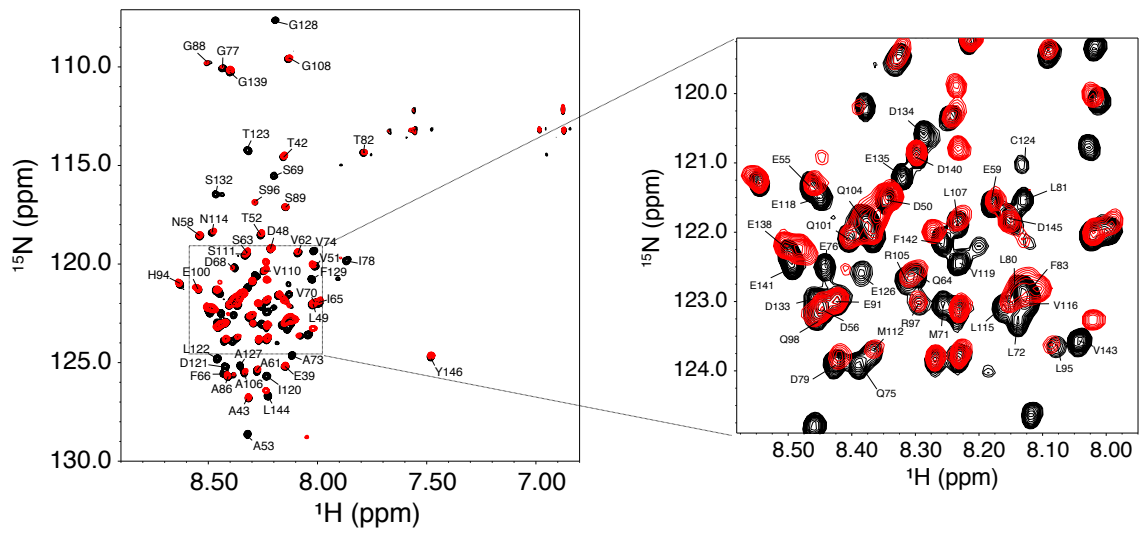


Figura A3.5. Superposición de espectros TROSY de AdE1A-ΔE libre a 690 μM (negro) y del complejo [RbAB:AdE1A-ΔE] 1:1 a 690 μM (rojo).

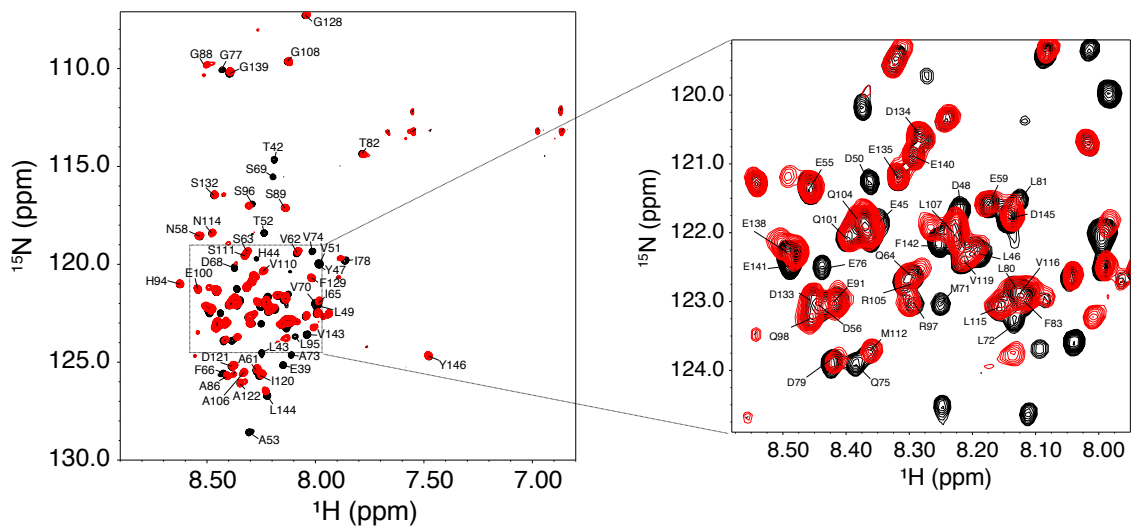


Figura A3.6. Superposición de espectros TROSY de AdE1A-ΔL libre a 680 μM (negro) y del complejo [RbAB:AdE1A-ΔL] 1:1 a 680 μM (rojo).

Tabla A3.13. Valores de Intensidad de resonancias para el experimento de titulación de AdE1A.WT con concentraciones crecientes de RbAB

#	AA	Puntos de la titulación													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14
39	E	2.599	2.436	2.325	2.096	1.935	1.496	1.293	0.529	0.331	0.264	0.011	0.025	0.014	0.003
42	T	2.415	2.168	2.224	2.071	1.962	1.561	1.372	0.571	0.368	0.278	0.003	0.000	0.009	0.005
43	L	2.551	2.411	2.372	2.134	1.924	1.467	1.301	0.533	0.333	0.244	0.006	0.009	0.012	0.013
44	H	1.168	1.131	1.077	1.029	0.952	0.740	0.678	0.425	0.384	0.374	0.308	0.307	0.277	0.298
45	E	2.032	2.001	1.915	1.842	1.785	1.550	1.236	0.785	0.636	0.620	0.387	0.373	0.350	0.315
46	L	3.802	3.613	3.456	3.243	2.841	2.186	1.909	0.818	0.532	0.434	0.105	0.101	0.100	0.100
47	Y	6.022	5.758	5.457	5.174	4.465	3.730	3.202	1.263	0.797	0.588	0.013	0.010	0.009	0.001
48	D	3.401	3.248	3.392	3.135	2.977	2.418	1.890	1.009	0.825	0.254	0.173	0.165	0.123	0.132
49	L	3.908	3.682	3.624	3.282	2.957	2.161	1.848	0.760	0.478	0.378	0.015	0.021	0.023	0.020
50	D	3.549	3.301	3.304	2.984	2.792	2.180	1.898	0.832	0.574	0.463	0.128	0.119	0.094	0.128
51	V	6.103	6.013	5.914	5.447	4.950	3.266	2.871	1.128	0.664	0.508	0.010	0.002	0.005	-0.003
52	T	3.027	3.058	2.974	2.773	2.561	1.963	1.733	0.718	0.454	0.354	-0.001	0.003	0.009	0.010
53	A	3.464	3.258	3.133	2.916	2.593	1.920	1.833	0.975	0.763	0.672	0.329	0.342	0.369	0.436
55	E	3.642	3.677	3.649	3.416	3.245	2.631	2.377	1.168	0.854	0.672	0.146	0.130	0.132	0.096
56	D	4.014	3.993	3.976	3.660	3.544	2.941	2.706	1.544	1.231	1.133	0.718	0.741	0.760	0.886
58	N	3.100	3.041	2.923	2.795	2.654	2.167	2.064	1.454	1.423	1.375	1.167	1.178	1.246	1.385
59	E	3.523	3.390	3.417	3.151	3.068	2.526	2.345	1.410	1.159	1.071	0.656	0.653	1.024	0.775
61	A	3.014	3.231	3.241	3.084	3.151	2.918	2.933	2.526	2.661	2.681	2.377	2.311	2.369	2.339
62	V	3.221	2.985	2.876	3.153	3.310	3.107	3.150	2.755	2.568	2.571	2.470	2.428	2.480	2.478
63	S	2.784	2.788	2.746	2.629	2.704	2.460	2.475	2.303	2.249	2.301	1.935	1.913	1.881	2.016
64	Q	2.626	2.585	2.549	2.451	2.602	2.488	2.542	2.497	2.494	2.580	2.312	2.486	2.641	2.890
65	I	4.167	4.256	4.359	3.973	4.129	3.807	3.710	2.676	2.236	2.175	2.514	2.423	2.450	2.369
66	F	2.974	2.854	2.745	2.607	2.372	2.059	1.969	1.393	1.268	1.072	0.853	0.831	0.859	0.890
68	D	2.912	2.886	2.731	2.691	2.564	2.201	2.155	1.821	1.690	1.692	1.344	1.297	1.310	1.260
69	S	2.273	2.200	2.213	2.124	2.076	1.796	1.773	1.604	1.555	1.565	1.109	1.092	1.151	1.137
70	V	4.010	3.774	5.245	5.172	4.901	4.073	3.929	3.380	3.336	3.237	2.489	2.400	2.488	2.505
71	M	3.057	3.042	2.972	2.794	2.673	2.052	2.093	1.861	1.787	1.795	0.921	0.891	0.910	0.880
72	L	3.359	3.365	3.188	3.028	3.087	2.674	2.555	1.893	1.712	1.646	1.437	1.390	1.427	1.406
73	A	3.261	3.046	2.841	2.632	2.675	2.225	2.327	2.107	2.039	2.047	0.509	0.475	0.526	0.472
74	V	3.389	3.279	3.272	3.035	2.991	2.934	2.945	2.680	2.636	2.636	0.932	0.874	0.902	0.856
75	Q	2.810	2.963	2.840	2.820	2.621	2.500	2.437	2.177	2.222	2.172	1.414	1.366	1.388	1.342
76	E	3.370	3.218	3.100	2.951	2.912	2.696	2.760	2.509	2.439	2.440	1.241	1.171	1.117	1.159
77	G	2.132	1.915	2.221	2.140	2.205	1.991	1.960	1.781	1.746	1.773	1.248	1.217	1.247	1.233
78	I	3.153	3.117	3.033	2.912	2.901	2.565	2.540	2.232	2.133	2.136	1.413	1.349	1.347	1.334
79	D	3.285	3.162	2.988	2.929	2.769	2.931	2.995	2.722	2.711	2.694	1.969	1.897	1.943	1.910
80	L	5.044	5.007	4.736	4.709	4.495	3.758	2.821	2.809	2.843	2.944	2.588	2.532	2.616	2.657
81	L	2.544	2.597	2.559	2.466	2.467	2.145	2.072	1.707	1.666	1.610	1.197	1.141	1.168	1.137
82	T	2.077	2.176	2.100	2.050	2.075	1.851	1.985	1.789	1.734	1.713	1.317	1.278	1.321	1.325
83	F	4.856	4.654	4.452	4.216	3.889	1.925	2.244	2.635	2.650	2.789	2.328	2.297	1.427	1.406
86	A	2.423	2.294	2.216	2.143	2.063	1.690	2.312	2.007	2.031	1.998	1.758	1.711	1.755	1.767
88	G	0.916	0.919	0.891	0.853	0.882	0.804	0.820	0.712	0.722	0.717	0.636	0.618	0.625	0.614
89	S	1.605	1.627	1.609	1.537	1.593	1.438	1.450	1.301	1.396	1.414	1.295	1.276	1.318	1.370
91	E	3.045	3.111	3.148	2.936	3.146	2.945	2.997	2.735	2.654	2.687	2.323	2.318	2.488	2.403
94	H	3.437	3.413	3.361	3.124	3.062	2.584	2.301	1.872	1.884	1.971	1.935	1.996	1.939	2.179
95	L	1.220	1.149	1.203	1.139	1.145	1.041	1.042	0.911	0.887	0.880	0.800	0.785	0.762	0.831
96	S	0.865	0.844	0.841	0.819	0.843	0.766	0.794	0.749	0.756	0.774	0.730	0.724	0.747	0.778
97	R	1.308	1.310	1.288	1.246	1.289	1.201	1.230	1.173	1.194	1.191	1.098	1.079	1.100	1.120
98	Q	1.948	1.917	1.801	1.819	1.803	1.607	1.711	1.722	1.825	1.765	1.611	1.519	1.793	1.457
100	E	2.906	2.930	2.978	2.785	2.955	2.779	2.810	2.570	2.695	2.655	2.463	2.404	2.475	2.232
101	Q	2.679	2.657	2.556	2.439	2.344	1.987	1.874	1.624	1.650	1.691	1.658	1.698	1.727	2.043
104	Q	3.788	3.632	3.275	3.304	2.859	4.790	4.361	2.803	2.466	2.309	1.658	1.708	1.818	1.344
105	R	2.462	2.394	2.265	2.261	2.284	2.004	1.934	1.880	2.076	2.083	1.986	1.970	2.641	1.349
106	A	2.101	2.033	1.965	1.854	1.719	1.359	1.295	1.183	1.190	1.248	1.262	1.291	1.390	1.627
107	L	3.174	3.062	2.890	2.757	2.487	2.213	2.033	1.654	1.638	1.658	1.610	1.672	1.834	2.180
108	G	2.999	2.907	2.777	2.637	2.462	1.964	1.785	1.667	1.691	1.758	1.832	1.868	2.007	2.305
110	V	2.891	2.774	2.698	2.593	2.472	2.083	2.051	1.725	1.676	1.777	1.886	1.898	1.991	2.203
111	S	1.984	1.905	1.775	1.757	1.660	1.689	1.633	1.421	1.408	1.402	1.150	1.133	2.049	1.022
112	M	1.875	1.915	1.845	1.786	1.792	1.681	1.661	1.566	1.530	1.550	1.050	0.987	0.814	1.005
114	N	1.706	1.650	1.597	1.479	1.406	1.239	1.264	1.059	1.053	1.017	0.900	0.858	0.885	0.857
115	L	2.802	2.627	2.478	2.327	2.039	1.979	1.847	1.272	1.214	1.103	1.113	1.050	1.125	0.349
116	V	1.825	1.649	1.416	1.546	1.438	3.078	2.790	2.176	2.061	2.029	2.024	2.005	1.687	1.138
118	E	4.136	3.892	3.900	3.532	3.305	2.575	2.197	0.954	0.600	0.515	0.049	0.060	0.050	0.061
119	V	2.939	2.928	2.940	2.728	2.540	1.908	1.590	0.694	0.445	0.364	0.055	0.041	0.041	0.054
120	I	2.683	2.433	2.660	2.489	2.385	1.927	1.697	0.755	0.493	0.392	0.040	0.039	0.022	0.024
121	D	3.141	2.968	2.857	2.660	2.390	1.906	1.701	1.054	0.625	0.534	0.053	0.054	0.521	0.600
122	L	2.856	2.737	2.691	2.455	2.230	1.640	1.507	0.644	0.436	0.323	0.016	0.021	0.012	0.017
123	T	2.484	2.482	2.504	2.299	2.199	1.756	1.561	0.661	0.424	0.330	0.017	0.010	0.005	0.010
124	C	1.172	1.114	1.174	1.095	1.060	0.830	0.766	0.451	0.371	0.320	0.213	0.189	0.194	0.164

126	E	1.329	1.278	1.220	1.165	1.070	0.844	0.785	0.441	0.330	0.296	0.097	0.100	0.090	0.093
127	A	2.320	2.259	2.143	2.064	1.852	1.489	1.313	0.570	0.342	0.270	0.001	-0.001	0.006	0.011
128	G	1.865	1.777	1.760	1.634	1.530	1.227	1.090	0.495	0.328	0.245	0.020	0.017	0.016	0.018
129	F	2.675	2.612	2.566	2.383	2.219	1.727	1.502	0.677	0.460	0.362	0.051	0.054	0.058	0.060
132	S	2.363	2.228	2.176	2.013	1.839	1.568	1.397	0.640	0.443	0.337	0.038	0.034	0.033	0.032
133	D	2.514	2.511	2.456	2.316	2.194	1.788	1.648	0.892	0.685	0.255	0.137	0.148	0.289	0.355
134	D	3.342	3.264	3.185	2.992	2.739	2.133	1.890	0.881	0.614	0.498	0.081	0.084	0.089	0.080
135	E	3.708	3.655	3.495	3.188	3.074	2.512	2.263	1.053	0.671	0.529	0.053	0.044	0.085	0.075
138	E	5.086	5.124	4.920	4.831	4.717	4.029	3.923	2.929	3.045	3.049	2.724	2.673	2.706	2.799
139	G	3.814	3.895	3.876	3.750	3.851	3.498	3.524	3.110	3.136	3.093	2.726	2.695	2.794	2.477
140	E	4.399	4.414	4.293	4.195	4.178	3.683	3.772	3.398	3.257	3.303	2.731	2.702	2.700	3.037
141	E	5.086	5.124	4.920	4.831	4.717	4.029	3.923	3.200	3.045	3.049	2.724	2.673	2.706	2.693
142	F	3.921	3.985	3.936	3.743	3.713	3.276	3.152	2.506	2.459	2.433	2.048	1.993	2.033	2.001
143	V	3.578	3.362	3.268	3.498	3.395	2.856	2.619	1.938	1.954	1.952	1.705	1.656	1.668	1.712
144	L	3.959	3.926	3.813	3.649	3.410	2.731	2.482	1.912	1.824	1.808	1.803	1.758	1.785	1.749
145	D	4.337	4.303	4.187	4.048	3.986	3.512	3.456	2.952	2.814	2.767	2.141	2.078	2.192	2.017
146	Y	3.547	3.414	3.315	3.173	3.008	3.090	3.162	2.948	2.921	2.921	2.555	2.515	2.552	2.548

Detalles de espectros TROSY por pico para AdE1A.WT y mutantes

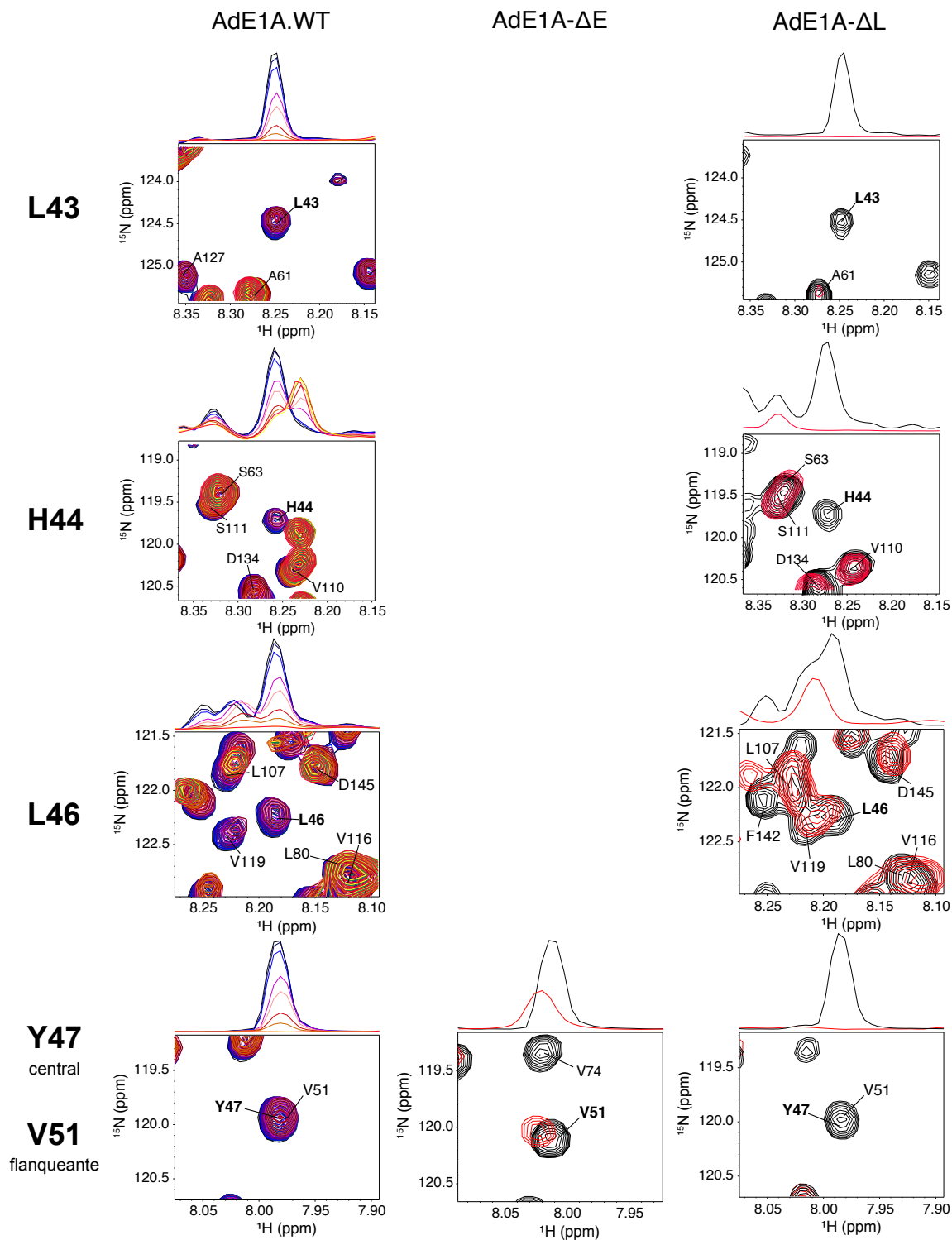


Figura A3.7. Detalle de los espectros TROSY para los residuos centrales (y un residuo flanqueante) del motivo E2F y sus proyecciones de intensidad de resonancia obtenidos punto a punto (código de colores en la Figura 3.13) en el experimento de titulación para AdE1A.WT (primer columna) y para las formas libres (en negro) y acomplejadas en relación molar 1:1 con RbAB (en rojo) de la mutante AdE1A-ΔE (segunda columna) y de la mutante AdE1A-ΔL (tercer columna). Nótese que los espectros faltantes corresponden a los residuos que han sido mutados en la variante AdE1A-ΔE.

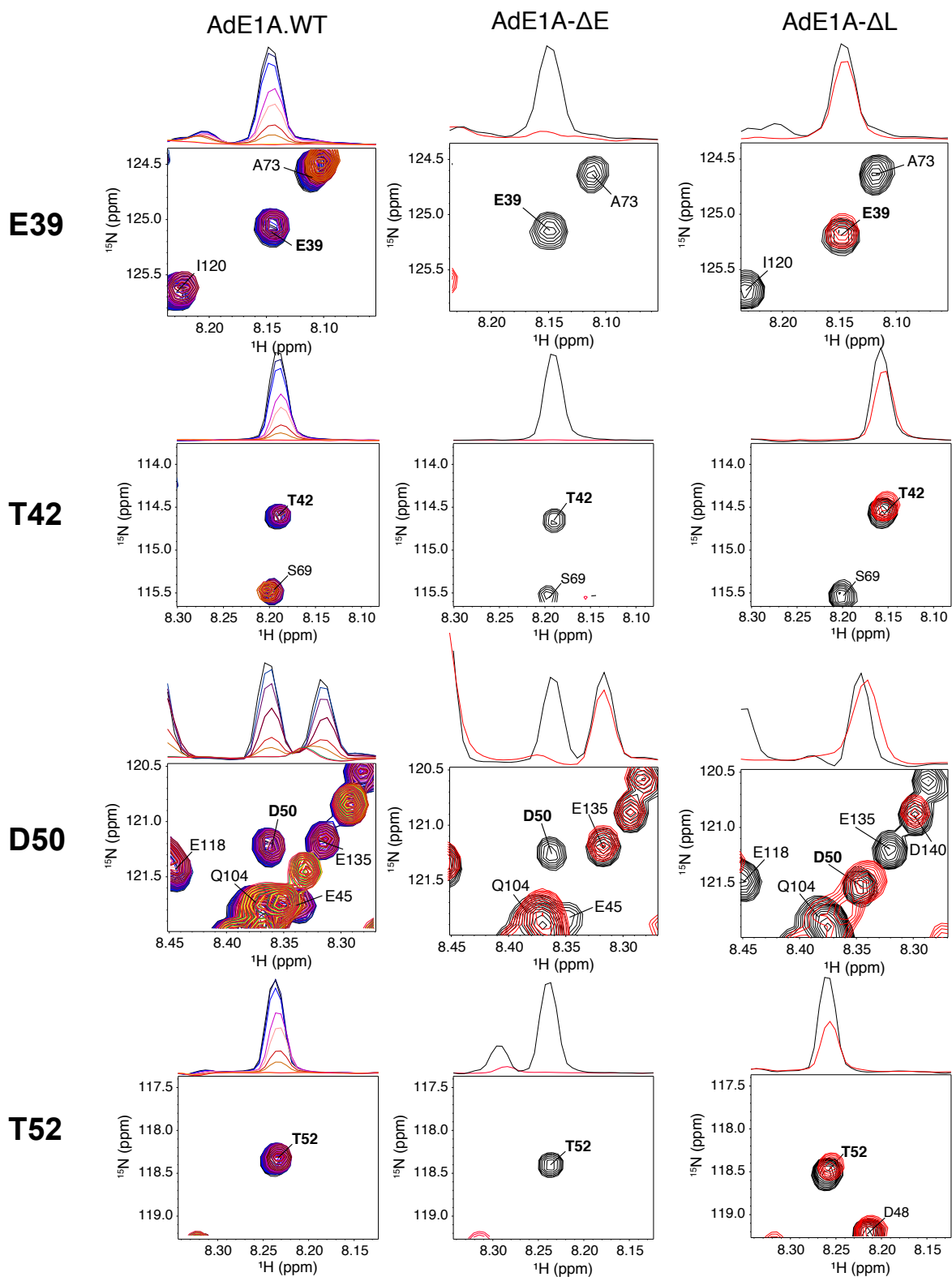


Figura A3.8. Detalle de los espectros TROSY para los residuos flanqueantes al motivo E2F y sus proyecciones de intensidad de resonancia obtenidos punto a punto (código de colores en la Figura 3.13) en el experimento de titulación para AdE1A.WT (primer columna) y para las formas libres (en negro) y acomplejadas en relación molar 1:1 con RbAB (en rojo) de la mutante AdE1A-ΔE (segunda columna) y de la mutante AdE1A-ΔL (tercer columna).

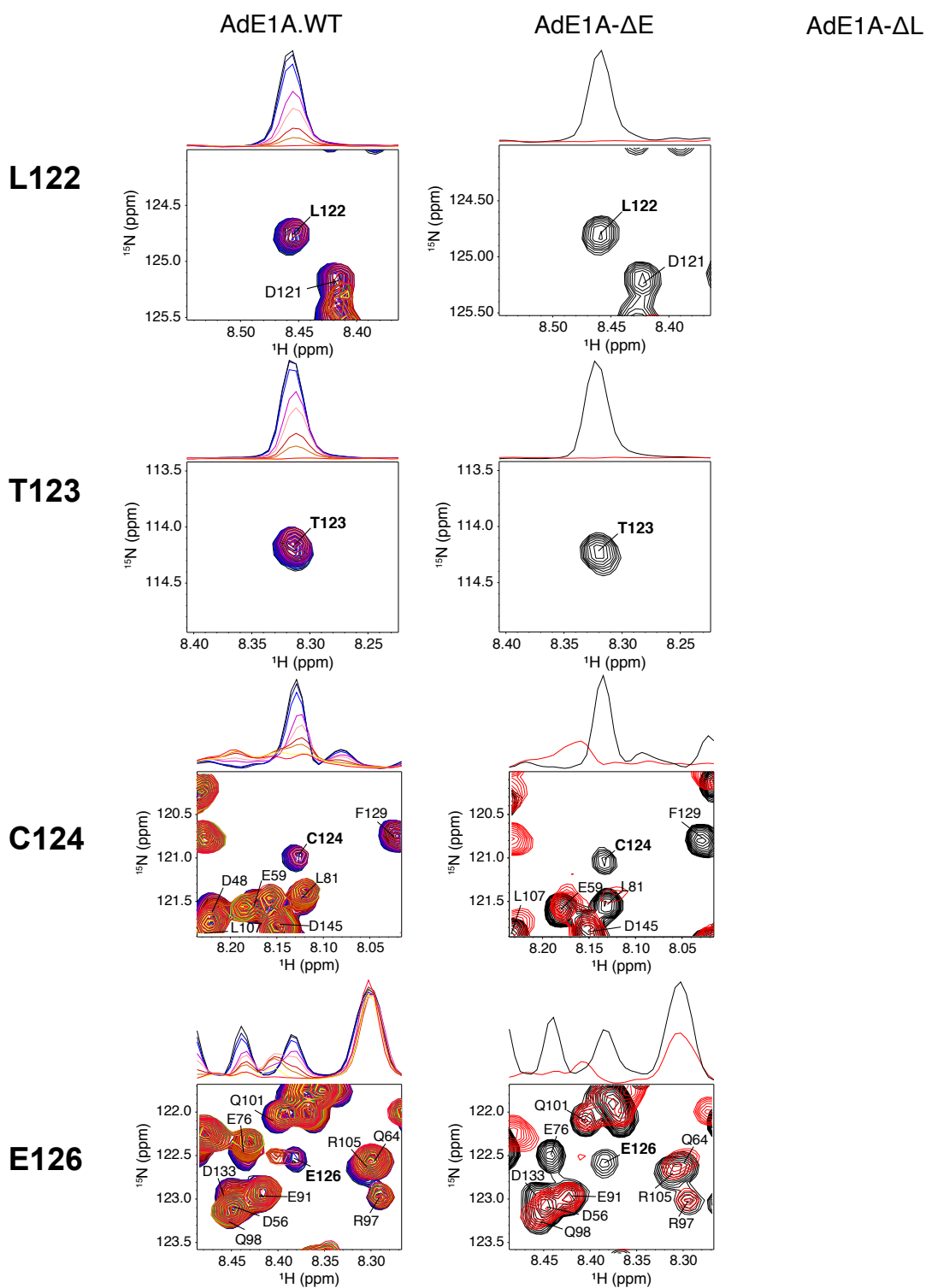


Figura A3.9. Detalle de los espectros TROSY para los residuos centrales del motivo LxCxE y sus proyecciones de intensidad de resonancia obtenidos punto a punto (código de colores en la Figura 3.13) en el experimento de titulación para AdE1A.WT (primer columna) y para las formas libres (en negro) y acomplejadas en relación molar 1:1 con RbAB (en rojo) de la mutante AdE1A- Δ E (segunda columna) y de la mutante AdE1A- Δ L (tercer columna). Nótese que los espectros faltantes corresponden a los residuos que han sido mutados en la variante AdE1A- Δ L.

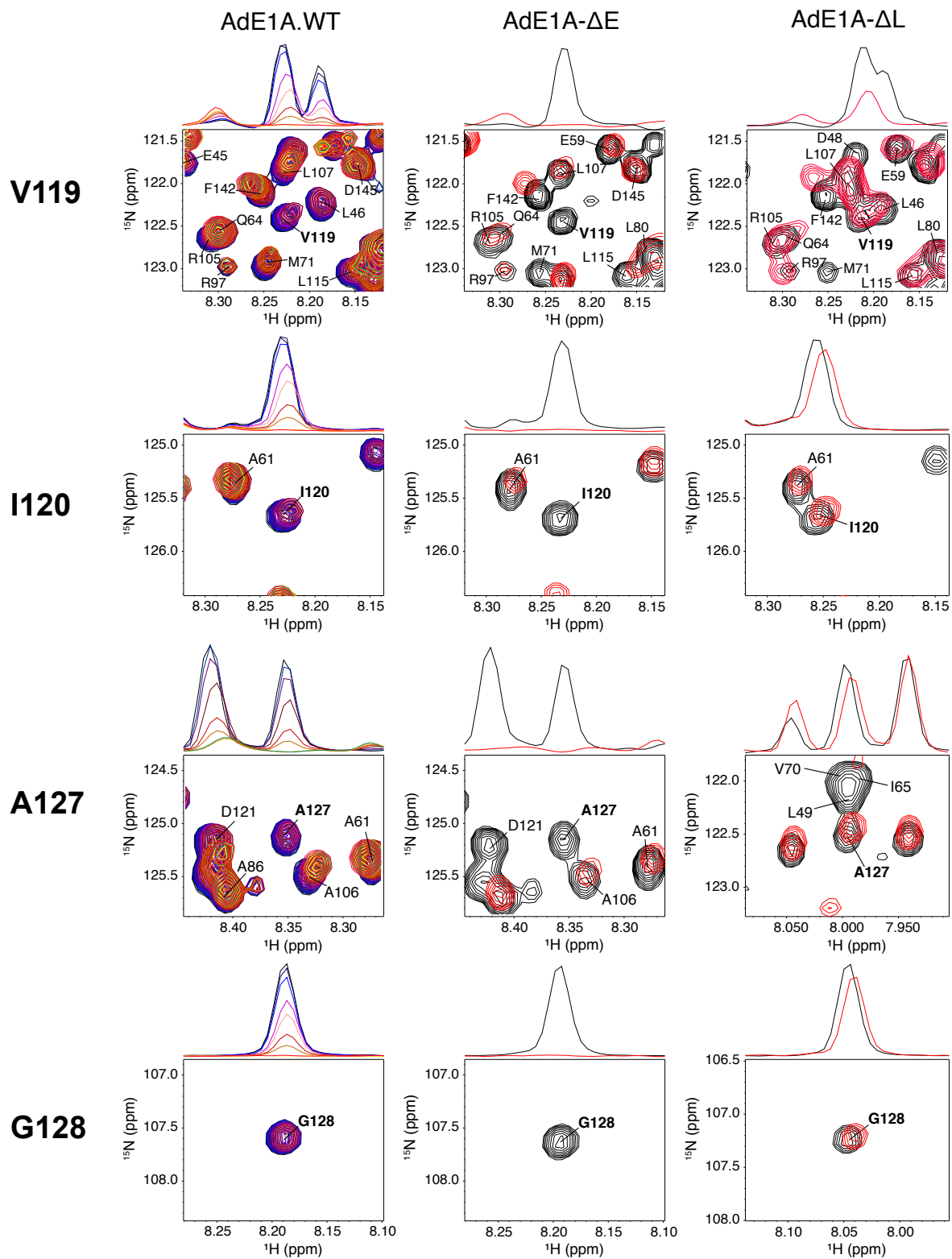


Figura A3.10. Detalle de los espectros TROSY para los residuos flanqueantes al motivo LxCxE y sus proyecciones de intensidad de resonancia obtenidos punto a punto (código de colores en la Figura 3.13) en el experimento de titulación para AdE1A.WT (primer columna) y para las formas libres (en negro) y acomplejadas en relación molar 1:1 con RbAB (en rojo) de la mutante AdE1A-ΔE (segunda columna) y de la mutante AdE1A-ΔL (tercer columna).

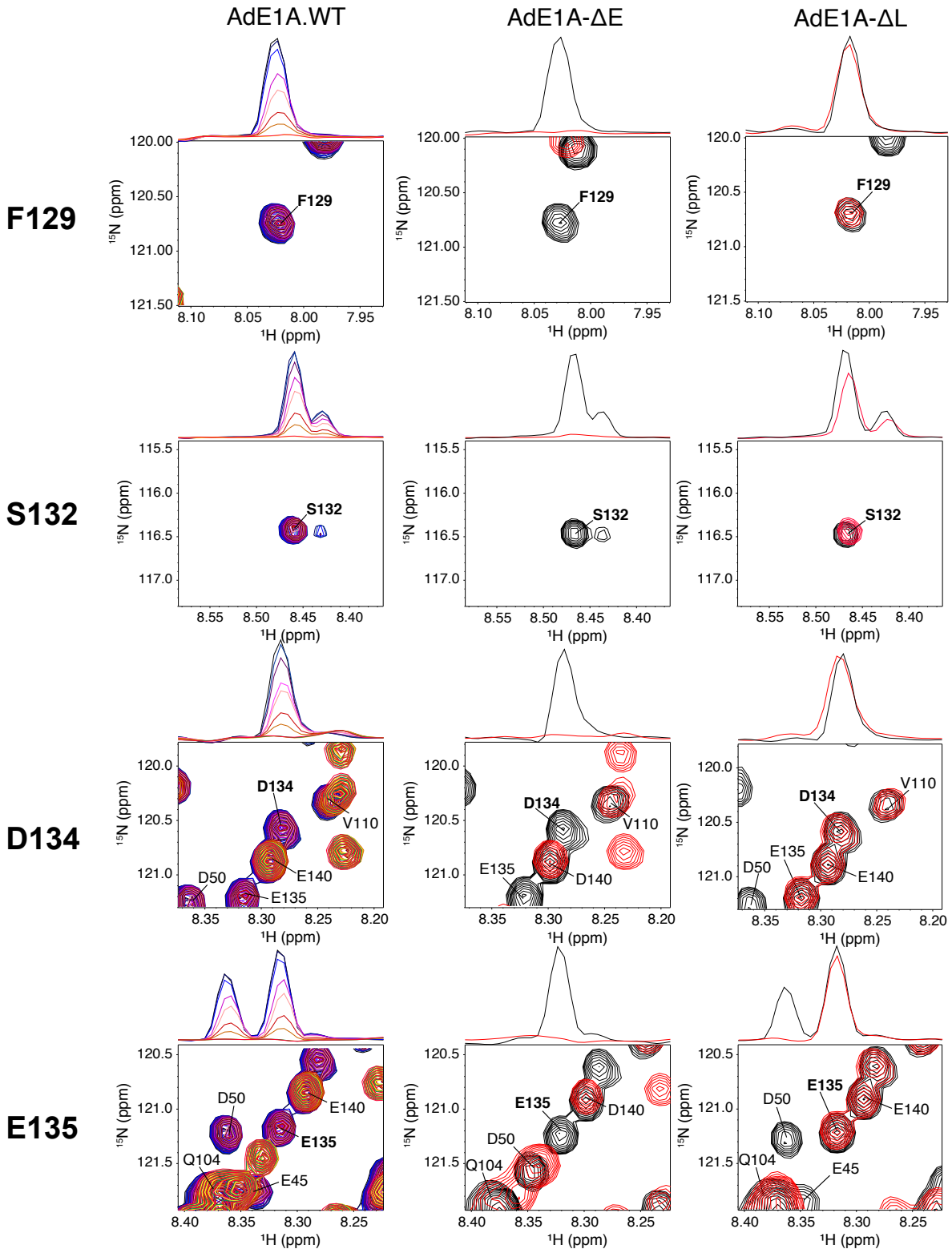


Figura A3.11. Detalle de los espectros TROSY para los residuos flanqueantes al motivo LxCxE y sus proyecciones de intensidad de resonancia obtenidos punto a punto (código de colores en la Figura 3.13) en el experimento de titulación para AdE1A.WT (primer columna) y para las formas libres (en negro) y acomplejadas en relación molar 1:1 con RbAB (en rojo) de la mutante AdE1A-ΔE (segunda columna) y de la mutante AdE1A-ΔL (tercer columna).

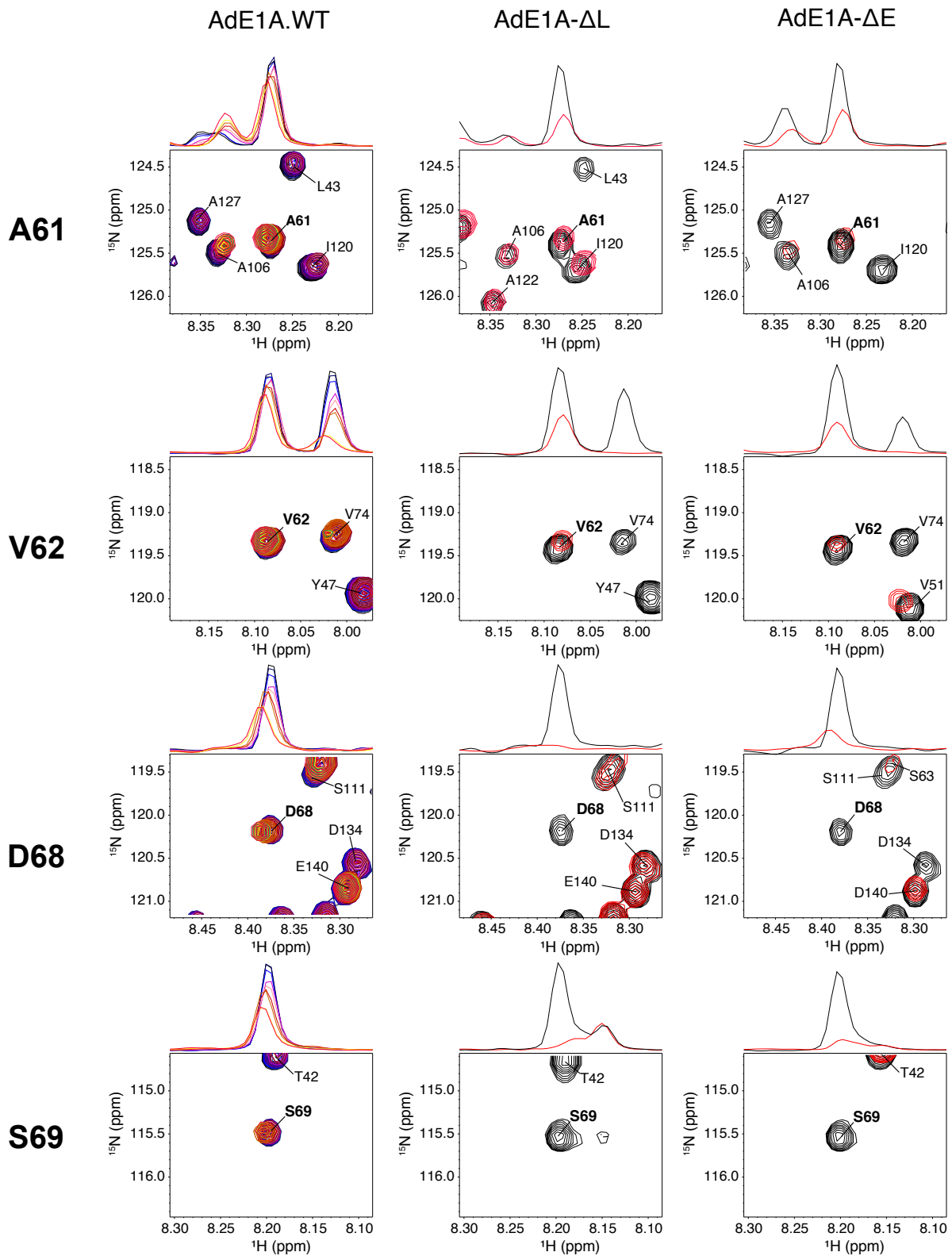


Figura A3.12. Detalle de los espectros TROSY para los residuos de la región linker entre I65 y F83 y sus proyecciones de intensidad de resonancia obtenidos punto a punto (código de colores en la Figura 3.13) en el experimento de titulación para AdE1A.WT (primer columna) y para las formas libres (en negro) y acomplejadas en relación molar 1:1 con RbAB (en rojo) de la mutante AdE1A-ΔL (segunda columna) y de la mutante AdE1A-ΔE (tercer columna).

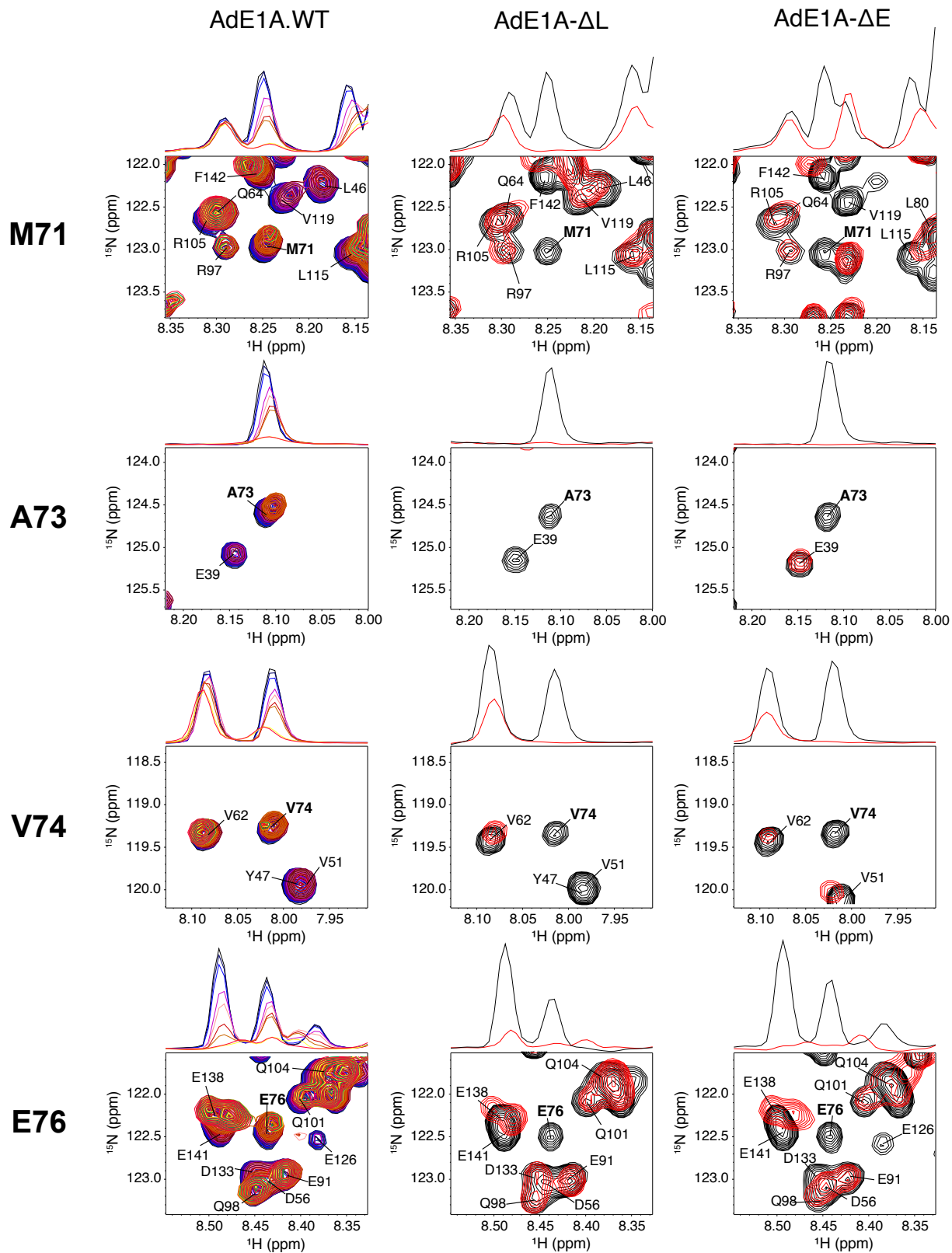


Figura A3.13. Detalle de los espectros TROSY para los residuos de la región linker entre I65 y F83 y sus proyecciones de intensidad de resonancia obtenidos punto a punto (código de colores en la Figura 3.13) en el experimento de titulación para AdE1A.WT (primer columna) y para las formas libres (en negro) y acomplejadas en relación molar 1:1 con RbAB (en rojo) de la mutante AdE1A-ΔL (segunda columna) y de la mutante AdE1A-ΔE (tercer columna).

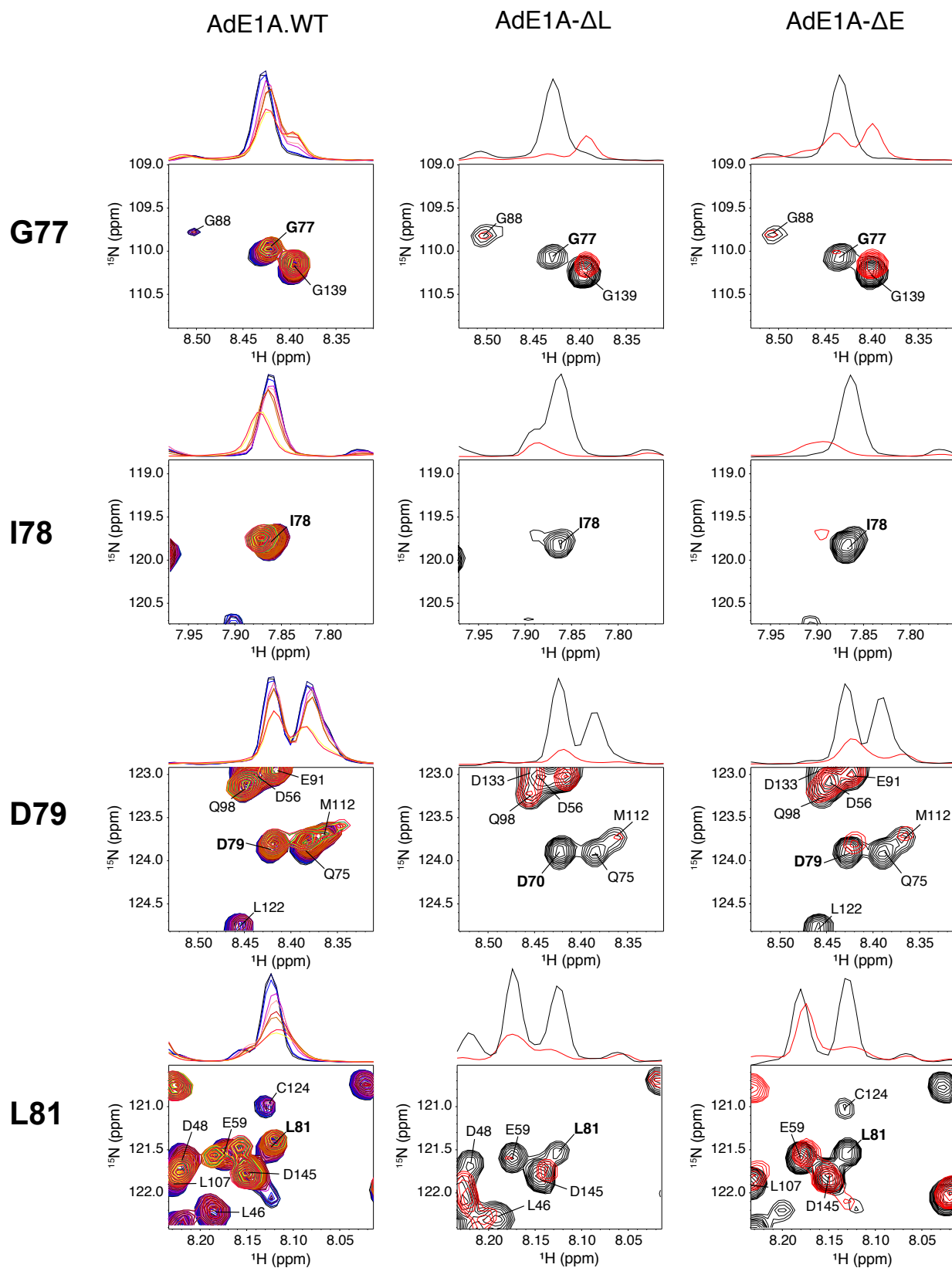


Figura A3.14. Detalle de los espectros TROSY para los residuos de la región linker entre I65 y F83 y sus proyecciones de intensidad de resonancia obtenidos punto a punto (código de colores en la Figura 3.13) en el experimento de titulación para AdE1A.WT (primer columna) y para las formas libres (en negro) y acomplejadas en relación molar 1:1 con RbAB (en rojo) de la mutante AdE1A-ΔL (segunda columna) y de la mutante AdE1A-ΔE (tercer columna).

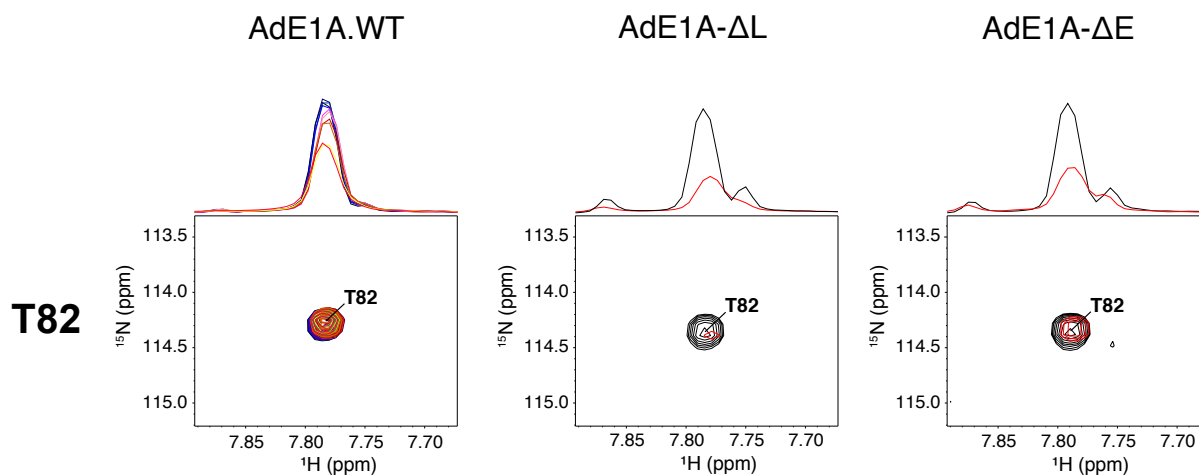


Figura A3.15. Detalle de los espectros TROSY para el residuo T82 de la región linker entre I65 y F83 y sus proyecciones de intensidad de resonancia obtenidos punto a punto (código de colores en la Figura 3.13) en el experimento de titulación para AdE1A.WT (primer columna) y para las formas libres (en negro) y acomplejadas en relación molar 1:1 con RbAB (en rojo) de la mutante AdE1A- Δ L (segunda columna) y de la mutante AdE1A- Δ E (tercer columna).

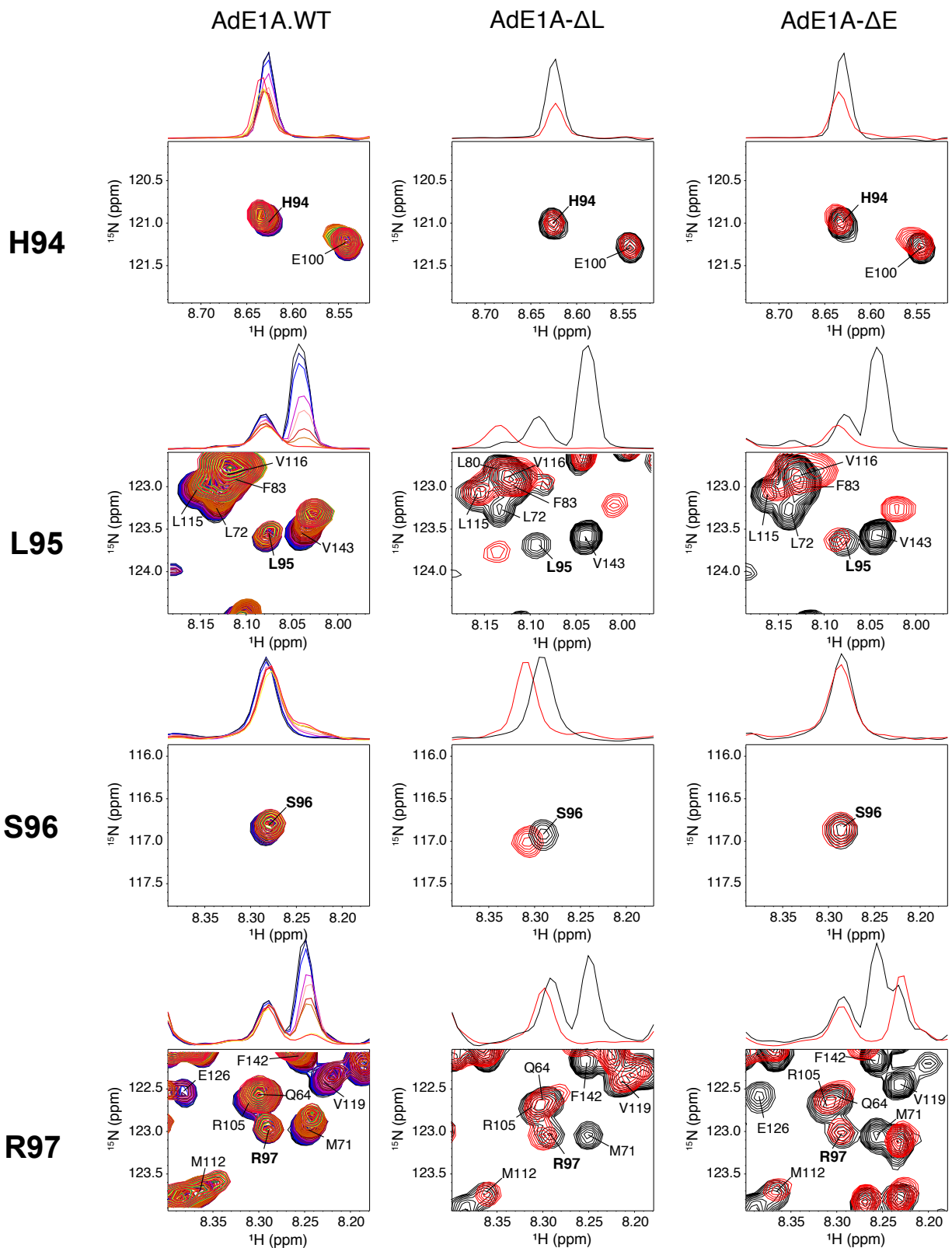


Figura A3.16. Detalle de los espectros TROSY para residuos de la región linker hacia el C-terminal y sus proyecciones de intensidad de resonancia obtenidos punto a punto (código de colores en la Figura 3.13) en el experimento de titulación para AdE1A.WT (primer columna) y para las formas libres (en negro) y acomplejadas en relación molar 1:1 con RbAB (en rojo) de la mutante AdE1A-ΔL (segunda columna) y de la mutante AdE1A-ΔE (tercer columna).

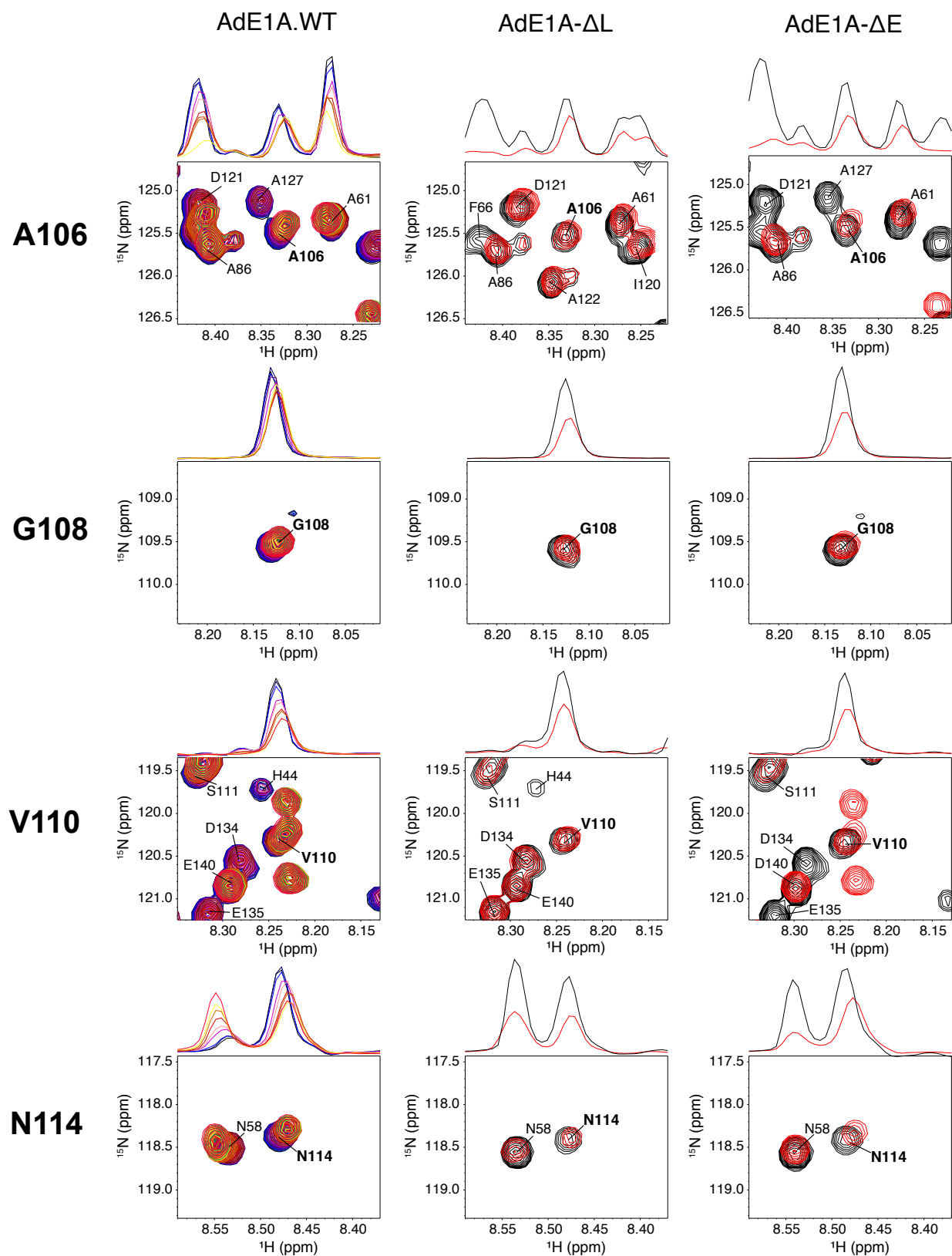


Figura A3.17. Detalle de los espectros TROSY para residuos de la región linker hacia el C-terminal y sus proyecciones de intensidad de resonancia obtenidos punto a punto (código de colores en la Figura 3.13) en el experimento de titulación para AdE1A.WT (primer columna) y para las formas libres (en negro) y acomplejadas en relación molar 1:1 con RbAB (en rojo) de la mutante AdE1A-ΔL (segunda columna) y de la mutante AdE1A-ΔE (tercer columna).

REFERENCIAS

1. Dyer, M.D., T.M. Murali, and B.W. Sobral, *The landscape of human proteins interacting with viruses and other pathogens*. PLoS Pathog, 2008. **4**(2): p. e32.
2. Goodacre, N., et al., *Protein-protein interactions of human viruses*. Semin Cell Dev Biol, 2018.
3. Brito, A.F. and J.W. Pinney, *Protein-Protein Interactions in Virus-Host Systems*. Front Microbiol, 2017. **8**: p. 1557.
4. King, C.R., et al., *Hacking the Cell: Network Intrusion and Exploitation by Adenovirus E1A*. MBio, 2018. **9**(3).
5. Davey, N.E., G. Trave, and T.J. Gibson, *How viruses hijack cell regulation*. Trends Biochem Sci, 2011. **36**(3): p. 159-69.
6. Kim, P.M., et al., *The role of disorder in interaction networks: a structural analysis*. Mol Syst Biol, 2008. **4**: p. 179.
7. Franzosa, E.A. and Y. Xia, *Structural principles within the human-virus protein-protein interaction network*. Proc Natl Acad Sci U S A, 2011. **108**(26): p. 10538-43.
8. Holmes, E.C., *Error thresholds and the constraints to RNA virus evolution*. Trends Microbiol, 2003. **11**(12): p. 543-6.
9. Drake, J.W., et al., *Rates of spontaneous mutation*. Genetics, 1998. **148**(4): p. 1667-86.
10. Gunasekaran, K., et al., *Extended disordered proteins: targeting function with less scaffold*. Trends Biochem Sci, 2003. **28**(2): p. 81-5.
11. Xue, B., et al., *Structural disorder in viral proteins*. Chem Rev, 2014. **114**(13): p. 6880-911.
12. Fischer, E., *Einfluss der Configuration auf die Wirkung der Enzyme*. Ber Dt Chem Ges, 1894. **27**: p. 2985–2993.
13. Berman, H.M., et al., *The Protein Data Bank*. Nucleic Acids Res, 2000. **28**(1): p. 235-42.
14. Kriwacki, R.W., et al., *Structural studies of p21Waf1/Cip1/Sdi1 in the free and Cdk2-bound state: conformational disorder mediates binding diversity*. Proc Natl Acad Sci U S A, 1996. **93**(21): p. 11504-9.
15. Romero, P.O., Z.; Kissinger, C. R.; Villafranca, J. E.; Garner, E.; Guilliot, S.; Dunker, A. K, *Pacific Symposium on Biocomputing R98*. 1998, World Scientific Publishing Co.: London. p. 437.
16. Wright, P.E. and H.J. Dyson, *Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm*. J Mol Biol, 1999. **293**(2): p. 321-31.
17. Dunker, A.K., et al., *What's in a name? Why these proteins are intrinsically disordered: Why these proteins are intrinsically disordered*. Intrinsically Disord Proteins, 2013. **1**(1): p. e24157.
18. Ward, J.J., et al., *Prediction and functional analysis of native disorder in proteins from the three kingdoms of life*. J Mol Biol, 2004. **337**(3): p. 635-45.
19. Schad, E., P. Tompa, and H. Hegyi, *The relationship between proteome size, structural disorder and organism complexity*. Genome Biol, 2011. **12**(12): p. R120.
20. Pushker, R., et al., *Marked variability in the extent of protein disorder within and between viral families*. PLoS One, 2013. **8**(4): p. e60724.
21. Habchi, J., et al., *Introducing protein intrinsic disorder*. Chem Rev, 2014. **114**(13): p. 6561-88.
22. Schweitzer-Stenner, R., *Conformational propensities and residual structures in unfolded peptides and proteins*. Mol Biosyst, 2012. **8**(1): p. 122-33.
23. van der Lee, R., et al., *Classification of intrinsically disordered regions and proteins*. Chem Rev, 2014. **114**(13): p. 6589-631.
24. Xie, H., et al., *Functional anthology of intrinsic disorder. 1. Biological processes and functions of proteins with long disordered regions*. J Proteome Res, 2007. **6**(5): p. 1882-98.
25. Uversky, V.N.L., S. , *Instrumental Analysis of Intrinsically Disordered Proteins*. 2010, John Wiley & Sons: Hoboken, NJ.
26. Uversky, V.N., J.R. Gillespie, and A.L. Fink, *Why are "natively unfolded" proteins unstructured under physiologic conditions?* Proteins, 2000. **41**(3): p. 415-27.
27. Uversky, V.N., *Natively unfolded proteins: a point where biology waits for physics*. Protein Sci, 2002. **11**(4): p. 739-56.
28. Sickmeier, M., et al., *DisProt: the Database of Disordered Proteins*. Nucleic Acids Res, 2007. **35**(Database issue): p. D786-93.
29. Das, R.K., K.M. Ruff, and R.V. Pappu, *Relating sequence encoded information to form and function of intrinsically disordered proteins*. Curr Opin Struct Biol, 2015. **32**: p. 102-12.

30. Das, R.K. and R.V. Pappu, *Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues*. Proc Natl Acad Sci U S A, 2013. **110**(33): p. 13392-7.
31. Mao, A.H., N. Lyle, and R.V. Pappu, *Describing sequence-ensemble relationships for intrinsically disordered proteins*. Biochem J, 2013. **449**(2): p. 307-18.
32. Dyson, H.J. and P.E. Wright, *Intrinsically unstructured proteins and their functions*. Nat Rev Mol Cell Biol, 2005. **6**(3): p. 197-208.
33. Borgia, A., et al., *Extreme disorder in an ultrahigh-affinity protein complex*. Nature, 2018. **555**(7694): p. 61-66.
34. Dunker, A.K., et al., *Intrinsic disorder and protein function*. Biochemistry, 2002. **41**(21): p. 6573-82.
35. Tompa, P., *The interplay between structure and function in intrinsically unstructured proteins*. FEBS Lett, 2005. **579**(15): p. 3346-54.
36. Gsponer, J. and M.M. Babu, *The rules of disorder or why disorder rules*. Prog Biophys Mol Biol, 2009. **99**(2-3): p. 94-103.
37. Gouw, M., et al., *The eukaryotic linear motif resource - 2018 update*. Nucleic Acids Res, 2018. **46**(D1): p. D428-D434.
38. Galea, C.A., et al., *Regulation of cell division by intrinsically unstructured proteins: intrinsic flexibility, modularity, and signaling conduits*. Biochemistry, 2008. **47**(29): p. 7598-609.
39. Shajani, Z., M.T. Sykes, and J.R. Williamson, *Assembly of bacterial ribosomes*. Annu Rev Biochem, 2011. **80**: p. 501-26.
40. Holt, C., *Unfolded phosphopolypeptides enable soft and hard tissues to coexist in the same organism with relative ease*. Curr Opin Struct Biol, 2013. **23**(3): p. 420-5.
41. Bardwell, J.C. and U. Jakob, *Conditional disorder in chaperone action*. Trends Biochem Sci, 2012. **37**(12): p. 517-25.
42. Chemes, L.B., G. de Prat-Gay, and I.E. Sanchez, *Convergent evolution and mimicry of protein linear motifs in host-pathogen interactions*. Curr Opin Struct Biol, 2015. **32**: p. 91-101.
43. Dinkel, H., et al., *The eukaryotic linear motif resource ELM: 10 years and counting*. Nucleic Acids Res, 2014. **42**(Database issue): p. D259-66.
44. Stein, A. and P. Aloy, *Contextual specificity in peptide-mediated protein interactions*. PLoS One, 2008. **3**(7): p. e2524.
45. Chica, C., F. Diella, and T.J. Gibson, *Evidence for the concerted evolution between short linear protein motifs and their flanking regions*. PLoS One, 2009. **4**(7): p. e6052.
46. Neduva, V. and R.B. Russell, *Linear motifs: evolutionary interaction switches*. FEBS Lett, 2005. **579**(15): p. 3342-5.
47. Mohan, A., et al., *Analysis of molecular recognition features (MoRFs)*. J Mol Biol, 2006. **362**(5): p. 1043-59.
48. Borchers, W., et al., *Structural divergence is more extensive than sequence divergence for a family of intrinsically disordered proteins*. Proteins, 2013. **81**(10): p. 1686-98.
49. Zor, T., et al., *Roles of phosphorylation and helix propensity in the binding of the KIX domain of CREB-binding protein by constitutive (c-Myb) and inducible (CREB) activators*. J Biol Chem, 2002. **277**(44): p. 42241-8.
50. Espinoza-Fonseca, L.M., *Reconciling binding mechanisms of intrinsically disordered proteins*. Biochem Biophys Res Commun, 2009. **382**(3): p. 479-82.
51. Wright, P.E. and H.J. Dyson, *Linking folding and binding*. Curr Opin Struct Biol, 2009. **19**(1): p. 31-8.
52. Tompa, P. and M. Fuxreiter, *Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions*. Trends Biochem Sci, 2008. **33**(1): p. 2-8.
53. Charlier, C., et al., *Structure and Dynamics of an Intrinsically Disordered Protein Region That Partially Folds upon Binding by Chemical-Exchange NMR*. J Am Chem Soc, 2017. **139**(35): p. 12219-12227.
54. Milles, S., et al., *Plasticity of an ultrafast interaction between nucleoporins and nuclear transport receptors*. Cell, 2015. **163**(3): p. 734-45.
55. Yoon, M.K., et al., *Cell cycle regulation by the intrinsically disordered proteins p21 and p27*. Biochem Soc Trans, 2012. **40**(5): p. 981-8.

56. Akimoto, M., et al., *Signaling through dynamic linkers as revealed by PKA*. Proc Natl Acad Sci U S A, 2013. **110**(35): p. 14231-6.
57. Borchers, W., et al., *Optimal Affinity Enhancement by a Conserved Flexible Linker Controls p53 Mimicry in MdmX*. Biophys J, 2017. **112**(10): p. 2038-2042.
58. Chong, P.A., et al., *Coupling of tandem Smad ubiquitination regulatory factor (Smurf) WW domains modulates target specificity*. Proc Natl Acad Sci U S A, 2010. **107**(43): p. 18404-9.
59. Davison, A.J., M. Benko, and B. Harrach, *Genetic content and evolution of adenoviruses*. J Gen Virol, 2003. **84**(Pt 11): p. 2895-908.
60. Harrach, B., Benk., M., Both, G.W., Brown, M., Davison, A.J., Echavarr.a, M., et al. , *Family Adenoviridae*, in *Virus Taxonomy. Ninth Report. International Committee Taxon*, A.M. King, Adams, M.J., Carstens, E.B., Lefkowitz, E.J., editors, Editor. 2011, Elsevier: San Diego, CA. p. 125–141.
61. Khanal, S., P. Ghimire, and A.S. Dhamoon, *The Repertoire of Adenovirus in Human Disease: The Innocuous to the Deadly*. Biomedicines, 2018. **6**(1).
62. Han, G., et al., *Identification and typing of respiratory adenoviruses in Guangzhou, Southern China using a rapid and simple method*. Virol Sin, 2013. **28**(2): p. 103-8.
63. Burrell, C.J.M., F A., , in *Fenner and White's Medical Virology 2017*, Elsevier: San Diego, CA.
64. Varga, M.J., C. Weibull, and E. Everitt, *Infectious entry pathway of adenovirus type 2*. J Virol, 1991. **65**(11): p. 6061-70.
65. Chardonnet, Y. and S. Dales, *Early events in the interaction of adenoviruses with HeLa cells. I. Penetration of type 5 and intracellular release of the DNA genome*. Virology, 1970. **40**(3): p. 462-77.
66. Waye, M.M.Y., & Sing, C. W, *Anti-Viral Drugs for Human Adenoviruses*. Pharmaceuticals, 2010. **3**(10): p. 3343-3354.
67. Wang, Y., et al., *Phylogenetic evidence for intratypic recombinant events in a novel human adenovirus C that causes severe acute respiratory infection in children*. Sci Rep, 2016. **6**: p. 23014.
68. Huebner, R.J., W.P. Rowe, and W.T. Lane, *Oncogenic effects in hamsters of human adenovirus types 12 and 18*. Proc Natl Acad Sci U S A, 1962. **48**: p. 2051-8.
69. Kobayashi, S. and N. Mukai, *Retinoblastoma-like tumors induced in rats by human adenovirus*. Invest Ophthalmol, 1973. **12**(11): p. 853-6.
70. Williams, J.F., et al., *E1A-based determinants of oncogenicity in human adenovirus groups A and C*. Curr Top Microbiol Immunol, 2004. **273**: p. 245-88.
71. Braithwaite, A.W., et al., *Adenovirus-induced alterations of the cell growth cycle: a requirement for expression of E1A but not of E1B*. J Virol, 1983. **45**(1): p. 192-9.
72. Radko, S., et al., *Effects of Adenovirus Type 5 E1A Isoforms on Viral Replication in Arrested Human Cells*. PLoS One, 2015. **10**(10): p. e0140124.
73. Kimelman, D., et al., *E1a regions of the human adenoviruses and of the highly oncogenic simian adenovirus 7 are closely related*. J Virol, 1985. **53**(2): p. 399-409.
74. Perricaudet, M., et al., *Structure of two spliced mRNAs from the transforming region of human subgroup C adenoviruses*. Nature, 1979. **281**(5733): p. 694-6.
75. Glavina, J., et al., *Interplay between sequence, structure and linear motifs in the adenovirus E1A hub protein*. Virology, 2018. **525**: p. 117-131.
76. Ferreon, J.C., et al., *Structural basis for subversion of cellular control mechanisms by the adenoviral E1A oncoprotein*. Proc Natl Acad Sci U S A, 2009. **106**(32): p. 13260-5.
77. Pelka, P., et al., *Intrinsic structural disorder in adenovirus E1A: a viral molecular hub linking multiple diverse processes*. J Virol, 2008. **82**(15): p. 7252-63.
78. Krippel, B., et al., *Functions of purified E1A protein microinjected into mammalian cells*. Proc Natl Acad Sci U S A, 1984. **81**(22): p. 6988-92.
79. Romero, P., Z. Obradovic, and A.K. Dunker, *Natively disordered proteins: functions and predictions*. Appl Bioinformatics, 2004. **3**(2-3): p. 105-13.
80. Ward, J.J., et al., *The DISOPRED server for the prediction of protein disorder*. Bioinformatics, 2004. **20**(13): p. 2138-9.
81. Hosek, T., et al., *Structural and Dynamic Characterization of the Molecular Hub Early Region 1A (E1A) from Human Adenovirus*. Chemistry, 2016. **22**(37): p. 13010-3.

82. Haberz, P., et al., *Mapping the interactions of adenoviral E1A proteins with the p160 nuclear receptor coactivator binding domain of CBP*. Protein Sci, 2016. **25**(12): p. 2256-2267.
83. Molloy, D.P., et al., *Structural determinants in adenovirus 12 E1A involved in the interaction with C-terminal binding protein 1*. Virology, 2000. **277**(1): p. 156-66.
84. Liu, X. and R. Marmorstein, *Structure of the retinoblastoma protein bound to adenovirus E1A reveals the molecular basis for viral oncoprotein inactivation of a tumor suppressor*. Genes Dev, 2007. **21**(21): p. 2711-6.
85. Dosztanyi, Z., et al., *IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content*. Bioinformatics, 2005. **21**(16): p. 3433-4.
86. Moran, E., et al., *Lytic and transforming functions of individual products of the adenovirus E1A gene*. J Virol, 1986. **57**(3): p. 765-75.
87. Yee, S.P. and P.E. Branton, *Detection of cellular proteins associated with human adenovirus type 5 early region 1A polypeptides*. Virology, 1985. **147**(1): p. 142-53.
88. Egan, C., et al., *Mapping of cellular protein-binding sites on the products of early-region 1A of human adenovirus type 5*. Mol Cell Biol, 1988. **8**(9): p. 3955-9.
89. DeCaprio, J.A., *How the Rb tumor suppressor structure and function was revealed by the study of Adenovirus and SV40*. Virology, 2009. **384**(2): p. 274-84.
90. Lee, W.H., et al., *The retinoblastoma susceptibility gene encodes a nuclear phosphoprotein associated with DNA binding activity*. Nature, 1987. **329**(6140): p. 642-5.
91. Whyte, P., Buchkovich, K.J., Horowitz, J.M., Friend, S.H., Raybuck, M., Weinberg, R.A., Harlow, E., , *Association between an oncogene and an anti-oncogene: the adenovirus E1A proteins bind to the retinoblastoma gene product*. Nature, 1988. **334**(6178): p. 124-129.
92. Dyson, N., et al., *The cellular 107K protein that binds to adenovirus E1A also associates with the large T antigens of SV40 and JC virus*. Cell, 1989. **58**(2): p. 249-55.
93. Cobrinik, D., et al., *Cell cycle-specific association of E2F with the p130 E1A-binding protein*. Genes Dev, 1993. **7**(12A): p. 2392-404.
94. Weinberg, R.A., *The retinoblastoma protein and cell cycle control*. Cell, 1995. **81**(3): p. 323-30.
95. Dick, F.A. and S.M. Rubin, *Molecular mechanisms underlying RB protein function*. Nat Rev Mol Cell Biol, 2013. **14**(5): p. 297-306.
96. Frolov, M.V. and N.J. Dyson, *Molecular mechanisms of E2F-dependent activation and pRB-mediated repression*. J Cell Sci, 2004. **117**(Pt 11): p. 2173-81.
97. Cobrinik, D., *Pocket proteins and cell cycle control*. Oncogene, 2005. **24**(17): p. 2796-809.
98. Liban, T.J., et al., *Structural Conservation and E2F Binding Specificity within the Retinoblastoma Pocket Protein Family*. J Mol Biol, 2016. **428**(20): p. 3960-3971.
99. Rubin, S.M., *Deciphering the retinoblastoma protein phosphorylation code*. Trends Biochem Sci, 2013. **38**(1): p. 12-9.
100. Chemes, L.B., et al., *Targeting mechanism of the retinoblastoma tumor suppressor by a prototypical viral oncoprotein. Structural modularity, intrinsic disorder and phosphorylation of human papillomavirus E7*. FEBS J, 2010. **277**(4): p. 973-88.
101. Rubin, S.M., et al., *Structure of the Rb C-terminal domain bound to E2F1-DP1: a mechanism for phosphorylation-induced E2F release*. Cell, 2005. **123**(6): p. 1093-106.
102. Dick, F.A., *Structure-function analysis of the retinoblastoma tumor suppressor protein - is the whole a sum of its parts?* Cell Div, 2007. **2**: p. 26.
103. Xiao, B., et al., *Crystal structure of the retinoblastoma tumor suppressor protein bound to E2F and the molecular basis of its regulation*. Proc Natl Acad Sci U S A, 2003. **100**(5): p. 2363-8.
104. Burkhart, D.L. and J. Sage, *Cellular mechanisms of tumour suppression by the retinoblastoma gene*. Nat Rev Cancer, 2008. **8**(9): p. 671-82.
105. Ferrari, R., et al., *Epigenetic reprogramming by adenovirus e1a*. Science, 2008. **321**(5892): p. 1086-8.
106. de Souza, R.F., L.M. Iyer, and L. Aravind, *Diversity and evolution of chromatin proteins encoded by DNA viruses*. Biochim Biophys Acta, 2010. **1799**(3-4): p. 302-18.
107. Helt, A.M. and D.A. Galloway, *Mechanisms by which DNA tumor virus oncoproteins target the Rb family of pocket proteins*. Carcinogenesis, 2003. **24**(2): p. 159-69.
108. Dyson, N., et al., *Adenovirus E1A makes two distinct contacts with the retinoblastoma protein*. J Virol, 1992. **66**(7): p. 4606-11.

109. Chemes, L.B., I.E. Sanchez, and G. de Prat-Gay, *Kinetic recognition of the retinoblastoma tumor suppressor by a specific protein target*. J Mol Biol, 2011. **412**(2): p. 267-84.
110. Sherry, K.P., et al., *Control of transcriptional activity by design of charge patterning in the intrinsically disordered RAM region of the Notch receptor*. Proc Natl Acad Sci U S A, 2017. **114**(44): p. E9243-E9252.
111. Zhou, H.X., *The affinity-enhancing roles of flexible linkers in two-domain DNA-binding proteins*. Biochemistry, 2001. **40**(50): p. 15069-73.
112. Bertagna, A., et al., *The effects of conformational heterogeneity on the binding of the Notch intracellular domain to effector proteins: a case of biologically tuned disorder*. Biochem Soc Trans, 2008. **36**(Pt 2): p. 157-66.
113. Di Fiore, B., et al., *The Mitotic Checkpoint Complex Requires an Evolutionary Conserved Cassette to Bind and Inhibit Active APC/C*. Mol Cell, 2016. **64**(6): p. 1144-1153.
114. Clements, A., et al., *Oligomerization properties of the viral oncoproteins adenovirus E1A and human papillomavirus E7 and their complexes with the retinoblastoma protein*. Biochemistry, 2000. **39**(51): p. 16033-45.
115. Alonso, L.G., et al., *High-risk (HPV16) human papillomavirus E7 oncoprotein is highly stable and extended, with conformational transitions that could explain its multiple cellular binding partners*. Biochemistry, 2002. **41**(33): p. 10510-8.
116. *Short protocols in molecular biology, in A compendium of methods from "current protocols in molecular biology"*, F.B. Ausubel, R; Kingston, R.E.; Moore, D.D.; Seidman, J.G.; Smith, J.A.; Struhl, K., Editor. 1995, John Wiley & Sons: New York, NY.
117. Marsh, J.A. and J.D. Forman-Kay, *Sequence determinants of compaction in intrinsically disordered proteins*. Biophys J, 2010. **98**(10): p. 2383-90.
118. Wilkins, D.K., et al., *Hydrodynamic radii of native and denatured proteins measured by pulse field gradient NMR techniques*. Biochemistry, 1999. **38**(50): p. 16424-31.
119. Tcherkasskaya, O. and V.N. Uversky, *Denatured collapsed states in protein folding: example of apomyoglobin*. Proteins, 2001. **44**(3): p. 244-54.
120. Uversky, V.N., *What does it mean to be natively unfolded?* Eur J Biochem, 2002. **269**(1): p. 2-12.
121. Gast, K.M., A.J., *Studying Protein Folding and Aggregation by Laser Light Scattering*, in *Protein Folding Handbook*, J.K. Buchner, T., Editor. 2008, Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim.
122. Chemes, L.B., et al., *Circular dichroism techniques for the analysis of intrinsically disordered proteins and domains*. Methods Mol Biol, 2012. **895**: p. 387-404.
123. Whitmore, L. and B.A. Wallace, *Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases*. Biopolymers, 2008. **89**(5): p. 392-400.
124. Lehrman, S.R., J.L. Tuls, and M. Lund, *Peptide alpha-helicity in aqueous trifluoroethanol: correlations with predicted alpha-helicity and the secondary structure of the corresponding regions of bovine growth hormone*. Biochemistry, 1990. **29**(23): p. 5590-6.
125. Jasanoff, A. and A.R. Fersht, *Quantitative determination of helical propensities from trifluoroethanol titration curves*. Biochemistry, 1994. **33**(8): p. 2129-35.
126. Munoz, V. and L. Serrano, *Elucidating the folding problem of helical peptides using empirical parameters. III. Temperature and pH dependence*. J Mol Biol, 1995. **245**(3): p. 297-308.
127. Zhu, F., et al., *Residual structure in disordered peptides and unfolded proteins from multivariate analysis and ab initio simulation of Raman optical activity data*. Proteins, 2008. **70**(3): p. 823-33.
128. Shi, Z., et al., *Conformation of the backbone in unfolded proteins*. Chem Rev, 2006. **106**(5): p. 1877-97.
129. Shi, Z., R.W. Woody, and N.R. Kallenbach, *Is polyproline II a major backbone conformation in unfolded proteins?* Adv Protein Chem, 2002. **62**: p. 163-240.
130. Adzhubei, A.A., M.J. Sternberg, and A.A. Makarov, *Polyproline-II helix in proteins: structure and function*. J Mol Biol, 2013. **425**(12): p. 2100-32.
131. Uversky, V.N., *A protein-chameleon: conformational plasticity of alpha-synuclein, a disordered protein involved in neurodegenerative disorders*. J Biomol Struct Dyn, 2003. **21**(2): p. 211-34.
132. Tompa, P., *Intrinsically unstructured proteins*. Trends Biochem Sci, 2002. **27**(10): p. 527-33.

133. Chemes, L.B., et al., *Folding of a cyclin box: linking multitarget binding to marginal stability, oligomerization, and aggregation of the retinoblastoma tumor suppressor AB pocket domain*. J Biol Chem, 2013. **288**(26): p. 18923-38.
134. Tompa, P., et al., *A million peptide motifs for the molecular biologist*. Mol Cell, 2014. **55**(2): p. 161-9.
135. Lee, J.O., A.A. Russo, and N.P. Pavletich, *Structure of the retinoblastoma tumour-suppressor pocket domain bound to a peptide from HPV E7*. Nature, 1998. **391**(6670): p. 859-65.
136. Lee, C., et al., *Structural basis for the recognition of the E2F transactivation domain by the retinoblastoma tumor suppressor*. Genes Dev, 2002. **16**(24): p. 3199-212.
137. Morris, E.J. and N.J. Dyson, *Retinoblastoma protein partners*. Adv Cancer Res, 2001. **82**: p. 1-54.
138. Defeo-Jones, D., et al., *Cloning of cDNAs for cellular proteins that bind to the retinoblastoma gene product*. Nature, 1991. **352**(6332): p. 251-4.
139. Fattaey, A.R., et al., *Characterization of the retinoblastoma binding proteins RBP1 and RBP2*. Oncogene, 1993. **8**(11): p. 3149-56.
140. Brehm, A., et al., *Retinoblastoma protein recruits histone deacetylase to repress transcription*. Nature, 1998. **391**(6667): p. 597-601.
141. Singh, M., et al., *Molecular determinants for the complex formation between the retinoblastoma protein and LXCXE sequences*. J Biol Chem, 2005. **280**(45): p. 37868-76.
142. Vortmeyer, A.O., et al., *Multiple endocrine neoplasia type 1: atypical presentation, clinical course, and genetic analysis of multiple tumors*. Mod Pathol, 1999. **12**(9): p. 919-24.
143. Zheng, L., et al., *Retinoblastoma protein enhances the fidelity of chromosome segregation mediated by hHec1p*. Mol Cell Biol, 2000. **20**(10): p. 3529-37.
144. Ciferri, C., et al., *Implications for kinetochore-microtubule attachment from the structure of an engineered Ndc80 complex*. Cell, 2008. **133**(3): p. 427-39.
145. Tevosian, S.G., et al., *HBP1: a HMG box transcriptional repressor that is targeted by the retinoblastoma family*. Genes Dev, 1997. **11**(3): p. 383-96.
146. Lavender, P., et al., *The HMG-box transcription factor HBP1 is targeted by the pocket proteins and E1A*. Oncogene, 1997. **14**(22): p. 2721-8.
147. MacLellan, W.R., et al., *A novel Rb- and p300-binding protein inhibits transactivation by MyoD*. Mol Cell Biol, 2000. **20**(23): p. 8903-15.
148. Laurita, K.R. and D.S. Rosenbaum, *Interdependence of modulated dispersion and tissue structure in the mechanism of unidirectional block*. Circ Res, 2000. **87**(10): p. 922-8.
149. Buyse, I.M., G. Shao, and S. Huang, *The retinoblastoma protein binds to RIZ, a zinc-finger protein that shares an epitope with the adenovirus E1A protein*. Proc Natl Acad Sci U S A, 1995. **92**(10): p. 4467-71.
150. Sun, Y., et al., *Structural and functional characterization of the acidic region from the RIZ tumor suppressor*. Biochemistry, 2015. **54**(6): p. 1390-400.
151. Wang, Y., et al., *A Novel Retinoblastoma Protein (RB) E3 Ubiquitin Ligase (NRBE3) Promotes RB Degradation and Is Transcriptionally Regulated by E2F1 Transcription Factor*. J Biol Chem, 2015. **290**(47): p. 28200-13.
152. Strober, B.E., et al., *Functional interactions between the hBRM/hBRG1 transcriptional activators and the pRB family of proteins*. Mol Cell Biol, 1996. **16**(4): p. 1576-83.
153. Trouche, D., et al., *RB and hbrm cooperate to repress the activation functions of E2F1*. Proc Natl Acad Sci U S A, 1997. **94**(21): p. 11268-73.
154. Dunaief, J.L., et al., *The retinoblastoma protein and BRG1 form a complex and cooperate to induce cell cycle arrest*. Cell, 1994. **79**(1): p. 119-30.
155. Day, P.J., et al., *Crystal structure of human CDK4 in complex with a D-type cyclin*. Proc Natl Acad Sci U S A, 2009. **106**(11): p. 4166-70.
156. Dowdy, S.F., et al., *Physical interaction of the retinoblastoma protein with human D cyclins*. Cell, 1993. **73**(3): p. 499-511.
157. Ewen, M.E., et al., *Functional interactions of the retinoblastoma protein with mammalian D-type cyclins*. Cell, 1993. **73**(3): p. 487-97.
158. Kato, J., et al., *Direct binding of cyclin D to the retinoblastoma gene product (pRb) and pRb phosphorylation by the cyclin D-dependent kinase CDK4*. Genes Dev, 1993. **7**(3): p. 331-42.

159. Richardson, B.P. and B. Berde, *Proceedings: The effects of cyproheptadine pretreatment on insulin release from isolated pancreatic islets*. Br J Pharmacol, 1976. **56**(3): p. 356P-357P.
160. Kadri, Z., et al., *Direct binding of pRb/E2F-2 to GATA-1 regulates maturation and terminal cell division during erythropoiesis*. PLoS Biol, 2009. **7**(6): p. e1000123.
161. deBaare, L., J. Lewis, and H. Sing, *Ultramicroscale determination of clinical chemical values for blood during the first four days of postnatal life*. Clin Chem, 1975. **21**(6): p. 746-50.
162. DeCaprio, J.A., et al., *SV40 large tumor antigen forms a specific complex with the product of the retinoblastoma susceptibility gene*. Cell, 1988. **54**(2): p. 275-83.
163. Borchert, S., et al., *High-affinity Rb binding, p53 inhibition, subcellular localization, and transformation by wild-type or tumor-derived shortened Merkel cell polyomavirus large T antigens*. J Virol, 2014. **88**(6): p. 3144-60.
164. Houben, R., et al., *An intact retinoblastoma protein-binding site in Merkel cell polyomavirus large T antigen is required for promoting growth of Merkel cell carcinoma cells*. Int J Cancer, 2012. **130**(4): p. 847-56.
165. Phelps, W.C., et al., *Structure-function analysis of the human papillomavirus type 16 E7 oncoprotein*. J Virol, 1992. **66**(4): p. 2418-27.
166. Corbeil, H.B. and P.E. Branton, *Functional importance of complex formation between the retinoblastoma tumor suppressor family and adenovirus E1A proteins as determined by mutational analysis of E1A conserved region 2*. J Virol, 1994. **68**(10): p. 6697-709.
167. Prichard, M.N., et al., *Human cytomegalovirus UL97 kinase activity is required for the hyperphosphorylation of retinoblastoma protein and inhibits the formation of nuclear aggresomes*. J Virol, 2008. **82**(10): p. 5054-67.
168. Ran, X., et al., *White spot syndrome virus IE1 and WSV056 modulate the G1/S transition by binding to the host retinoblastoma protein*. J Virol, 2013. **87**(23): p. 12576-82.
169. Mohr, S., et al., *Targeting the retinoblastoma protein by MC007L, gene product of the molluscum contagiosum virus: detection of a novel virus-cell interaction by a member of the poxviruses*. J Virol, 2008. **82**(21): p. 10625-33.
170. Liu, L., et al., *Bean yellow dwarf virus RepA, but not rep, binds to maize retinoblastoma protein, and the virus tolerates mutations in the consensus binding motif*. Virology, 1999. **256**(2): p. 270-9.
171. Xie, Q., P. Suarez-Lopez, and C. Gutierrez, *Identification and analysis of a retinoblastoma binding motif in the replication protein of a plant DNA virus: requirement for efficient viral DNA replication*. EMBO J, 1995. **14**(16): p. 4073-82.
172. McGivern, D.R., et al., *An intact RBR-binding motif is not required for infectivity of Maize streak virus in cereals, but is required for invasion of mesophyll cells*. J Gen Virol, 2005. **86**(Pt 3): p. 797-801.
173. Aronson, M.N., et al., *Clink, a nanovirus-encoded protein, binds both pRB and SKP1*. J Virol, 2000. **74**(7): p. 2967-72.
174. Wanitchakorn, R., et al., *Functional analysis of proteins encoded by banana bunchy top virus DNA-4 to -6*. J Gen Virol, 2000. **81**(Pt 1): p. 299-306.
175. Guiley, K.Z., et al., *Structural mechanisms of DREAM complex assembly and regulation*. Genes Dev, 2015. **29**(9): p. 961-74.
176. Flemington, E.K., S.H. Speck, and W.G. Kaelin, Jr., *E2F-1-mediated transactivation is inhibited by complex formation with the retinoblastoma susceptibility gene product*. Proc Natl Acad Sci U S A, 1993. **90**(15): p. 6914-8.
177. Helin, K., E. Harlow, and A. Fattaey, *Inhibition of E2F-1 transactivation by direct binding of the retinoblastoma protein*. Mol Cell Biol, 1993. **13**(10): p. 6501-8.
178. Burke, J.R., G.L. Hura, and S.M. Rubin, *Structures of inactive retinoblastoma protein reveal multiple mechanisms for cell cycle control*. Genes Dev, 2012. **26**(11): p. 1156-66.
179. Burke, J.R., et al., *Phosphorylation-induced conformational changes in the retinoblastoma protein inhibit E2F transactivation domain binding*. J Biol Chem, 2010. **285**(21): p. 16286-93.
180. Ferreon, A.C., et al., *Modulation of allostery by protein intrinsic disorder*. Nature, 2013. **498**(7454): p. 390-4.
181. Jones, R.E., et al., *Identification of HPV-16 E7 peptides that are potent antagonists of E7 binding to the retinoblastoma suppressor protein*. J Biol Chem, 1990. **265**(22): p. 12782-5.

182. Palopoli, N., et al., *Short linear motif core and flanking regions modulate retinoblastoma protein binding affinity and specificity*. Protein Eng Des Sel, 2018. **31**(3): p. 69-77.
183. Kim, H.Y., B.Y. Ahn, and Y. Cho, *Structural basis for the inactivation of retinoblastoma tumor suppressor by SV40 large T antigen*. EMBO J, 2001. **20**(1-2): p. 295-304.
184. Fattaey, A.R., E. Harlow, and K. Helin, *Independent regions of adenovirus E1A are required for binding to and dissociation of E2F-protein complexes*. Mol Cell Biol, 1993. **13**(12): p. 7267-77.
185. Whalen, S.G., et al., *Importance of the Ser-132 phosphorylation site in cell transformation and apoptosis induced by the adenovirus type 5 E1A protein*. J Virol, 1996. **70**(8): p. 5373-83.
186. Schrama, D., et al., *Serine 220 phosphorylation of the Merkel cell polyomavirus large T antigen crucially supports growth of Merkel cell carcinoma cells*. Int J Cancer, 2016. **138**(5): p. 1153-62.
187. Liban, T.J., et al., *Conservation and divergence of C-terminal domain structure in the retinoblastoma protein family*. Proc Natl Acad Sci U S A, 2017. **114**(19): p. 4942-4947.
188. Inoue, Y., M. Kitagawa, and Y. Taya, *Phosphorylation of pRB at Ser612 by Chk1/2 leads to a complex between pRB and E2F-1 after DNA damage*. EMBO J, 2007. **26**(8): p. 2083-93.
189. Hensey, C.E., et al., *Identification of discrete structural domains in the retinoblastoma protein. Amino-terminal domain is required for its oligomerization*. J Biol Chem, 1994. **269**(2): p. 1380-7.
190. Rogers, J.M., A. Steward, and J. Clarke, *Folding and binding of an intrinsically disordered protein: fast, but not 'diffusion-limited'*. J Am Chem Soc, 2013. **135**(4): p. 1415-22.
191. Balog, E.R., et al., *Crystal structure of the unliganded retinoblastoma protein pocket domain*. Proteins, 2011. **79**(6): p. 2010-4.
192. Perozzo, R., G. Folkers, and L. Scapozza, *Thermodynamics of protein-ligand interactions: history, presence, and future aspects*. J Recept Signal Transduct Res, 2004. **24**(1-2): p. 1-52.
193. Van Roey, K., et al., *Short linear motifs: ubiquitous and functionally diverse protein interaction modules directing cell regulation*. Chem Rev, 2014. **114**(13): p. 6733-78.
194. Ward, L.D., *Measurement of ligand binding to proteins by fluorescence spectroscopy*. Methods Enzymol, 1985. **117**: p. 400-14.
195. Duff, M.R., Jr., J. Grubbs, and E.E. Howell, *Isothermal titration calorimetry for measuring macromolecule-ligand affinity*. J Vis Exp, 2011(55).
196. Sahu, D., et al., *Assessing Coupled Protein Folding and Binding Through Temperature-Dependent Isothermal Titration Calorimetry*. Methods Enzymol, 2016. **567**: p. 23-45.
197. Cooper, A., *Protein Heat Capacity: An Anomaly that Maybe Never Was*. Journal of Physical Chemistry Letters, 2010. **1**(22): p. 3298-3304.
198. Spolar, R.S. and M.T. Record, Jr., *Coupling of local folding to site-specific binding of proteins to DNA*. Science, 1994. **263**(5148): p. 777-84.
199. Dill, K.A., *Dominant forces in protein folding*. Biochemistry, 1990. **29**(31): p. 7133-55.
200. Lostao, A., et al., *Dissecting the energetics of the apoflavodoxin-FMN complex*. J Biol Chem, 2000. **275**(13): p. 9518-26.
201. Whalen, S.G., et al., *Phosphorylation within the transactivation domain of adenovirus E1A protein by mitogen-activated protein kinase regulates expression of early region 4*. J Virol, 1997. **71**(5): p. 3545-53.
202. Mal, A., A. Piotrkowski, and M.L. Harter, *Cyclin-dependent kinases phosphorylate the adenovirus E1A protein, enhancing its ability to bind pRb and disrupt pRb-E2F complexes*. J Virol, 1996. **70**(5): p. 2911-21.
203. Garcia-Alai, M.M., L.G. Alonso, and G. de Prat-Gay, *The N-terminal module of HPV16 E7 is an intrinsically disordered domain that confers conformational and recognition plasticity to the oncoprotein*. Biochemistry, 2007. **46**(37): p. 10405-12.
204. Prestel, A., et al., *Characterization of Dynamic IDP Complexes by NMR Spectroscopy*. Methods Enzymol, 2018. **611**: p. 193-226.
205. Konrat, R., *Interpreting protein chemical shift data*. Progress in Nuclear Magnetic Resonance Spectroscopy 2014. **58**: p. 62-87.
206. Showalter, S.A., *Intrinsically Disordered Proteins: Methods for Structure and Dynamics Studies*. eMagRes, 2014. **3**: p. 181-190.
207. Kosol, S., et al., *Structural characterization of intrinsically disordered proteins by NMR spectroscopy*. Molecules, 2013. **18**(9): p. 10802-28.

208. Dyson, H.J. and P.E. Wright, *Insights into the structure and dynamics of unfolded proteins from nuclear magnetic resonance*. Adv Protein Chem, 2002. **62**: p. 311-40.
209. Alexandrescu, A.T.E., P.A.; Pitkeathly, M.; Baum, J.; Dobson, C.M. , *Structure and dynamic of the molten globule state of guinea pig a-lactalbumin: a two-dimensional NMR study*. Biochemistry, 1993. **32**: p. 1707-1718.
210. Wishart, D.S., *Interpreting protein chemical shift data*. Prog Nucl Magn Reson Spectrosc, 2011. **58**(1-2): p. 62-87.
211. Pascal, S., *NMR Primer: An HSQC-based Approach with Vector Animations*. 2008: IM Publications LLP.
212. Wishart, D.S. and D.A. Case, *Use of chemical shifts in macromolecular structure determination*. Methods Enzymol, 2001. **338**: p. 3-34.
213. Tamiola, K. and F.A. Mulder, *Using NMR chemical shifts to calculate the propensity for structural order and disorder in proteins*. Biochem Soc Trans, 2012. **40**(5): p. 1014-20.
214. Bai, Y., et al., *Structural and dynamic characterization of an unfolded state of poplar apo-plastocyanin formed under nondenaturing conditions*. Protein Sci, 2001. **10**(5): p. 1056-66.
215. Tamiola, K., B. Acar, and F.A. Mulder, *Sequence-specific random coil chemical shifts of intrinsically disordered proteins*. J Am Chem Soc, 2010. **132**(51): p. 18000-3.
216. Wishart, D.S. and B.D. Sykes, *Chemical shifts as a tool for structure determination*. Methods Enzymol, 1994. **239**: p. 363-92.
217. Wishart, D.S. and A.M. Nip, *Protein chemical shift analysis: a practical guide*. Biochem Cell Biol, 1998. **76**(2-3): p. 153-63.
218. Daughdrill, G.W., L.J. Hanely, and F.W. Dahlquist, *The C-terminal half of the anti-sigma factor FlgM contains a dynamic equilibrium solution structure favoring helical conformations*. Biochemistry, 1998. **37**(4): p. 1076-82.
219. Camilloni, C., et al., *Determination of secondary structure populations in disordered states of proteins using nuclear magnetic resonance chemical shifts*. Biochemistry, 2012. **51**(11): p. 2224-31.
220. Kim, Y., *NMR spin relaxation approaches for characterizing protein motions*. BioDesign, 2015. **3**(1): p. 14-25.
221. Kleckner, I.R. and M.P. Foster, *An introduction to NMR-based approaches for measuring protein dynamics*. Biochim Biophys Acta, 2011. **1814**(8): p. 942-68.
222. Kay, L.E., D.A. Torchia, and A. Bax, *Backbone dynamics of proteins as studied by ¹⁵N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease*. Biochemistry, 1989. **28**(23): p. 8972-9.
223. Maciejewski, M.W., et al., *Backbone dynamics and refined solution structure of the N-terminal domain of DNA polymerase beta. Correlation with DNA binding and dRP lyase activity*. J Mol Biol, 2000. **296**(1): p. 229-53.
224. Pagano, K., et al., *Direct and allosteric inhibition of the FGF2/HSPGs/FGFR1 ternary complex formation by an antiangiogenic, thrombospondin-1-mimic small molecule*. PLoS One, 2012. **7**(5): p. e36990.
225. Cabral, K.M., et al., *Biophysical Studies on BEX3, the p75NTR-Associated Cell Death Executor, Reveal a High-Order Oligomer with Partially Folded Regions*. PLoS One, 2015. **10**(9): p. e0137916.
226. Vise, P.D., et al., *NMR chemical shift and relaxation measurements provide evidence for the coupled folding and binding of the p53 transactivation domain*. Nucleic Acids Res, 2005. **33**(7): p. 2061-77.
227. Hellman, M., et al., *Characterization of intrinsically disordered prostate associated gene (PAGE5) at single residue resolution by NMR spectroscopy*. PLoS One, 2011. **6**(11): p. e26633.
228. Farina, B., et al., *NMR backbone dynamics studies of human PED/PEA-15 outline protein functional sites*. FEBS J, 2010. **277**(20): p. 4229-40.
229. Chiliveri, S.C. and M.V. Deshmukh, *Recent excitements in protein NMR: Large proteins and biologically relevant dynamics*. J Biosci, 2016. **41**(4): p. 787-803.
230. Pervushin, K., et al., *Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution*. Proc Natl Acad Sci U S A, 1997. **94**(23): p. 12366-71.

231. Fernandez, C. and G. Wider, *TROSY in NMR studies of the structure and function of large biological macromolecules*. *Curr Opin Struct Biol*, 2003. **13**(5): p. 570-80.
232. Pellecchia, M., D.S. Sem, and K. Wuthrich, *NMR in drug discovery*. *Nat Rev Drug Discov*, 2002. **1**(3): p. 211-9.
233. Bernado, P., et al., *A structural model for unfolded proteins from residual dipolar couplings and small-angle x-ray scattering*. *Proc Natl Acad Sci U S A*, 2005. **102**(47): p. 17002-7.
234. Tompa, P., *On the supertertiary structure of proteins*. *Nat Chem Biol*, 2012. **8**(7): p. 597-600.
235. Tria, G., et al., *Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering*. *IUCrJ*, 2015. **2**(Pt 2): p. 207-17.
236. Kachala, M., E. Valentini, and D.I. Svergun, *Application of SAXS for the Structural Characterization of IDPs*. *Adv Exp Med Biol*, 2015. **870**: p. 261-89.
237. Pace, C.N., et al., *How to measure and predict the molar absorption coefficient of a protein*. *Protein Sci*, 1995. **4**(11): p. 2411-23.
238. Aslam, M. and S.J. Perkins, *Folded-back solution structure of monomeric factor H of human complement by synchrotron X-ray and neutron scattering, analytical ultracentrifugation and constrained molecular modelling*. *J Mol Biol*, 2001. **309**(5): p. 1117-38.
239. Zhou, H.X., *Quantitative account of the enhanced affinity of two linked scFvs specific for different epitopes on the same antigen*. *J Mol Biol*, 2003. **329**(1): p. 1-8.
240. Zhou, H.X., *Polymer models of protein stability, folding, and interactions*. *Biochemistry*, 2004. **43**(8): p. 2141-54.
241. Dunker, A.K., et al., *Intrinsically disordered protein*. *J Mol Graph Model*, 2001. **19**(1): p. 26-59.
242. McClendon, S., C.C. Rospigliosi, and D. Eliezer, *Charge neutralization and collapse of the C-terminal tail of alpha-synuclein at low pH*. *Protein Sci*, 2009. **18**(7): p. 1531-40.
243. Oldfield, C.J., et al., *Coupled folding and binding with alpha-helix-forming molecular recognition elements*. *Biochemistry*, 2005. **44**(37): p. 12454-70.
244. Daughdrill, G.W., et al., *Dynamic behavior of an intrinsically unstructured linker domain is conserved in the face of negligible amino acid sequence conservation*. *J Mol Evol*, 2007. **65**(3): p. 277-88.
245. Toth-Petroczy, A., et al., *Structured States of Disordered Proteins from Genomic Sequences*. *Cell*, 2016. **167**(1): p. 158-170 e12.
246. Avvakumov, N., et al., *Comprehensive sequence analysis of the E1A proteins of human and simian adenoviruses*. *Virology*, 2004. **329**(2): p. 477-92.
247. Ansieau, S. and A. Leutz, *The conserved Mynd domain of BS69 binds cellular and oncoviral proteins through a common PXLXP motif*. *J Biol Chem*, 2002. **277**(7): p. 4906-10.
248. Sherry, K.P., et al., *Effects of Linker Length and Transient Secondary Structure Elements in the Intrinsically Disordered Notch RAM Region on Notch Signaling*. *J Mol Biol*, 2015. **427**(22): p. 3587-3597.
249. Bradford, M.M., *A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding*. *Anal Biochem*, 1976. **72**: p. 248-54.
250. Hermanson, G., *Bioconjugate Techniques*. First ed. ed. 1996, San Diego, CA: Academic Press.
251. Chen, Y.H., J.T. Yang, and K.H. Chau, *Determination of the helix and beta form of proteins in aqueous solution by circular dichroism*. *Biochemistry*, 1974. **13**(16): p. 3350-9.
252. Uversky, V.N., *Use of fast protein size-exclusion liquid chromatography to study the unfolding of proteins which denature through the molten globule*. *Biochemistry*, 1993. **32**(48): p. 13288-98.
253. de Gennes, P.-G., *Scaling Concepts in Polymer Physics*. 1979, Ithaca, NY: Cornell University Press.
254. Flory, P.J., *Principles of Polymer Chemistry*. 1953, Ithaca, NY: Cornell University Press.
255. Baldwin, R.L., *Temperature dependence of the hydrophobic interaction in protein folding*. *Proc Natl Acad Sci U S A*, 1986. **83**(21): p. 8069-72.
256. Murphy, K.P., et al., *Molecular basis of co-operativity in protein folding. III. Structural identification of cooperative folding units and folding intermediates*. *J Mol Biol*, 1992. **227**(1): p. 293-306.
257. Luque, I. and E. Freire, *Structural parameterization of the binding enthalpy of small ligands*. *Proteins*, 2002. **49**(2): p. 181-90.
258. Muhandiram, D.R.K., L.E. , *Gradient-enhanced triple-resonance three-dimensional NMR experiments with improved sensitivity*. *J. Magn. Reson*, 1994. **103**(3): p. 203-216.

259. Wittekind, M.M., L. , *HNCACB, a high-sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the alpha- and beta-carbon resonances in proteins* J. Magn. Reson, 1993. **101**: p. 201-205.
260. Kay, L.E.K., P.; Saarinen, T. , *Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity* J. Am. Chem. Soc, 1992. **114**(10663).
261. Johnson, B.A.R.B., R.A. , *NMRView: a computer program for the visualization and analysis of NMR data*. J. Biomol. NMR, 1994. **4**: p. 603-614.
262. Mulder, F.A., et al., *Altered flexibility in the substrate-binding site of related native and engineered high-alkaline Bacillus subtilisins*. J Mol Biol, 1999. **292**(1): p. 111-23.
263. Estaña, A.S., N.; Delaforge, E.; Vaisset, M.; Cortés, J.; Bernadó, P., *Realistic ensemble models of intrinsically disordered proteins using a structure-encoding coil database*. In press.
264. Cortes, J., et al., *Geometric algorithms for the conformational analysis of long protein loops*. J Comput Chem, 2004. **25**(7): p. 956-67.
265. Garcia De La Torre, J., M.L. Huertas, and B. Carrasco, *Calculation of hydrodynamic properties of globular proteins from their atomic-level structure*. Biophys J, 2000. **78**(2): p. 719-30.
266. Ortega, A., D. Amoros, and J. Garcia de la Torre, *Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models*. Biophys J, 2011. **101**(4): p. 892-8.
267. Kikhney, A.G. and D.I. Svergun, *A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins*. FEBS Lett, 2015. **589**(19 Pt A): p. 2570-7.
268. Franke, D., et al., *ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions*. J Appl Crystallogr, 2017. **50**(Pt 4): p. 1212-1225.
269. Guinier, A., *Diffraction of x-rays of very small angles-application to the study of ultramicroscopic phenomenon*. Ann. Phys, 1939. **12**: p. 161-237.
270. Svergun, D.I., Semenyuk, A. V.; Feigin, L.A. , *Small-angle-scattering-data treatment by the regularization method*. Acta Crystallogr. Sect. A Found. Crystallogr., 1988. **44**: p. 244-250.
271. Schneidman-Duhovny, D., et al., *Accurate SAXS profile computation and its assessment by contrast variation experiments*. Biophys J, 2013. **105**(4): p. 962-74.
272. Schneidman-Duhovny, D., et al., *FoXS, FoXSDock and MultiFoXS: Single-state and multi-state structural modeling of proteins and their complexes based on SAXS profiles*. Nucleic Acids Res, 2016. **44**(W1): p. W424-9.
273. Weinkam, P., J. Pons, and A. Sali, *Structure-based model of allostery predicts coupling between distant sites*. Proc Natl Acad Sci U S A, 2012. **109**(13): p. 4875-80.
274. Bernado, P., et al., *Structural characterization of flexible proteins using small-angle X-ray scattering*. J Am Chem Soc, 2007. **129**(17): p. 5656-64.
275. Svergun, D., Barberato, C.; Koch, M.H.J. , *CRY SOL – a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates*. J. Appl. Crystallogr, 1995. **28**: p. 768-773.